

OPTICAL SCIENCES

Jay N. Damask

Polarization Optics in Telecommunications



Springer

Founded by H.K.V. Lotsch

Editor-in-Chief: W.T. Rhodes, Atlanta

Editorial Board: T. Asakura, Sapporo
K.-H. Brenner, Mannheim
T.W. Hänsch, Garching
T. Kamiya, Tokyo
F. Krausz, Vienna and Garching
B. Monemar, Linköping
H. Venghaus, Berlin
H. Weber, Berlin
H. Weinfurter, Munich

Springer Series in OPTICAL SCIENCES

The Springer Series in Optical Sciences, under the leadership of Editor-in-Chief *William T. Rhodes*, Georgia Institute of Technology, USA, provides an expanding selection of research monographs in all major areas of optics: lasers and quantum optics, ultrafast phenomena, optical spectroscopy techniques, optoelectronics, quantum information, information optics, applied laser technology, industrial applications, and other topics of contemporary interest.

With this broad coverage of topics, the series is of use to all research scientists and engineers who need up-to-date reference books.

The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to the Editor-in-Chief or one of the Editors.

Editor-in-Chief

William T. Rhodes

Georgia Institute of Technology
School of Electrical and Computer Engineering
Atlanta, GA 30332-0250, USA
E-mail: bill.rhodes@ece.gatech.edu

Ferenc Krausz

Max-Planck-Institut für Quantenoptik
Hans-Kopfermann-Straße 1
85748 Garching, Germany
E-mail: ferenc.krausz@mpq.mpg.de
and
Institute for Photonics
Gußhausstraße 27/387
1040 Wien, Austria

Editorial Board

Toshimitsu Asakura

Hokkai-Gakuen University
Faculty of Engineering
1-1, Minami-26, Nishi 11, Chuo-ku
Sapporo, Hokkaido 064-0926, Japan
E-mail: asakura@eli.hokkai-s-u.ac.jp

Bo Monemar

Department of Physics
and Measurement Technology
Materials Science Division
Linköping University
58183 Linköping, Sweden
E-mail: bom@ifm.liu.se

Karl-Heinz Brenner

Chair of Optoelectronics
University of Mannheim
Institute of Computer Engineering
B6, 26
68131 Mannheim, Germany
E-mail: brenner@uni-mannheim.de

Herbert Venghaus

Heinrich-Hertz-Institut
für Nachrichtentechnik Berlin GmbH
Einsteinufer 37
10587 Berlin, Germany
E-mail: venghaus@hhi.de

Theodor W. Hänsch

Max-Planck-Institut für Quantenoptik
Hans-Kopfermann-Straße 1
85748 Garching, Germany
E-mail: t.w.haensch@physik.uni-muenchen.de

Horst Weber

Technische Universität Berlin
Optisches Institut
Straße des 17. Juni 135
10623 Berlin, Germany
E-mail: weber@physik.tu-berlin.de

Takeshi Kamiya

Ministry of Education, Culture, Sports
Science and Technology
National Institution for Academic Degrees
3-29-1 Otsuka, Bunkyo-ku
Tokyo 112-0012, Japan
E-mail: kamiyatk@niad.ac.jp

Harald Weinfurter

Ludwig-Maximilians-Universität München
Sektion Physik
Schellingstraße 4/III
80799 München, Germany
E-mail: harald.weinfurter@physik.uni-muenchen.de

Jay N. Damask

Polarization Optics in Telecommunications

With 202 Figures

 Springer

Jay N. Damask
damask@polarization-optics.com

Library of Congress Cataloging-in-Publication Data

Damask, Jay N.

Polarization optics in telecommunications / Jay N. Damask.

p. cm — (Springer series in optical sciences, ISSN 0342-4111 ; v. 101)

Includes bibliographical references and index.

ISBN 0-387-22493-9

1. Optical communication systems. 2. Fiber optics. 3. Polarization (Light) I. Title. II. Series.

TK5103.592.F52.D36 2004

621.382'7—dc22

2004056603

ISBN 0-387-22493-9

ISSN 0342-4111

Printed on acid-free paper.

© 2005 Springer Science+Business Media, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America. (SBA)

9 8 7 6 5 4 3 2 1

SPIN 10949047

springeronline.com

To Diana Castelnovo-Tedesco,
to my Family,
and in loving memory of A. C. Damask

Preface

I have written this book to fill a void between theory and practice, a void that I perceived while conducting my own research and development of components and instruments over the last five years. In the chapters that follow I have pulled materials from the technical and patent literature that are relevant to the understanding and practice of polarization optics in telecommunications, material that is often known by the respective experts in industry and academia but is rarely if ever found in one place. By bringing this material into one monograph, and by applying a single formalism throughout, I hope to create a “base level” upon which future research and development can grow.

Polarization optics in telecommunications is an ever-evolving field. Each year significant advancements are made, punctuated by important discoveries. The references upon which this book is based are only a snap-shot in time. Areas that remain unresolved at the time of publication may very well be clarified in the years to come. Moreover, the focus of the field changes in time: for instance, there have been few passive nonreciprocal component advancements reported in the last few years, but PMD and PDL advancement continues with only modest abatement.

The framework used throughout the monograph is the spin-vector calculus of polarization. The spin-vector calculus as applied to telecommunications optics has long been advocated by N. Frigo, N. Gisin, and J. Gordon. The calculus has its origins in the quantum mechanical description of electron spin and in classical dynamics of rotating bodies. While this calculus may be unfamiliar to the reader, the advantage is its inherent geometric nature and its compact form. Spin-vector calculus abstracts the matrix algebra generally used to describe polarization into a purely vector form. Compound operations are evaluated on the vector field before being resolved onto a local coordinate system. Without exception I have found every derivation in this book shorter, more intuitive, and sometimes surprisingly revealing when using spin-vector calculus. Chapter 2 is entirely dedicated to this formalism. I assure the reader that the time invested learning this material will be rewarding.

The monograph is divided into three logical sections: theory, components, and fiber polarization. The three sections can be treated with some independence. Chapters 1–3 present the basic theory of Maxwell’s equations, polarization, and the classical interaction of light with dielectric media. Next, Chapters 4–7 detail passive optical components, their design, and the building blocks upon which they are based. Special to this section is Chapter 4, which attempts to bridge theory and practice by tabulating known properties of the most commonly used materials and offering practical explanation of simple optical combinations. Lastly, Chapters 8–10 present aspects of polarization-mode dispersion and polarization-dependent loss.

Even though this monograph is entitled, “Polarization Optics in Telecommunications,” the reader should be cognizant of subjects that are missing. Notably absent are, for example, electro-optic effects, used in polarization controllers; liquid-crystal elements, used for switching and attenuation; and interleaver filters, used in wavelength-division multiplexing. These omissions are a measure of my limited experience rather than the fertility of the fields.

I have been fortunate to have a number of experts read various chapters of this book. Their help and dedication have clarified a variety of points and helped prevent mistakes. I am indebted to Dr. C. R. Doerr, Distinguished Member Technical Staff, Bell Laboratories, Lucent Technologies; Dr. N. J. Frigo, Division Manager, AT&T Laboratories; Dr. J. P. Mattia, Co-Founder, Big Bear Networks; Prof. T. E. Murphy, Assistant Professor, University of Maryland, College Park; Dr. K. R. Rochford, Division Chief, Optoelectronics, National Institute of Standards and Technology; Dr. M. Shirasaki, Co-Founder and Chief Scientist, Arasor; and Dr. P. Westbrook, Technical Manager, Photonics Device Research, OFS Labs. Complementing my Readers, Dr. P. A. Williams of the National Institute of Standards and Technology has carefully answered my questions throughout the entire writing of this book – I am pleased to acknowledge his great support.

I have also contacted many other experts when I needed clarification on particular topics. I would like to thank Mr. M. Alexandrovich, Prof. H. Ammari, Dr. N. Bergano, Mr. A. Boschi, Dr. S. Evangelides, Prof. A. Eyal, Dr. V. Fratello, Prof. D. Hagen, Dr. D. Harris, Dr. G. Harvey, Prof. E. Ippen, Dr. P. Leo, Dr. J. Livas, Dr. C. Madsen, Prof. A. Meccozi, Prof. C. Menyuk, Mr. P. Myers, Dr. J. Nagel, Dr. K. Nordsieck, Dr. B. Nyman, Dr. C. Poole, Dr. G. Shtengel, Mr. G. Simer, and Dr. P. Xie.

While I am indebted to these contributors, all mistakes are my responsibility alone. You can contact me at damask@polarization-optics.com and I look forward to receiving your feedback.

I wish to thank The MathWorks Company, and especially C. Esposito, for generous support through the MathWork’s Authors’ program. Many of the code pieces I used to generate the figures will be available courtesy of the MathWorks at www.mathworks.com.

The people at Springer, New York, have generously given their time and encouragement over the last eighteen months. In particular, I am indebted to

my editor Dr. H. Koelsch, and to F. Ganz and M. Mitchell. Their professionalism and expertise has made this project a pleasure for me.

I wish also to thank the library services at the Massachusetts Institute of Technology. The M.I.T. technical library is a national resource and is second to none. The professional staff and on-line databases have helped me find original references of all sorts.

I wish to remember M.I.T. Institute Professor Hermann A. Haus, who, over a decade, supported my pursuit into the beauties of optics.

Finally, I am indebted to my family, especially Mary and John, and to my friends, who encouraged me throughout this project. Special acknowledgement goes to my wife D. C.-T., without whose unwavering support this book would not have been written.

New York City
July 2004

Jay N. Damask

Contents

1	Vectorial Propagation of Light	1
1.1	Maxwell's Equations and Free-Space Solutions	2
1.2	The Vector and Scalar Potentials	8
1.3	Time-Harmonic Solutions	10
1.4	Classical Description of Polarization	12
1.4.1	Stokes Vectors, Jones and Muller Matrices	17
1.4.2	The Poincaré Sphere	20
1.5	Partial Polarization	22
1.5.1	Coherently Polarized Waves	24
1.5.2	Incoherently Depolarized Waves	28
1.5.3	Pseudo-Depolarized Waves	31
1.5.4	A Heterogeneous Ray Bundle	33
	References	36
2	The Spin-Vector Calculus of Polarization	37
2.1	Motivation	37
2.2	Vectors, Length, and Direction	39
2.2.1	Bra and Ket Vectors	39
2.2.2	Length and Inner Products	41
2.2.3	Projectors and Outer Products	42
2.2.4	Orthonormal Basis	43
2.3	General Vector Transformations	44
2.3.1	Operator Relations	44
2.4	Eigenstates, Hermitian and Unitary Operators	46
2.4.1	Hermitian Operators	47
2.4.2	Unitary Operators	48
2.4.3	Connection between Hermitian and Unitary Matrices	49
2.4.4	Similarity Transforms	49
2.4.5	Construction of General Unitary Matrix	50
2.4.6	Group Properties of $SU(2)$	51
2.5	Vectors Cast in Jones and Stokes Spaces	52

2.5.1	Complete Measurement of the Polarization Ellipse	52
2.5.2	Pauli Spin Matrices	54
2.5.3	The Pauli Spin Vector	55
2.5.4	Spin-Vector Identities	56
2.5.5	Conservation of Length	58
2.5.6	Orthogonal Polarization States	59
2.5.7	Non-Orthogonal Polarization States	60
2.5.8	Pauli Spin Operators	61
2.6	Equivalent Unitary Transformations	63
2.6.1	Group Properties of $SU(2)$ and $O(3)$	65
2.6.2	Matrix Entries of R in a Fixed Coordinate System	66
2.6.3	Vector Expression of R in a Local Coordinate System	67
2.6.4	Select Vector Identities	70
2.6.5	Euler Rotations	71
2.6.6	Some Relevant Transformation Applications	72
	References	78
3	Interaction of Light and Dielectric Media	79
3.1	Introduction of Media Terms into Maxwell's Equations	80
3.2	Constitutive Relation Tensors	85
3.3	The kDB System	87
3.4	The Lorentz Force	90
3.5	Isotropic Materials	90
3.5.1	Permittivity of Isotropic Materials	91
3.5.2	Propagation in Isotropic Materials	94
3.5.3	Refraction at an Interface	96
3.5.4	Reflection and Transmission for TE Waves	96
3.5.5	Reflection and Transmission for TM Waves	99
3.5.6	Total Internal Reflection	101
3.6	Birefringent Materials	105
3.6.1	Propagation in Uniaxial Materials	106
3.6.2	Refraction at an Interface	112
3.6.3	Total Internal Reflection	117
3.6.4	Polarization Transformation	120
3.7	Gyrotropic Materials	122
3.7.1	Magnetic Material Classes	123
3.7.2	Permittivity of Diamagnetic Materials	124
3.7.3	Propagation in Gyrotropic Materials	126
3.7.4	Faraday Rotation	129
3.7.5	The Verdet Constant	132
3.7.6	Faraday Rotation in Ferrous Materials	133
3.8	Optically Active Materials	135
3.8.1	Propagation in Bi-Isotropic Media	138
	References	142

4	Elements and Basic Combinations	143
4.1	Wavelength-Division Multiplexed Frequency Grid	143
4.2	Properties of Select Materials	146
4.2.1	Isotropic Glass Materials	146
4.2.2	Birefringent Crystals	147
4.2.3	Iron Garnets	150
4.2.4	Packaging Alloys	153
4.3	Fabry-Perot and Gires-Tournois Interferometers	154
4.3.1	Fabry-Perot Response	157
4.3.2	Gires-Tournois Response	161
4.4	Temperature Dependence of Select Birefringent Crystals	163
4.4.1	Experimental Setup	163
4.4.2	Quadratic Temperature-Dependence Model	166
4.4.3	Association of Resonant Peak Shift With Temperature Coefficients	167
4.4.4	Group Index and Thermal-Optic Coefficients	168
4.4.5	Passive Temperature Compensation	170
4.5	Compound Crystals For Off-Axis Delay	173
4.6	Polarization Retarders	178
4.6.1	Half-Wave and Quarter-Wave Waveplates	179
4.6.2	Birefringent Waveplate Technologies	182
4.6.3	Waveplate Combinations	184
4.6.4	Elementary Polarization Control	191
4.6.5	TIR Polarization Retarders	196
4.7	Single and Compound Prisms	198
4.7.1	Wollaston and Rochon Prisms	199
4.7.2	Kaifa Prism	202
4.7.3	Shirasaki Prism	204
	References	208
5	Collimator Technologies	211
5.1	Collimator Assemblies	213
5.2	Gaussian Optics	219
5.2.1	q Transformation and ABCD Matrices	224
5.2.2	ABCD Ray Tracing	227
5.2.3	Action of a Single Lens	228
5.2.4	Action of a GRIN Lens	230
5.2.5	Some Limitations of the ABCD Matrix	232
5.3	Select Collimators Analyzed with the ABCD Matrix	234
5.4	Fiber-to-Fiber Coupling by a Lens Pair	239
5.4.1	Coupling Coefficients	242
	References	245

6	Isolators	247
6.1	Polarizing Isolator	247
6.2	Comparison of Lens Systems	252
6.3	Deflection-Type Isolators	254
6.4	Displacement-Type Isolators	259
6.5	Two-Stage Isolators	263
6.6	PMD-Compensated Isolators	266
	References	271
7	Circulators	273
7.1	Polarizing Circulator	274
7.2	Historical Development	277
7.3	Displacement Circulators	279
7.4	Deflection Circulators	284
7.5	Summary	294
	References	295
8	Properties of PDL and PMD	297
8.1	Polarization-Dependent Loss	298
8.1.1	Definitions	299
8.1.2	Change of Polarization State	304
8.1.3	Repolarization	306
8.1.4	PDL Evolution Equations	308
8.2	Polarization-Mode Dispersion	312
8.2.1	A PMD Primer	313
8.2.2	Fundamental Derivations	327
8.2.3	Connection Between Jones and Stokes Space	330
8.2.4	Concatenation Rules for PMD	333
8.2.5	PMD Evolution Equations	338
8.2.6	Time-Domain Representation	342
8.2.7	Fourier Analysis of the DGD Spectrum	364
8.3	Combined Effects of PMD and PDL	371
8.3.1	Frequency-Dependence of the Polarization State	372
8.3.2	Non-Orthogonality of PSP's	374
8.3.3	PMD and PDL Evolution Equations	376
8.3.4	Separation of PMD and PDL	378
	References	381
9	Statistical Properties of Polarization in Fiber	385
9.1	Polarization Evolution Model	388
9.1.1	Random Birefringent Orientation	389
9.1.2	Random Component Birefringence	391
9.2	Polarization Diffusion	392
9.3	RMS Differential-Group Delay Evolution	397
9.4	PMD Statistics	399

9.4.1	Probability Densities	401
9.4.2	Autocorrelation Functions	408
9.4.3	Mean-DGD Measurement Uncertainty	414
9.4.4	Discrete Waveplate Model	417
9.4.5	Karhunen-Loève Expansion of Brownian Motion	419
9.5	PDL Statistics	422
	References	425
10	Review of Polarization Test and Measurement	429
10.1	SOP Measurement	430
10.2	PDL Measurement	432
10.3	PMD Measurement	436
10.3.1	Mean DGD Measurement	438
10.3.2	PMD Vector Measurement	440
10.3.3	Polarization OTDR	450
10.4	Programmable PMD Sources	451
10.4.1	Sources of DGD and Depolarization	454
10.4.2	ECHO Sources	463
10.5	Receiver Performance Validation	478
	References	483
A	Addition of Multiple Coherent Waves	491
B	Select Magnetic Field Profiles	493
	References	496
C	Efficient Calculation of PMD Spectra	497
D	Multidimensional Gaussian Deviates	505
	Index	509

Vectorial Propagation of Light

Maxwell's equations are the basis of all optical studies. In vacuum the equations can be stripped to a pure form where the wave motion is most easily described. Moreover, as the equations in vacuum are linear, each Fourier component of a wave can be individually studied and subsequently superimposed to construct a composite wavefront or ray bundle. When the electromagnetic wave propagates through media, additional terms are added to Maxwell's equations to account for the interaction. These terms come in as constitutive laws of the media. Constitutive laws can encompass lossy, charged, dielectric, nonlinear, or relativistic media. There is almost no end to the studies on optical interactions already undertaken over the last several hundred years.

The purpose of this chapter and that of Chapters 2 and 3 is to derive the necessary governing equations for studies of birefringent media, birefringent components, and birefringent effects in optical fiber. This chapter exclusively deals with Maxwell's equations in vacuum. The classical description of polarization motion and the degree of polarization is emphasized. Chapter 2 presents a modern description of polarization that adopts well-developed mathematical formalisms from quantum mechanics to polarization studies. Chapter 3 adds interaction terms to Maxwell's equations to describe optical propagation through birefringent linear dielectrics.

1.1 Maxwell's Equations and Free-Space Solutions

The four vectorial Maxwell's equations are

$$\text{Faraday's law:} \quad \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mu_o \mathbf{H}(\mathbf{r}, t) - \frac{\partial}{\partial t} \mu_o \mathbf{M}(\mathbf{r}, t)$$

$$\text{Ampère's law:} \quad \nabla \times \mathbf{H}(\mathbf{r}, t) = \varepsilon_o \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t) + \frac{\partial}{\partial t} \mathbf{P}(\mathbf{r}, t) + \mathbf{J}(\mathbf{r}, t)$$

$$\text{Gauss's electric law:} \quad \nabla \cdot \varepsilon_o \mathbf{E}(\mathbf{r}, t) = -\nabla \cdot \mathbf{P}(\mathbf{r}, t) + \rho_f(\mathbf{r}, t)$$

$$\text{Gauss's magnetic law:} \quad \nabla \cdot \mu_o \mathbf{H}(\mathbf{r}, t) = -\nabla \cdot \mu_o \mathbf{M}(\mathbf{r}, t)$$

where the vector quantities $\mathbf{E}$, $\mathbf{H}$, $\mathbf{P}$, $\mathbf{M}$, $\mathbf{J}$, and ρ_f are real functions of time t and the three-dimensional spatial vector $\mathbf{r}$. These vector quantities are

$$\mathbf{E}(\mathbf{r}, t) : \quad \text{electric field strength} \quad (\text{V/m})$$

$$\mathbf{H}(\mathbf{r}, t) : \quad \text{magnetic field strength} \quad (\text{A/m})$$

$$\mathbf{P}(\mathbf{r}, t) : \quad \text{polarization density} \quad (\text{C/m}^2)$$

$$\mathbf{M}(\mathbf{r}, t) : \quad \text{magnetization density} \quad (\text{A/m})$$

$$\mathbf{J}(\mathbf{r}, t) : \quad \text{current density} \quad (\text{A/m}^3)$$

$$\rho_f(\mathbf{r}, t) : \quad \text{electric charge density} \quad (\text{C/m}^3)$$

where V is volts, A is amperes, and C is coulombs. The free electric charge density ρ_f is distinguished from the bound charge density ρ_b as the bound density is the generator of the polarization density vector $\mathbf{P}$.

The physical constants ε_o and μ_o are the permittivity and permeability of vacuum, respectively. The values and units are [8]

$$\varepsilon_o \simeq 8.854187817 \times 10^{-12} \quad (\text{F/m})$$

$$\mu_o = 4\pi \times 10^{-7} \quad (\text{H/m})$$

where F is Farads and H is Henries.

Maxwell's equations completely describe the propagation and spatial extent of electromagnetic waves in free-space and in any medium. Faraday's law states that the curl of the electric field is generated by the temporal change of the magnetic field and the magnetization density vector. Ampère's law states that the curl of the magnetic field is generated by the temporal change of the electric field and the polarization density vector, as well as by currents of

charged particles. Gauss's two laws govern the divergence of the electric and magnetic fields. The divergence is zero except in the presence of dipoles and electric charges.

It is customary when considering a restricted class of problems to eliminate various non-essential terms from the equations. As this text is predominantly focused on passive birefringent optical components, including interaction with fixed electric and magnetic fields, the current density $\mathbf{J}(\mathbf{r}, t)$, and the free electric charge density $\rho_f(\mathbf{r}, t)$ are set to zero. The reduced equations are

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\mu_o \frac{\partial}{\partial t} \mathbf{H}(\mathbf{r}, t) - \mu_o \frac{\partial}{\partial t} \mathbf{M}(\mathbf{r}, t) , \quad (1.1.1)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \varepsilon_o \frac{\partial}{\partial t} \mathbf{E}(\mathbf{r}, t) + \frac{\partial}{\partial t} \mathbf{P}(\mathbf{r}, t) , \quad (1.1.2)$$

$$\nabla \cdot \varepsilon_o \mathbf{E}(\mathbf{r}, t) = -\nabla \cdot \mathbf{P}(\mathbf{r}, t) , \quad (1.1.3)$$

$$\nabla \cdot \mathbf{H}(\mathbf{r}, t) = -\nabla \cdot \mathbf{M}(\mathbf{r}, t) \quad (1.1.4)$$

The field terms $\mathbf{E}$ and $\mathbf{H}$ are the two complementary components of an electromagnetic wave. The polarization and magnetization density vectors $\mathbf{P}$ and $\mathbf{M}$, respectively, are the means to describe the interaction of the electromagnetic field with matter. The density vectors $\mathbf{P}$ and $\mathbf{M}$ are related to the field quantities $\mathbf{E}$ and $\mathbf{H}$ by constitutive relations. The constitutive relations for various dielectric materials will be presented in detail in Chapter 3. This chapter details the most simple solutions to Maxwell's equations, the field solutions in vacuum.

In a vacuum, vectors $\mathbf{P}$ and $\mathbf{M}$ are zero. The wave equation for the electric field is derived by taking the curl of Faraday's law, substituting in Ampère's law, and reordering the temporal derivative $\frac{\partial}{\partial t}$ since it commutes with ∇ . The wave equation for the magnetic field is similarly found. The electric-field wave equation is

$$\nabla \times \nabla \times \mathbf{E} = -\mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \mathbf{E} \quad (1.1.5)$$

Application of the vector identity $\nabla \times \nabla \times = \nabla (\nabla \cdot) - \nabla^2 ()$, and recognition that Gauss' law (1.1.3) dictates zero electric-field divergence in the absence of a fixed charge density, (1.1.5) simplifies to the Helmholtz wave equation:

$$\nabla^2 \mathbf{E} = \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \mathbf{E} \quad (1.1.6)$$

The Helmholtz equation relates the spatial curvature of the electric field $\mathbf{E}(\mathbf{r}, t)$ to its temporal second derivative, the factor of proportionality being $\mu_o \varepsilon_o$. The wave equation is otherwise invariant to spatial and temporal translation, spatial rotation, time reversal, and coordinate system selection. Moreover, the wave equation is linear in that

$$\nabla^2 (\mathbf{E}_1 + \mathbf{E}_2) = \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} (\mathbf{E}_1 + \mathbf{E}_2) \quad (1.1.7)$$

The linear property of the wave equation allows arbitrarily complex field distributions $\mathbf{E}$ to be constructed by Fourier synthesis or the method of superposition.

A monochromatic solution to (1.1.6) is

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_o \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (1.1.8)$$

where $\mathbf{E}_o$, $\mathbf{k}$, and $\mathbf{r}$ are three-dimensional real-valued vectors and ω is the radial oscillatory frequency of the wave. $\mathbf{E}_o$ is the field amplitude at time and distance zero. $\mathbf{k}$ is the propagation vector of the field. The magnitude of $\mathbf{k}$, having units of inverse length, is the wavenumber $k = |\mathbf{k}|$. The monochromatic wave (1.1.8) is a travelling plane wave that propagates in the direction of $\mathbf{k}$ and oscillates at frequency ω .

When an underlying coordinate system is chosen so that the propagation direction of the wave $\mathbf{k}$ is coincident with a coordinate axis $\hat{r}$, i.e. $\mathbf{k} \parallel \hat{r}$, the spatial argument of (1.1.8) reduces to $\mathbf{k} \cdot \mathbf{r} = kr$. The monochromatic solution simplifies to $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_o \cos(\omega t - kr)$. This is the equation of a plane wave whose phase fronts are constant in the plane perpendicular to $\hat{r}$ and whose amplitude is likewise constant in that plane. Picking a fixed phase position along the wavefront as it propagates along $\hat{r}$, $\omega t - kr = \text{constant}$, it is found that the phase front travels at phase velocity $v_{\text{ph}} = \omega/k$. The wavelength λ , as defined by the length along $\hat{r}$ between two adjacent field maxima, is $\lambda = 2\pi/k$.

Substitution of the monochromatic plane wave solution (1.1.8) into the wave equation (1.1.6) yields the dispersion relation that relates the wavenumber to the radial frequency:

$$k = \omega \sqrt{\mu_o \varepsilon_o} \quad (1.1.9)$$

As wave equation (1.1.6) is written for vacuum, the electromagnetic wavefront velocity is the speed of light, c . Using the dispersion relation, the speed of light is related to the free-space permittivity and permeability as

$$c = 1/\sqrt{\mu_o \varepsilon_o} \quad (1.1.10)$$

The speed of light in vacuum is [8]

$$c \simeq 299,792,458 \text{ (m/s)}$$

The wavenumber is therefore related to frequency and the speed of light via $k = \omega/c$.

The monochromatic wave (1.1.8) can be resolved into cartesian coordinates as follows. The field amplitude vector is resolved into three scalar components $\mathbf{E}_o = [E_x \ E_y \ E_z]^T$; the coordinate vector $\mathbf{r}$ is resolved as

$$\mathbf{r} = \hat{x}x + \hat{y}y + \hat{z}z \quad (1.1.11)$$

the wave vector $\mathbf{k}$ is resolved as

$$\mathbf{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z \quad (1.1.12)$$

A particular vector component of (1.1.8) takes associated elements from $\mathbf{E}$ and $\mathbf{k} \cdot \mathbf{r}$, e.g. $E(x, t) = E_x \cos(\omega t - k_x x)$. From (1.1.12), the wavenumber in cartesian coordinates is

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2} \quad (1.1.13)$$

The monochromatic electric-field solution (1.1.8) has a magnetic field counterpart. Introduction of (1.1.8) into Faraday's law and solving for $\mathbf{H}$ by taking the time integral yields the magnetic field monochromatic solution

$$\mathbf{H}(\mathbf{r}, t) = \sqrt{\frac{\varepsilon_o}{\mu_o}} \hat{k} \times \mathbf{E}_o \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (1.1.14)$$

where the value of the wavenumber k has been pulled through by writing $\mathbf{k} = k\hat{k}$ and where $\hat{k}$ is a unit vector pointing in the direction of $\mathbf{k}$. The magnetic field has the same spatial and temporal dependence as the associated electric field. The scalar constant that relates the two field amplitudes is $\sqrt{\varepsilon_o/\mu_o}$. This physical constant is called the characteristic admittance of vacuum. The characteristic impedance, the inverse of the admittance, is approximately [8]

$$\sqrt{\frac{\mu_o}{\varepsilon_o}} \simeq 376.730313461 \text{ (ohms)}$$

Substitution of the field equations (1.1.8) and (1.1.14) into Maxwell's equations (1.1.1-1.1.4) for vacuum yields

$$\mathbf{k} \times \mathbf{E} = \omega \mu_o \mathbf{H} \quad (1.1.15a)$$

$$\mathbf{k} \times \mathbf{H} = -\omega \varepsilon_o \mathbf{E} \quad (1.1.15b)$$

$$\mathbf{k} \cdot \mathbf{E} = 0 \quad (1.1.15c)$$

$$\mathbf{k} \cdot \mathbf{H} = 0 \quad (1.1.15d)$$

These equations show the relation of the electric and magnetic field oscillations with respect to one another and with respect to the propagation direction $\mathbf{k}$. The divergence equations for the electric and magnetic fields (1.1.15c,d) show that there are no field components in the direction of propagation. That is, the longitudinal field components are zero; only transverse components exist. Both the electric and magnetic field oscillations are therefore perpendicular to $\mathbf{k}$. Moreover, the electric and magnetic field oscillations are mutually perpendicular. Calculation of $\mathbf{E} \cdot \mathbf{H}$ via (1.1.15a,b) results in

$$\mathbf{E} \cdot \mathbf{H} = -\frac{1}{\omega^2 \mu_o \varepsilon_o} (\mathbf{k} \times \mathbf{H}) \cdot (\mathbf{k} \times \mathbf{E})$$

Application of the vector relation $\mathbf{a} \times \mathbf{b} \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ shows that $\mathbf{E} \cdot \mathbf{H} = 0$.

Combination of Faraday's and Ampère's laws has led to the wave equation (1.1.6), which in turn yielded a monochromatic plane-wave solution for both field components (1.1.8) and (1.1.14). Substitution of these field expressions into Maxwell's equations for vacuum leads to the conclusion that the vectors $(\mathbf{E}, \mathbf{H}, \mathbf{k})$ are mutually perpendicular. What remains is the calculation of energy flow of the propagating electromagnetic wave.

Poynting's theorem shows explicitly that conservation of energy is an immediate result of Maxwell's equations. The theorem states that the electromagnetic power flow into a volume must equal the rate of increase of stored electric and magnetic energy plus the total power dissipated. To arrive at the conservation equation, take the dot product of $\mathbf{H}$ with Faraday's law and the dot product of $\mathbf{E}$ with Ampère's law, and use the vector identity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot \mathbf{a} \times \mathbf{b} - \mathbf{b} \cdot \mathbf{a} \times \mathbf{c}$. Poynting's energy conservation equation is

$$\begin{aligned} \nabla \cdot (\mathbf{E} \times \mathbf{H}) + \frac{\partial}{\partial t} \left(\frac{1}{2} \varepsilon_o \mathbf{E} \cdot \mathbf{E} \right) + \frac{\partial}{\partial t} \left(\frac{1}{2} \mu_o \mathbf{H} \cdot \mathbf{H} \right) + \\ \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mu_o \mathbf{M}}{\partial t} + \mathbf{E} \cdot \mathbf{J} = 0 \quad (1.1.16) \end{aligned}$$

The Poynting theorem introduces a new vector quantity: $\mathbf{E} \times \mathbf{H}$. This is called the Poynting vector and represents the electromagnetic power flow density and has units of (W/m²). It is customary to represent the Poynting vector by the symbol $\mathbf{S}$:

$$\mathbf{S}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) \times \mathbf{H}(\mathbf{r}, t) \quad (1.1.17)$$

The direction of $\mathbf{S}$ is the direction of power flow. The power flow direction is always orthogonal to both the $\mathbf{E}$ and $\mathbf{H}$ fields. Recalling Gauss' integral theorem,

$$\int_V \nabla \cdot \mathbf{F} dV = \oint_S \mathbf{F} \cdot d\mathbf{a}$$

the divergence of $\mathbf{F}$ enclosed by volume V equals the power flow through surface S out of the volume. Accordingly, $\nabla \cdot \mathbf{S}$ represents the power flow out of a differential volume. This power flow is balanced by the increase of stored electromagnetic energy W and by the power dissipated P_d . Symbolically [2],

$$\nabla \cdot \mathbf{S} + \frac{\partial W}{\partial t} + P_d = 0$$

The energy stored in the system is recoverable; the stored energy is reactive rather than resistive. The power dissipated is non-recoverable. In terms of the conservation equation, energy that can be grouped after the $\partial/\partial t$ operator is stored while the fixed power is dissipated. As an example, consider a volume V through which electric energy $W_e = 1/2 \varepsilon_o \mathbf{E} \cdot \mathbf{E}$ flows. Denote the temporal profile as $W_e(t) = W_o f(t)$ where $f(t)$ is a positive, bounded scalar function of time and W_o is the maximum electric energy. The profile function is zero at $t = \pm\infty$. The time-integrated reactive power is

$$\int_{-\infty}^{+\infty} \frac{\partial}{\partial t} (W_o f(t)) dt = 0$$

Integration over all time shows that no net power was left in volume V . Shown another way [1], for any intermediate time t_o , the energy into the volume V is

$$\int_{-\infty}^{t_o} \frac{\partial}{\partial t} (W_o f(t)) dt = +W_o f(t_o)$$

After t_o , the energy into the volume V is

$$\int_{t_o}^{\infty} \frac{\partial}{\partial t} (W_o f(t)) dt = -W_o f(t_o)$$

The energy that flows into V up to time t_o is fully recovered as $t \rightarrow +\infty$. However, consider the $\mathbf{E} \cdot \mathbf{J}$ term. Using Ohm's law relating current density to electric field, $\mathbf{J} = \sigma \mathbf{E}$, where σ is the charge density, the power dissipated is

$$\int_{-\infty}^{\infty} \sigma E^2(t) dt = \sigma E_o^2 \mathcal{I}_p$$

where $\mathcal{I}_p$ is the integral of the square electric field $E^2(t)$ over all time and E_o is the maximum field amplitude assuming a bounded field-amplitude time profile. Only if $E(t) = 0$ for all time for finite σ will the dissipated power vanish, but this is the trivial case.

With this understanding of what constitutes stored energy and dissipated power, the stored energy present in Poynting's theorem is identified with

$$W = W_e + W_m = \frac{1}{2} \varepsilon_o \mathbf{E} \cdot \mathbf{E} + \frac{1}{2} \mu_o \mathbf{H} \cdot \mathbf{H} \quad (1.1.18)$$

and the power dissipated is identified with

$$P_d = \mathbf{E} \cdot \mathbf{J} \quad (1.1.19)$$

This leaves the remaining terms $\mathbf{E} \cdot \partial \mathbf{P} / \partial t$ and $\mathbf{H} \cdot \partial \mu_o \mathbf{M} / \partial t$ open to interpretation as energy storage terms or power dissipative terms. In general these two terms can be either; the particulars depend on the nature of the matter with which the electromagnetic field interacts. For example, in the case of linear dielectrics, $\mathbf{P} = \varepsilon_o \chi_e \mathbf{E}$, the dipole density follows the electric field instantaneously. The change of energy of the polarization density is then

$$\mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} = \frac{\partial}{\partial t} \left(\frac{1}{2} \varepsilon_o \chi_e \mathbf{E} \cdot \mathbf{E} \right)$$

where the energy stored in the polarization density is clearly reactive. If, on the other hand, the dipole density exhibits a delayed reaction to the electric field, as can be the case in highly resistive media, then one could write $d\mathbf{P}/dt = a\mathbf{E}$ where a is a scaling parameter [2]. Then,

$$\mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} = a \mathbf{E} \cdot \mathbf{E}$$

and the system is dissipative.

Earlier in this section the general plane-wave monochromatic field solutions in vacuum were found for both the electric and magnetic fields. The power flow density is found by $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. Taking the cross of (1.1.8) and (1.1.14) yields

$$\mathbf{S}(\mathbf{r}, t) = \hat{k} \sqrt{\frac{\varepsilon_o}{\mu_o}} E_o^2 \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (1.1.20)$$

The time average of the Poynting vector yields the average power flow of the electromagnetic field:

$$\langle \mathbf{S}(\mathbf{r}, t) \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{S}(\mathbf{r}, t) d(\omega t) = \hat{k} \frac{1}{2} \sqrt{\frac{\varepsilon_o}{\mu_o}} E_o^2 \quad (1.1.21)$$

The time-average power flow of the electromagnetic field in vacuum is along the $\hat{k}$ direction, where $\hat{k}$ is perpendicular to planes of constant phase along the wave front. In the following chapters, dielectric anisotropy is introduced. The anisotropy will, in general, break the apparent identity that $\mathbf{S}$ and $\mathbf{k}$ run parallel to one another and instead induce the power flow and wave-front propagation directions to diverge.

1.2 The Vector and Scalar Potentials

In the absence of currents, free charges, and electric and magnetic dipoles, Maxwell's equations reduce to

$$\nabla \times \mathbf{E} = -\mu_o \frac{\partial}{\partial t} \mathbf{H} \quad (1.2.1a)$$

$$\nabla \cdot \mu_o \mathbf{H} = 0 \quad (1.2.1b)$$

$$\nabla \times \mathbf{H} = \varepsilon_o \frac{\partial}{\partial t} \mathbf{E} \quad (1.2.1c)$$

$$\nabla \cdot \varepsilon_o \mathbf{E} = 0 \quad (1.2.1d)$$

Under these circumstances, the magnetic and electric fields are solenoidal (having zero divergence). It is appealing to find the class of fields that *a priori* guarantee the solenoidal nature. Note the following vectors identities:

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0 \quad (1.2.2a)$$

$$\nabla \times (\nabla \psi) = 0 \quad (1.2.2b)$$

that is, the divergence of an arbitrary field curl $\nabla \times \mathbf{F}$ is solenoidal and the curl of an arbitrary potential gradient $\nabla \psi$ is irrotational.

The solenoidal nature of $\mu_o \mathbf{H}$ is guaranteed by equating it with the curl of the vector potential $\mathbf{A}$:

$$\mu_o \mathbf{H} = \nabla \times \mathbf{A} \quad (1.2.3)$$

Substitution of (1.2.3) into (1.2.1a) yields

$$\nabla \times \left(\mathbf{E} + \frac{\partial}{\partial t} \mathbf{A} \right) = 0 \quad (1.2.4)$$

Following (1.2.2b), (1.2.4) is guaranteed by defining $\mathbf{E}$ as

$$\mathbf{E} = -\nabla \Phi - \frac{\partial}{\partial t} \mathbf{A} \quad (1.2.5)$$

where Φ is the scalar potential. Maxwell's equations (1.2.1a,b) are guaranteed to be satisfied when $\mathbf{E}$ and $\mathbf{H}$ are expressed in terms of the vector potential $\mathbf{A}$ and scalar potential Φ as above. That said, $\mathbf{A}$ is not yet uniquely determined, as any field is defined by both its curl and divergence. The divergence of $\mathbf{A}$ has not yet been established. Without this, a shift of the vector potential by an arbitrary gradient, e.g. $\mathbf{A}' = \mathbf{A} + \nabla \psi$, would not change either $\mathbf{E}$ nor $\mathbf{H}$ but would indeed change Φ .

The divergence of $\mathbf{A}$ must be set with an eye toward guaranteeing the solutions to the remaining Maxwell's equations (1.2.1c,d). Substitution of (1.2.3, 1.2.5) into (1.2.1c) gives

$$\nabla \times (\nabla \times \mathbf{A}) = \mu_o \varepsilon_o \frac{\partial}{\partial t} \left(-\nabla \Phi - \frac{\partial}{\partial t} \mathbf{A} \right) \quad (1.2.6)$$

Expanding the double-curl on the left side and rearranging terms makes

$$\nabla^2 \mathbf{A} = \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \mathbf{A} + \nabla \left(\nabla \cdot \mathbf{A} + \mu_o \varepsilon_o \frac{\partial}{\partial t} \Phi \right) \quad (1.2.7)$$

The selection of the vector potential divergence is arbitrary since $\mathbf{E}$ and $\mathbf{H}$ are invariant. Therefore the most convenient choice is suitable. Accordingly, a wave equation for the vector potential can be established given the definition

$$\nabla \cdot \mathbf{A} + \mu_o \varepsilon_o \frac{\partial}{\partial t} \Phi = 0 \quad (1.2.8)$$

This choice is called the Lorentz gauge. This gauge in turn is used to generate a wave equation for the scalar potential through substitution into (1.2.1d). Together the wave equations are

$$\nabla^2 \mathbf{A} = \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \mathbf{A} \quad (1.2.9a)$$

$$\nabla^2 \Phi = \mu_o \varepsilon_o \frac{\partial^2}{\partial t^2} \Phi \quad (1.2.9b)$$

In summary, the vector and scalar potentials are self-consistent fields that are constructed to satisfy all of Maxwell's equations by definition. The divergence and curl of the vector potential is completely specified, through which the link to the scalar potential is defined. The vector and scalar potentials provide an alternative means to find solutions to Maxwell's equations. In particular, plane wave solutions exemplified by (1.1.8) are highly convenient when the electromagnetic source is modelled infinitely far away and any dielectric or magnetic media are piece-wise uniform; Fourier techniques can be used to assemble a ray bundle that satisfies some boundary condition. In contrast, point sources generate nonuniform field patterns that cannot be modelled by plane waves. The vector and scalar potentials are necessary to find the requisite field solutions. As a particularly relevant example, Gaussian beam optics grants the adiabatic expansion of a ray bundle as fundamental. In this paraxial limit, the eigen-waves have a spherical phase curvature that is not present in a plane wave. In practice, which formalism is used, field solutions or vector potential solutions, is determined by the problem and the required degree of accuracy.

1.3 Time-Harmonic Solutions

The above developments of Maxwell's equations, monochromatic field solutions, and Poynting's theorem were performed in vector notation with only passing reference to an underlying coordinate system. Pure vector notation provides the most compact form of the equations, provides for direct comparison of the vector quantities, and allows for resolution onto any convenient coordinate system. In an analogous manner, complex exponential notation is like vector notation because there is no *a priori* selection to an underlying time reference. The use of cosine solutions in the previous section is certainly acceptable, but choice of (sin, cos) requires selecting an underlying time reference from the beginning. To keep with real-valued functions at this point will lead to unnecessary analytic complexity when adding phases or multiplying frequencies. The equations and solutions of the preceding section will be recast into complex exponential notation to simplify the analytics.

One problem with complex exponential notation is that there is no customary sign of the argument. Physics texts usually use $\exp(-i\omega t)$, while engineering text usually use $\exp(j\omega t)$. Either selection is fine, as long as the derivations, particularly those regarding polarization, are consistent. This text chooses to use $\exp(j\omega t)$.

The operators $\Re$ and $\Im$ are used to translate between real functions and complex exponential functions. For a complex exponential $z = \exp(j\phi)$, the following relations are defined:

$$\Re\{z\} = \frac{z + z^*}{2}, \quad \Im\{z\} = \frac{z - z^*}{2} \quad (1.3.1)$$

and

$$z = \Re\{z\} + j\Im\{z\} \quad (1.3.2)$$

where z^* is the complex conjugate of z .

The real-valued electric field is defined using complex exponential notation as

$$\mathbf{E}(\mathbf{r}, t) = \Re \left\{ \mathbf{E} e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})} \right\} \quad (1.3.3)$$

where $\mathbf{E}$ is a complex vector. Moreover, $\mathbf{E}$ is written rather than $\mathbf{E}_o$ only for compactness of notation, but it is recognized that $\mathbf{E}$ is evaluated at $t = 0$ and $r = 0$. The real part of (1.3.3) is the same as (1.1.8). The remaining field, dipole, and current terms in Maxwell's equations undergo a similar substitution. Operations ∇ and $\frac{\partial}{\partial t}$ on the complex field undergo the following mapping:

$$\begin{aligned} \nabla &\rightarrow -j\mathbf{k} \\ \frac{\partial}{\partial t} &\rightarrow j\omega \end{aligned}$$

Substitution of (1.3.3) and like terms into Faraday's law yields [7]

$$\Re \left\{ -j \left[\mathbf{k} \times \mathbf{E} - \omega (\mu_o \mathbf{H} + \mu_o \mathbf{M}) \right] e^{j(\omega t - \mathbf{k} \cdot \mathbf{r})} \right\} = 0$$

This equation must hold true for all time and position. As the real part of the exponential term can take any value between $-1 \leq \Re(\exp(j\phi)) \leq 1$, the remaining expression must equal zero. To summarize, Maxwell's equations in time-harmonic, plane-wave form are

$$\mathbf{k} \times \mathbf{E} = \omega \mu_o (\mathbf{H} + \mathbf{M}) \quad (1.3.4)$$

$$\mathbf{k} \times \mathbf{H} = -\omega (\varepsilon_o \mathbf{E} + \mathbf{P}) \quad (1.3.5)$$

$$\mathbf{k} \cdot (\varepsilon_o \mathbf{E} + \mathbf{P}) = 0 \quad (1.3.6)$$

$$\mathbf{k} \cdot (\mu_o \mathbf{H} + \mu_o \mathbf{M}) = 0 \quad (1.3.7)$$

where the fixed charge and current densities have been excluded. It is particularly relevant to remark that since the electric and magnetic Gaussian laws show zero divergence, (1.3.4 and 1.3.5) describe the field motion exclusively in the plane perpendicular to $\mathbf{k}$.

The Poynting theorem can likewise be recast into complex notation. The theorem is

$$\mathbf{k} \cdot (\mathbf{E} \times \mathbf{H}^*) = \omega \mu_o |H|^2 + \omega \varepsilon_o |E|^2 + \omega \mathbf{H}^* \cdot \mu_o \mathbf{M} + \omega \mathbf{E} \cdot \mathbf{P}^* \quad (1.3.8)$$

As long as there is no phase between $\mathbf{H}^*$ and $\mu_o \mathbf{M}$, and similarly between $\mathbf{E}$ and $\mathbf{P}^*$, then the power flow density experiences no gain or loss. However, a lead or lag of $\mathbf{M}$ to $\mathbf{H}^*$, or $\mathbf{P}^*$ to $\mathbf{E}$, introduces gain or loss into the system.

The complex Poynting vector is defined as

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}^* \quad (1.3.9)$$

and the time-average of $\mathbf{S}$ is found by

$$\langle \mathbf{S} \rangle = \frac{1}{2} \Re \{ \mathbf{E} \times \mathbf{H}^* \} \quad (1.3.10)$$

The following identities are useful for time-harmonic calculations:

$$\langle \mathbf{a}(\mathbf{r}, t) \mathbf{b}^*(\mathbf{r}, t) \rangle = \frac{1}{2} \Re \{ \mathbf{a}(\mathbf{r}) \mathbf{b}^*(\mathbf{r}) \} \quad (1.3.11a)$$

$$\langle \mathbf{a}(\mathbf{r}, t) \cdot \mathbf{b}^*(\mathbf{r}, t) \rangle = \frac{1}{2} \Re \{ \mathbf{a}(\mathbf{r}) \cdot \mathbf{b}^*(\mathbf{r}) \} \quad (1.3.11b)$$

$$\langle \mathbf{a}(\mathbf{r}, t) \times \mathbf{b}^*(\mathbf{r}, t) \rangle = \frac{1}{2} \Re \{ \mathbf{a}(\mathbf{r}) \times \mathbf{b}^*(\mathbf{r}) \} \quad (1.3.11c)$$

1.4 Classical Description of Polarization

Thus far the study of the vectorial nature of light has shown that a planar electro-magnetic wave is a solution to Maxwell's equations in free space, and that the wave has a phase velocity, wavelength, and dispersion relation. Moreover, the relation between electric and magnetic fields and the power flow of the wave have been determined. This section is addressed to the evolution of the electric field in the plane perpendicular to the propagation direction. The motion of the electric field in this plane governs the polarization of the wave. Separate discussion of the magnetic wave motion is redundant as the magnetic field is immediately derived from the electric field using Faraday's law.

Consider a time-harmonic monochromatic plane wave (1.3.3) that travels in the $\hat{z}$ direction ($\mathbf{k} \cdot \mathbf{r} = kz$), Fig. 1.1. Since $\mathbf{k} \cdot \mathbf{E} = 0$ in vacuum, so there is no $\hat{z}$ component to the electric field. The most general form of the electric field vector is then

$$\mathbf{E}(z, t) = \begin{pmatrix} E_x e^{j\phi_x} \\ E_y e^{j\phi_y} \end{pmatrix} e^{j(\omega t - kz)} \quad (1.4.1)$$

where $E_{x,y}$ are signed real numbers. The complex 2-row column vector is called the Jones polarization vector [5].

This plane wave propagates along the z -axis with wavelength $2\pi/k$ and phase velocity c . The two field components lie in the (x, y) plane and complete full cycles at rate ω . The polarization of the wave is governed by the electric-field evolution in the xy *Basis* plane. For convenience of notion but without loss of generality, $kz = \phi_x$. Using this reference plane and converting (1.4.1) to its real-valued counterpart, the electric field vector is

$$\mathbf{E}(x, y, t) = \hat{x}E_x \cos(\omega t) + \hat{y}E_y \cos(\omega t + \phi) \quad (1.4.2)$$

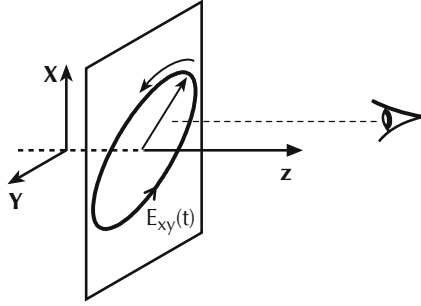


Fig. 1.1. In a vacuum, $\mathbf{k} \cdot \mathbf{E} = 0$, restricting the electric field to lie in the plane perpendicular to the propagation direction. Polarization is the motion of the electric field in the perpendicular plane.

where $\phi = \phi_y - \phi_x$. Equation (1.4.2) describes an ellipse in the plane perpendicular to $\hat{z}$. The convention used in this text to describe the state and handedness of the polarization ellipse is: the field is observed as it propagates towards the observer; that is, the observer faces in the $-\hat{z}$ direction, (see Fig. 1.1). The field is right-hand polarized when one's right-hand thumb points along $+z$ and one's fingers curl in the direction of electric-field vector motion.

The elliptical equation is derived from (1.4.2) as follows. The field amplitudes as projected along the $\hat{x}$ and $\hat{y}$ directions are

$$x = E_x \cos(\omega t) \quad (1.4.3a)$$

$$y = E_y \cos(\omega t + \phi) \quad (1.4.3b)$$

Taking the square of the parametric equations, adding and absorbing terms by identification with $xy/E_x E_y$ yields the elliptical equation

$$\frac{x^2}{E_x^2} + \frac{y^2}{E_y^2} - \frac{2xy}{E_x E_y} \cos \phi = \sin^2 \phi \quad (1.4.4)$$

There are three independent variables that govern the shape of the ellipse: E_x , E_y , and ϕ .

Figure 1.2 illustrates a general polarization ellipse resolved onto two coordinate systems. A general ellipse is one where there is no zero component in the (E_x, E_y, ϕ) triplet. In Fig. 1.2(a), $E_{x,y}$ mark the projections of the ellipse onto the (x, y) basis, and the angle χ is defined as $\tan \chi = E_y/E_x$ [4]. From the tangent relation between E_y and E_x , the Jones vector can be rewritten in normalized form:

$$\mathbf{E} = E_o \begin{pmatrix} \cos \chi \\ \sin \chi e^{j\phi} \end{pmatrix} \quad (1.4.5)$$

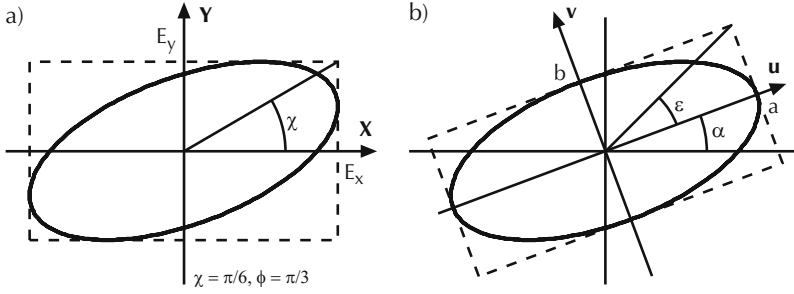


Fig. 1.2. Analysis of a general polarization ellipse onto the (x, y) and (u, v) coordinate systems. a) $E_{x,y}$ show maximum extent of elliptical motion on (x, y) basis. b) Same ellipse but where (u, v) basis is aligned to the major and minor elliptical axes. The angle between (x, y) and (u, v) is α .

where $E_o = \sqrt{E_x^2 + E_y^2}$ is the field amplitude irrespective of coordinate system. With this normalization, the state of polarization is described uniquely by the (χ, ϕ) pair of polarimetric parameters.

Now, as any ellipse has a major and minor axis, a coordinate system can be defined to align to these axes. Call this basis (u, v) , Fig. 1.2(b). In the (u, v) basis the elliptical equation is

$$\frac{u^2}{a^2} + \frac{v^2}{b^2} = 1 \quad (1.4.6)$$

where (a, b) , the major and minor axes of the ellipse, are the projections onto the u and v axes, respectively. The parametric time-evolution equations that result in ellipse (1.4.6) are

$$u = a \cos \omega t \quad (1.4.7a)$$

$$v = b \sin \omega t \quad (1.4.7b)$$

As χ is defined as the tangent angle between E_y and E_x , ε is likewise defined as $\tan \varepsilon = b/a$. The ellipses described by (1.4.4) and (1.4.6) are related by a rotation in the plane through angle α . That is,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.4.8)$$

Substituting the elliptical projections (1.4.3) and (1.4.7) into the above rotation, the angle of rotation α is

$$\tan 2\alpha = \tan 2\chi \cos \phi \quad (1.4.9)$$

To verify that the rotation was unitary, one can show that $a^2 + b^2 = E_x^2 + E_y^2$. An important conclusion is that while the (u, v) basis is the natural coordinate

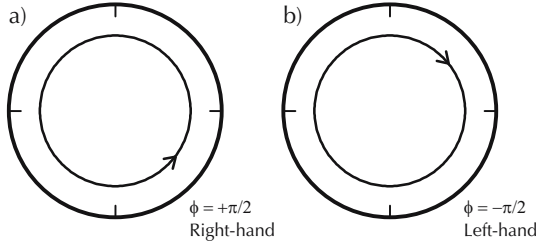


Fig. 1.3. Two states of circular polarization, counterclockwise (right-hand circular, or R) and clockwise (left-hand circular, or L). Right- and left-hand circular states are distinguished by the curl of one's fingers with the thumb pointing along the $+\hat{z}$ direction. Circular polarization exists when $\chi = \pm 45^\circ$ and $\phi = \pm\pi/2$. a) Counterclockwise (R) corresponds to $\phi = \pi/2$. b) Clockwise (L) corresponds to $\phi = -\pi/2$.

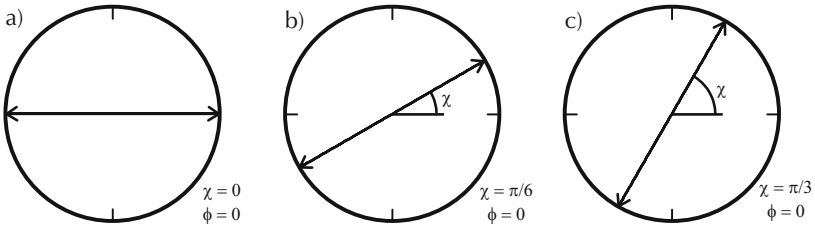


Fig. 1.4. Linear states of polarization exist when $\phi = m\pi$, where m is an integer. The orientation of the state is determined by χ , or alternatively by α . From a) to c), the value of α increases.

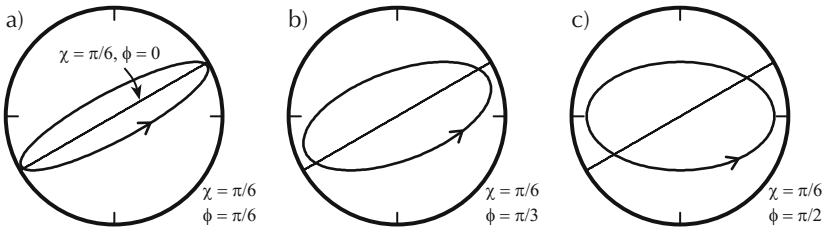


Fig. 1.5. Three elliptical polarization states. All three states have same value of χ . The phase difference ϕ increases: a) $\phi = \pi/6$, b) $\phi = \pi/3$, and c) $\phi = \pi/2$. Both χ and ϕ play a role in the orientation α of the ellipse, as governed by $\tan 2\alpha = \tan 2\chi \cos \phi$.

system for an ellipse having arbitrary rotation α , any unit ellipse may equally well be described on an arbitrary (x, y) basis by the (χ, ϕ) pair. The coordinate pairs (χ, ϕ) and (ε, α) are in one-to-one correspondence.

The parametric electric field described by (1.4.2) exhibits a handedness that depends on the sign of ϕ . For the range $-\pi \leq \phi < 0$, the evolution of the ellipse is in the clockwise (cw) direction and the handedness is left (L). For the range $0 < \phi \leq \pi$, the evolution is in the counterclockwise (ccw) direction and the handedness is right (R). The sense of the handedness is lost in elliptical equation (1.4.4) since $\cos \phi$ is an even function and $\sin^2 \phi$ is positive definite. The same loss of handedness shows, however, that the shape of the ellipse is independent of the rotary sense.

There are three general categories of polarization state: circular, linear, and elliptical. Taken as a progression, circular is the most restrictive on the possible (χ, ϕ) values, linear is less restrictive, and elliptical places no restrictions on (χ, ϕ) . In particular, circular polarization requires $\chi = \pm\pi/4$ and $\phi = \pm\pi/2$. Handedness is the only distinguishing property. When (χ, ϕ) have the same sign, the sense is R ; when the signs are opposite the sense is L . Linear polarization lets χ take any value and requires $\phi = m\pi$, where m is an integer. Elliptical polarization includes circular and linear states as well as all other possible values of (χ, ϕ) . Figures 1.3–1.5 provide examples of these three categories.

The polarization ellipse is completely described by the (χ, ϕ) pair. The question is how to determine these polarimetric parameters uniquely for an arbitrary state having arbitrary intensity. The following series of seven measurements will uniquely determine the state. The first measurement is for the overall time-averaged intensity. For a fixed polarization state

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y e^{j\phi} \end{pmatrix} \quad (1.4.10)$$

where E_x and E_y are real, the time-averaged intensity is¹

$$I_o = \frac{1}{2} \Re(\mathbf{E}^* \cdot \mathbf{E}) \quad (1.4.11)$$

$$= (E_x^2 + E_y^2)/2 \quad (1.4.12)$$

The remaining six measurements use a linear polarizer and, in two cases, a quarter-wave waveplate, to make the measurements. The projection matrix is a suitable model of a linear polarizer [10]

$$\mathcal{P} = \begin{pmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{pmatrix} \quad (1.4.13)$$

¹ The time-average here is only over a few optical cycles. Partial polarization takes time-averages over longer periods.

The origin of this matrix is derived in Chapter 2. The angle θ is the angle of the polarizer to the horizontal axis. Any particular component intensity is calculated from $I_k \propto \mathbf{E}^\dagger \mathcal{P}(\theta) \mathbf{E}$. The first pair of measurements orient the polarizer in the $\hat{x}$ direction and $\hat{y}$ direction. The component intensities are

$$I_x = E_x^2/2 \quad (1.4.14a)$$

$$I_y = E_y^2/2 \quad (1.4.14b)$$

The second pair of measurements orient the polarizer in the $+45^\circ$ and -45° directions. The component intensities are

$$I_{+45} = (E_x^2 + E_y^2)/4 + (E_x E_y/2) \cos \phi \quad (1.4.15a)$$

$$I_{-45} = (E_x^2 + E_y^2)/4 - (E_x E_y/2) \cos \phi \quad (1.4.15b)$$

One more measurement pair is necessary because handedness cannot be determined since $\cos \phi$ is an even function of ϕ . To complete the measurements, the optical beam is passed through a $+45^\circ$ -oriented quarter-wave waveplate and an $\hat{x}$ - or $\hat{y}$ -oriented polarizer so as to convert R and L hand circular polarizations to linear horizontal and vertical, respectively. The resulting intensities are

$$I_R = (E_x^2 + E_y^2)/4 + (E_x E_y/2) \sin \phi \quad (1.4.16a)$$

$$I_L = (E_x^2 + E_y^2)/4 - (E_x E_y/2) \sin \phi \quad (1.4.16b)$$

These seven measurements can be succinctly combined into four terms called Stokes parameters, which are defined by the equations

$$\begin{aligned} S_0 &= I_x + I_y &= (E_x^2 + E_y^2)/2 &= \frac{1}{2} E_o^2 \\ S_1 &= I_x - I_y &= (E_x^2 - E_y^2)/2 &= \frac{1}{2} E_o^2 \cos 2\chi \\ S_2 &= I_{+45} - I_{-45} &= E_x E_y \cos \phi &= \frac{1}{2} E_o^2 \sin 2\chi \cos \phi \\ S_3 &= I_R - I_L &= E_x E_y \sin \phi &= \frac{1}{2} E_o^2 \sin 2\chi \sin \phi \end{aligned} \quad (1.4.17)$$

From these equations the polarization coordinates (χ, ϕ) can be uniquely determined. Table 1.1 displays representative states in Jones and Stokes form.

1.4.1 Stokes Vectors, Jones and Muller Matrices

The Stokes vector $\mathbf{S}$ is defined by the projector construct (1.4.17). In general, one can write

$$\mathbf{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (1.4.18)$$

The Stokes vector is the analogue to the Jones vector (1.4.5) on page 13. One must recognize that directly underlying the Jones vector are Maxwell's equations. The problem is that the Jones vector cannot be directly measured, but the Stokes vector can. The Jones vector is reconstructed from a Stokes vector to within a complex c constant by inverting (1.4.17):

$$\mathbf{E} = c \begin{pmatrix} \sqrt{\frac{1}{2}(1 + S_1/S_0)} \\ \sqrt{\frac{1}{2}(1 - S_1/S_0)} \exp(j \tan^{-1} S_3/S_2) \end{pmatrix} \quad (1.4.19)$$

Other than the undetermined complex constant c , there are three free variables in (1.4.19). A Jones vector, however, has four free variables: two amplitudes and two phases. The fourth free variable is the common phase of the two polarization components; this common phase is lost in the intensity measurements.

When light propagates through a medium, the interaction between medium and light can impart a change in the polarization state. In Stokes space, the change of state to $\mathbf{S}'$ from $\mathbf{S}$ is determined by the Mueller matrix $\mathbf{M}$. The general transformation is

$$\begin{pmatrix} S'_0 \\ S'_1 \\ S'_2 \\ S'_3 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (1.4.20)$$

In matrix form one writes $\mathbf{S}' = \mathbf{M}\mathbf{S}$. The Mueller matrix is a 4×4 matrix with real-valued entries. Polarimetric measurements find the Mueller matrix elements directly.

Underlying a Stokes-state transformation $\mathbf{M}$ is the Jones-state transformation $\mathbf{J}$. As with vectors, the Jones transformation matrix comes directly from Maxwell's equations; were it not for the natural advantages of polarimetric measurements the Mueller matrix would simply be a tautology. The Mueller matrix is in any case the analogue to the Jones matrix. In Jones space, an output vector $\mathbf{E}'$ is related to the input vector $\mathbf{E}$ through

$$\mathbf{E}' = \mathbf{J}\mathbf{E} \quad (1.4.21)$$

The Jones matrix $\mathbf{J}$ is a 2×2 matrix with complex-valued entries. The connection between Jones and Mueller matrices is derived using Pauli matrices (cf. §2.6.2). The Mueller matrix is derived from the Jones matrix via

$$\mathbf{M}_{i+1,j+1} = \frac{1}{2} \text{Tr}(\mathbf{J}\sigma_j \mathbf{J}^\dagger \sigma_i) \quad (1.4.22)$$

where $i, j = 0, 1, 2$ or 3 , σ_i is the i^{th} Pauli matrix, and Tr is the trace operator. The derivation of this expression is given in §2.6.2 starting on page 66.

Equation (1.4.22) is not invertible directly. However, R. C. Jones prescribes the way to reconstruct a Jones matrix from output Stokes vectors after three measurements [3, 6]. The three input states for the measurement are $\mathbf{S}_a = (1, 1, 0, 0)^T$, $\mathbf{S}_b = (1, -1, 0, 0)^T$, and $\mathbf{S}_c = (1, 0, 1, 0)^T$. Three output Jones vectors are constructed from the sequence:

$$\begin{pmatrix} \mathbf{S}_a \\ \mathbf{S}_b \\ \mathbf{S}_c \end{pmatrix} \xrightarrow{\mathbf{M}} \begin{pmatrix} \mathbf{S}'_a \\ \mathbf{S}'_b \\ \mathbf{S}'_c \end{pmatrix} \xrightarrow{\text{to Jones}} \begin{pmatrix} \mathbf{E}'_a \\ \mathbf{E}'_b \\ \mathbf{E}'_c \end{pmatrix} \quad (1.4.23)$$

From these three Jones vectors four complex ratios are calculated:

$$k_1 = E'_{xa}/E'_{ya}, \quad k_2 = E'_{xb}/E'_{yb}, \quad k_3 = E'_{xc}/E'_{yc}, \quad k_4 = \frac{k_3 - k_2}{k_1 - k_3} \quad (1.4.24)$$

To within a complex constant c , as before, the reconstructed Jones matrix is

$$\mathbf{J} = c \begin{pmatrix} k_1 k_4 & k_2 \\ k_4 & 1 \end{pmatrix} \quad (1.4.25)$$

Two classes of Jones matrices are particularly important for polarization studies: the Hermitian matrix and unitary matrix. Either matrix is written in the form

$$\mathbf{J} = \begin{pmatrix} a_0 + a_1 & a_2 - ja_3 \\ a_2 + ja_3 & a_0 - a_1 \end{pmatrix} \quad (1.4.26)$$

A Hermitian matrix represents a measurement of the polarization state and thus has real-valued eigenvalues. All four coefficients $a_{0,1,2,3}$ in (1.4.26) are real numbers. A unitary matrix represents a coordinate transformation of the Stokes vectors but imparts no loss or gain. Its eigenvalues are related through the matrix exponential, cf. §2.4.3. The Mueller equivalents to these matrices depend on the details, but the characteristic matrix forms are

$$\mathbf{J}_H \longrightarrow \mathbf{M}_H = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{pmatrix}, \quad \mathbf{J}_U \longrightarrow \mathbf{M}_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix} \quad (1.4.27)$$

While a Hermitian matrix scatters energy to all elements of the Mueller matrix a unitary matrix keeps all of the light within the three spherical Stokes coordinates; the vector length S_0 remains unchanged. This characteristic form shows that $\mathbf{J}_U$ imparts only a rotation.

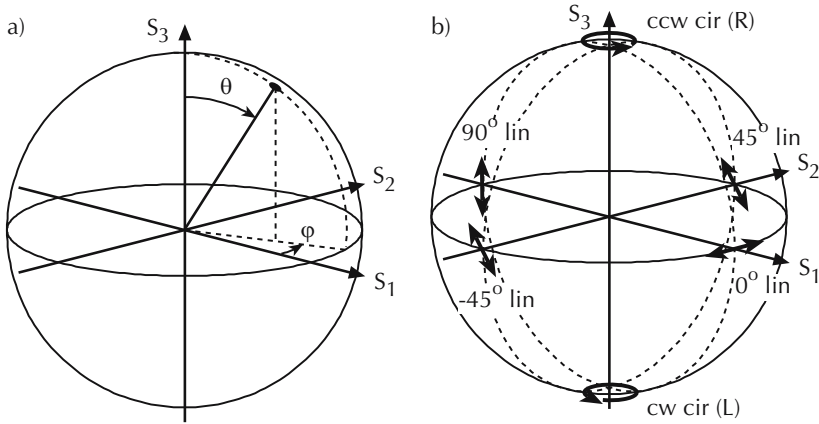


Fig. 1.6. Spherical representation of polarization states. a) The cartesian basis is (S_1, S_2, S_3) . The equivalent spherical basis is (r, θ, φ) . On a unit sphere, $r = 1$, so (θ, φ) coordinates uniquely determine position. b) Identification of particular polarization states on the Poincaré sphere. Along the equator lie linear states. At the north and south poles lie ccw (R) and cw (L) circular states. All remaining points are elliptical states. Orthogonal states are point pairs on opposite sides of the sphere connected by a cord that runs through the origin.

1.4.2 The Poincaré Sphere

Every possible polarization state can be represented on the surface of a unit sphere. The unit sphere is called the Poincaré sphere after H. Poincaré, its creator. A unit sphere is made by normalizing the three-directional Stokes components $S_{1,2,3}$ by the intensity component S_0 . On a unit sphere, the declination and azimuth angles θ and φ describe any point on the surface. Referring to the polar coordinates illustrated in Fig. 1.6(a), the azimuth and declination angles are projected onto the (S_1, S_2, S_3) basis as

$$\begin{aligned} S_1 &= \sin \theta \cos \varphi \\ S_2 &= \sin \theta \sin \varphi \\ S_3 &= \cos \theta \end{aligned} \tag{1.4.28}$$

Associating spherical parameters to ellipse parameters $\theta = 2\varepsilon$ and $\varphi = 2\alpha$, the normalized Stokes components $S_{1,2,3}$ of (1.4.17) are related to the spherical coordinates as

$$\begin{aligned} S_1/S_0 &= \sin 2\varepsilon \cos 2\alpha = \cos 2\chi \\ S_2/S_0 &= \sin 2\varepsilon \sin 2\alpha = \sin 2\chi \cos \phi \\ S_3/S_0 &= \cos 2\varepsilon = \sin 2\chi \sin \phi \end{aligned} \tag{1.4.29}$$

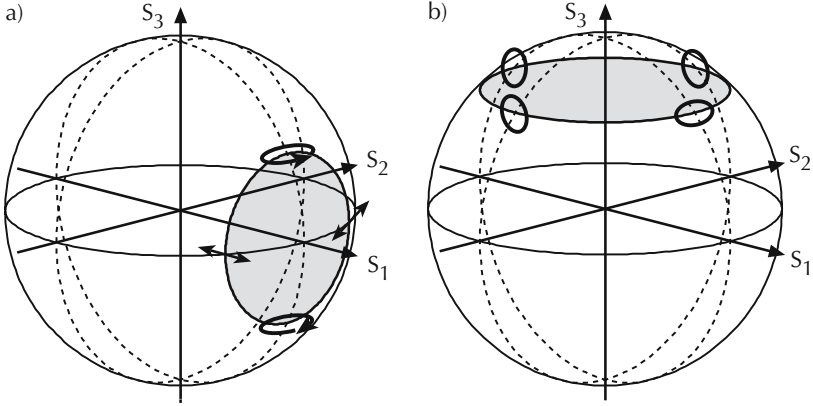


Fig. 1.7. Polarization contours. a) Contour of states for fixed χ and $-\pi \leq \phi \leq \pi$. The phase slips through a full revolution. This effect can be achieved physically by transmission through a waveplate. b) Contour of states for fixed ε and $-\pi/2 \leq \alpha \leq \pi/2$. The ellipse does a full rotation while maintaining its eccentricity. This effect can be achieved physically by transmission through an optically active waveplate.

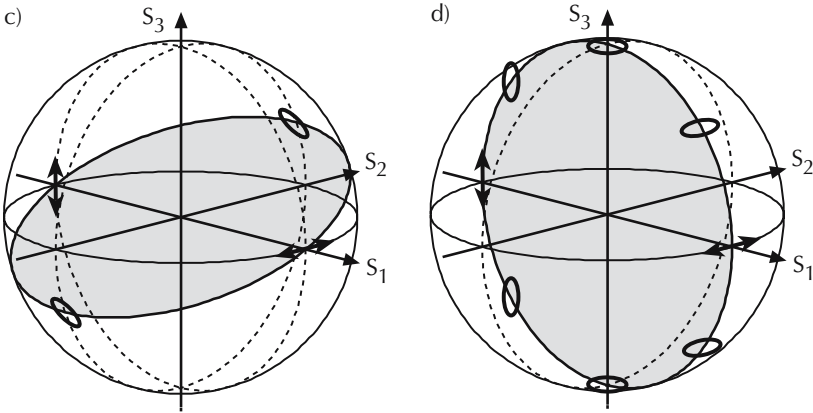


Fig. 1.7. Polarization contours. c) Contour of states for fixed ϕ and for $-\pi/2 \leq \chi \leq \pi/2$. χ determines the tilt of the plane. Any two orthogonal states lie on such a contour, the states being separated by 180° . d) Contour of states for fixed α and $-\pi \leq \varepsilon \leq \pi$. The eccentricity of the ellipse varies between linear and circular, but the pointing direction remains either vertical or horizontal.

Figure 1.6(b) illustrates the polarization states on the coordinate axes. Figure 1.7(a–d) illustrates various contours on the Poincaré sphere and their associations with ε , α , χ , and ϕ .

It is significant that the variables χ , ε , and α have a multiplier of two in (1.4.29) while ϕ does not. Physically, any full 2π phase slip of ϕ yields the identical polarization state; distinct optical phases within a 2π range correspond to distinct polarization states. In contrast, a π change in the χ , ε , and α parameters does not change the state. This is physically reasonable as an ellipse is preserved under 180° rotation, and $(E_x, E_y) \rightarrow (-E_x, -E_y)$ or $(a, b) \rightarrow (-a, -b)$ inversion. Jones space includes a built in degeneracy of elliptical parameters χ , ε , and α .

The spherical representation provides a geometric interpretation of the transformations that polarization states undergo when propagating through birefringent media. This representation will be used extensively throughout the text. There are, however, two drawbacks to the geometric interpretation. First, as the Stokes parameters are determined through measurements of intensity, only the polarization phase ϕ modulo 2π can be determined. In the study of polarization-mode dispersion, two orthogonally polarized waves can accrue thousands of 2π phase revolutions. As delay τ is defined as $\tau = \partial\phi/\partial\omega$, is it essential to track the total number of phase revolutions as well as any partial slip. Polarization-mode dispersion requires a modification to the Stokes calculus to treat the delay as well as the phase. Second, the polarization of a state by an arbitrarily oriented polarizer is difficult to picture in Stokes space. The projection due to the polarizer is more easily pictured in physical space. It is good practice to intuit a polarization state seamlessly in both Stokes and Jones space as a more robust understanding is achieved.

1.5 Partial Polarization

A wave is fully polarized when all component polarizations of a ray-bundle oscillate coherently. Such is the case with a laser. By contrast, “natural” light, such as light from the sun, is fully depolarized: the components of a ray-bundle are completely incoherent and the instantaneous polarization over a differential bandwidth can point in any direction on the Poincaré sphere. Partially polarized light can be “naturally” partially polarized in that some fraction of the ray-bundle is polarized and the remaining part “naturally” depolarized, or can be “pseudo” depolarized in that all components individually remain fully polarized but the polarization of the sum is not. The instantaneous polarization of pseudo-depolarized light touches a limited loci of points on the Poincaré sphere.

There are two ways to express partial polarization: the degree of polarization (DOP, denoted $\mathcal{D}$) and the Jones coherency matrix J . DOP is a scalar value between zero and one and can be expressed in terms of Stokes or Jones parameters. The Jones coherency matrix is derived from the dyadic form of

the Jones vector and is used to trace depolarization through a system in Jones space. The coherency matrix is a necessary augmentation to Jones calculus because the 16 free variables of the Mueller matrix are enough to include depolarization directly, while that eight free variables of the Jones matrix do not provide enough freedom.

In terms of Stokes parameters, DOP is defined as

$$\mathcal{D} = \frac{\sqrt{\langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2}}{\langle S_0 \rangle} \quad (1.5.1)$$

where the time averages are given by

$$\langle S(t) \rangle = \frac{1}{T} \int_0^T S(t) dt$$

The time average is taken over all time-varying quantities, i.e. ωt , $\chi(t)$, $\phi(t)$, etc. $\mathcal{D} = 1$ means that all waves that make up a ray bundle each have fully determined, time-invariant polarizations. $\mathcal{D} = 0$ means the polarimetric terms of the ray bundle have vanishing time averages, but the underlying cause, e.g. whether from incoherence or pseudo-depolarization, cannot be discerned using $\mathcal{D}$ alone. An intermediate value of $\mathcal{D}$ means that some of the optical power is polarized and the remaining power is not.

In terms of the coherency matrix, DOP is defined as

$$\mathcal{D} = \sqrt{1 - \frac{4 \det(J)}{\text{Tr}(J)^2}} \quad (1.5.2)$$

The coherency matrix is defined by $J = \langle \mathbf{E} \mathbf{E}^\dagger \rangle$ [9], where

$$\mathbf{E}(t) = \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix}, \quad \text{and} \quad J = \begin{pmatrix} \langle e_x^* e_x \rangle & \langle e_x e_y^* \rangle \\ \langle e_x^* e_y \rangle & \langle e_y^* e_y \rangle \end{pmatrix} \quad (1.5.3)$$

and where (e_x, e_y) are complex numbers. Finally, the time-averaged Stokes parameters in terms of the coherency-matrix elements are

$$\begin{pmatrix} \langle S_0 \rangle \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -j & j \end{pmatrix} \begin{pmatrix} J_{xx} \\ J_{yy} \\ J_{xy} \\ J_{yx} \end{pmatrix} \quad (1.5.4)$$

Both $\mathcal{D}$ and J are inherently time-average measures. The integration period can affect the reported values. For instance, a monochromatic source that has a coherence time of 0.1 sec certainly produces polarized waves on time-scales $T \ll 0.1$ sec. However, polarization states separated by $T > 0.1$ sec are uncorrelated. A $\mathcal{D}$ measure taken over a long time scale would produce a subunity value, while a $\mathcal{D}$ measure over a short time scale would produce $\mathcal{D} \rightarrow 1$.

Both answers are technically correct and the issue reduces to what is a relevant time scale. That will depend on the application.

The following studies of partial polarization are grouped into ray bundles comprised of coherent, or polarized, components; incoherent, or depolarized, components; heterogeneous combinations of coherent and incoherent components; and pseudo-depolarized components. In all cases the ray-bundle components are collinear. In the following calculations, the electric-field spectrum is denoted as

$$\mathbf{E}(\omega) = E_o G(\omega) \sum_n \hat{p}_n(\omega) \quad (1.5.5)$$

where $G(\omega)$ is the spectral profile, E_o is complex, and $\hat{p}_n(\omega)$ is the n^{th} polarization at ω . The time-dependent field $\mathbf{E}(t)$ is the inverse Fourier transform of $\mathbf{E}(\omega)$:

$$\mathbf{E}(t) = \sum_n \int E_o G(\omega) \hat{p}_n(\omega) e^{j\omega t} d\omega \quad (1.5.6)$$

1.5.1 Coherently Polarized Waves

The common feature of the four cases studied below is that the polarization of each component is time-invariant and independent of frequency. The study begins with a single monochromatic wave and generalizes to narrowband ray bundles having either discrete or continuous spectra. The studies show that for coherently polarized waves, only pseudo-depolarization can reduce the degree of polarization below unity.

A Monochromatic Polarized Wave

The simplest case is a single monochromatic polarized plane wave. The field spectrum is

$$\mathbf{E}(\omega) = E_o \delta(\omega - \omega_o) \hat{p} \quad (1.5.7)$$

where $\delta(\omega - \omega_o)$ is the Dirac delta function centered at ω_o . In the time domain, the plane wave is

$$\mathbf{E}(t) = E_o e^{j\omega_o t} \begin{pmatrix} \cos \chi \\ \sin \chi e^{j\phi} \end{pmatrix}$$

The corresponding Stokes parameters are

$$\langle S \rangle = |E_o|^2 \begin{pmatrix} 1 \\ \cos 2\chi \\ \sin 2\chi \cos \phi \\ \sin 2\chi \sin \phi \end{pmatrix} \quad (1.5.8)$$

As χ and ϕ are fixed in time, substitution of (1.5.8) into (1.5.1) yields $\mathcal{D} = 1$. The coherency matrix is

$$J = \begin{pmatrix} \cos^2 \chi & e^{-j\phi} \sin \chi \cos \chi \\ e^{j\phi} \sin \chi \cos \chi & \sin^2 \chi \end{pmatrix} \quad (1.5.9)$$

The polarization state of this wave is completely determined.

A Monochromatic Wave Having Multiple Polarizations

The spectrum of a ray bundle that comprises multiple monochromatic polarized waves of multiple polarization components is written as

$$\mathbf{E}(\omega) = \sum_n E_{on} \delta(\omega - \omega_o) \hat{p}_n \quad (1.5.10)$$

The time-domain field of the ray bundle is

$$\mathbf{E}(t) = e^{j\omega_o t} \sum_n E_{on} \begin{pmatrix} \cos \chi_n \\ \sin \chi_n e^{-j\phi_n} \end{pmatrix}$$

While the polarimetric parameters of the combined wave may be complicated, they do not vary in time. One can verify that

$$\langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2 = (e_x^* e_x + e_y^* e_y)^2$$

and thus $\mathcal{D} = 1$. A ray bundle that is constituted from multiple monochromatic coherent waves has a polarization state that is completely determined. The intensity of the ray bundle is calculated from $\langle S_1 \rangle$, or

$$I_{\text{coh}} = \langle S_0 \rangle = \sum_n |E_{on}|^2 \quad (1.5.11)$$

where I_{coh} denotes the intensity of the coherent waves.

Narrowband Polarized Waves with Discrete Spectrum

Consider an extension of (1.5.7) where the spectrum comprises multiple frequency components, each component itself being polarized:

$$\mathbf{E}(\omega) = \sum_n E_{on} \delta(\omega - \omega_n) \hat{p}_n \quad (1.5.12)$$

In the time domain, this discretely polychromatic wave is

$$\mathbf{E}(t) = \sum_n e^{j\omega_n t} E_{on} \begin{pmatrix} \cos \chi_n \\ \sin \chi_n e^{j\phi_n} \end{pmatrix} \quad (1.5.13)$$

The polarimetric parameters χ_n and ϕ_n for each frequency component are fixed in time ($\mathcal{D}_n = 1$) and the frequencies ω_n are distinct. After summation, however, the composite polarimetric parameters do depend on time. In general this leads to $\mathcal{D}_{\text{total}} < 1$.

The depolarization is calculated as follows. Consider first the S_1 Stokes parameter:

$$\begin{aligned} S_1 &= e_x^* e_x - e_y^* e_y \\ &= \sum_m e^{-j\omega_m t} E_{om}^* \cos \chi_m \sum_n e^{j\omega_n t} E_{on} \cos \chi_n \\ &\quad - \sum_m e^{-j\omega_m t} E_{om}^* \sin \chi_m e^{-j\phi_m} \sum_n e^{j\omega_n t} E_{on} \sin \chi_n e^{j\phi_n} \end{aligned}$$

The time averages on normalized components are

$$\left\langle \sum_m e^{-j\omega_m t} \cos \chi_m \sum_n e^{j\omega_n t} \cos \chi_n \right\rangle = \sum_n \cos^2 \chi_n$$

and

$$\left\langle \sum_m e^{-j\omega_m t} \sin \chi_m e^{-j\phi_m} \sum_n e^{j\omega_n t} \sin \chi_n e^{j\phi_n} \right\rangle = \sum_n \sin^2 \chi_n$$

where the time-average window is $T \gg [\min(\omega_n - \omega_m)]^{-1}$. All cross terms are eliminated upon averaging, and the same holds for S_2 and S_3 . In general the three time-averaged Stokes parameters are

$$\langle S_k \rangle = \sum_n \langle S_{kn} \rangle \quad (1.5.14)$$

Accordingly, the degree of polarization is

$$\mathcal{D} = \frac{\sqrt{(\sum_n \langle S_{1n} \rangle)^2 + (\sum_n \langle S_{2n} \rangle)^2 + (\sum_n \langle S_{3n} \rangle)^2}}{\sum_n \langle S_{0n} \rangle}$$

Since $\mathcal{D}_n$ of each component is unity, it follows that $\langle S_{0n} \rangle^2 = \langle S_{1n} \rangle^2 + \langle S_{2n} \rangle^2 + \langle S_{3n} \rangle^2$. By iterating the triangle inequality

$$|\mathbf{r}_1 + \mathbf{r}_2| \leq |\mathbf{r}_1| + |\mathbf{r}_2|$$

one concludes that

$$\sqrt{(\sum_n \langle S_{1n} \rangle)^2 + (\sum_n \langle S_{2n} \rangle)^2 + (\sum_n \langle S_{3n} \rangle)^2} \leq \sum_n \langle S_{0n} \rangle$$

Therefore, in general, the DOP for a discretely polychromatic ray bundle is

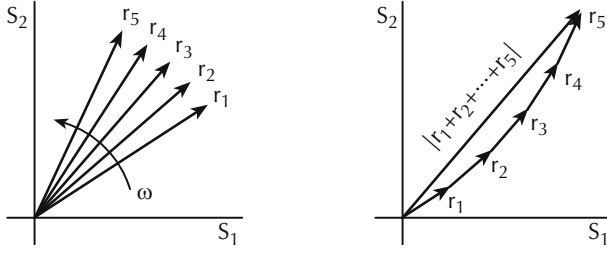


Fig. 1.8. Stokes vectors r_k in a plane. On the left, individual vector components: the vector direction is a function of frequency. On the right, the length of the vector sum is generally less than the arithmetic sum of the vector lengths.

$$\mathcal{D} = \frac{\sqrt{(\sum_n \langle S_{1n} \rangle)^2 + (\sum_n \langle S_{2n} \rangle)^2 + (\sum_n \langle S_{3n} \rangle)^2}}{I_{\text{coh}}} \leq 1 \quad (1.5.15)$$

where I_{coh} is given by (1.5.11). Equation (1.5.15) does provide some physical insight even though a specific expression is lacking. As Fig. 1.8 illustrates, when the Stokes vectors for the various frequencies are nearly aligned, then $\mathcal{D} \sim 1$. However, when the vector components are not aligned the overall DOP is reduced. Passage through a birefringent element can pseudo-depolarize this ray bundle (more detail is found in §1.5.3), but otherwise the addition of more coherent components in and of itself does not decrease the degree of polarization of the total.

A Narrowband Polarized Wave With Continuous Spectrum

A narrowband polarized wave is one where a modulation has been imprinted on a carrier. The broadening of the spectrum in this way does not entail a frequency-dependent polarization rotation. Accordingly, the spectrum is written as

$$\mathbf{E}(\omega) = E_o |G(\omega)| e^{j\theta_G(\omega)} \hat{p} \quad (1.5.16)$$

where $G(\omega)$ is the modulated spectral profile. $G(\omega)$ is continuous for broadband modulation and discrete for harmonic modulation. The polarization direction is fixed along $\hat{p}$ and the profile amplitude is taken as a bound function which goes to zero outside a bandwidth of $\Delta\omega$. The time-domain electric field is

$$\mathbf{E}(t) = \hat{p} \int_{\Delta\omega} E_o |G(\omega)| e^{j\theta_G(\omega)} e^{j\omega t} d\omega$$

Consider first the $e_x^* e_x$ product:

$$\begin{aligned} e_x^* e_x &= \int_{\Delta\omega} \int_{\Delta\omega} |E_o|^2 |G(\omega_1)| |G(\omega_2)| e^{j(\theta_G(\omega_2) - \theta_G(\omega_1))} \times \\ &\quad e^{j(\omega_2 - \omega_1)t} \cos^2 \chi \, d\omega_1 d\omega_2 \end{aligned}$$

Simplification comes with the time-average operation, where

$$\langle e_x^* e_x \rangle = \frac{1}{T} \int_0^T e_x^* e_x dt$$

generates a Dirac delta $\delta(\omega_2 - \omega_1)$ once the temporal integral is moved through to the $\exp j(\omega_2 - \omega_1)t$ term. Therefore,

$$\begin{aligned} \langle e_x^* e_x \rangle &= \int_{\Delta\omega} \int_{\Delta\omega} |E_o|^2 |G(\omega_1)| |G(\omega_2)| e^{j(\theta_G(\omega_2) - \theta_G(\omega_1))} \times \\ &\quad \cos^2 \chi \delta(\omega_2 - \omega_1) d\omega_1 d\omega_2 \\ &= |E_o|^2 \mathcal{I}_G \cos^2 \chi \end{aligned} \quad (1.5.17)$$

where the integral $\mathcal{I}_G$ is

$$\mathcal{I}_G = \int_{\Delta\omega} |G(\omega)|^2 d\omega \quad (1.5.18)$$

Following the same procedure,

$$\langle e_y^* e_y \rangle = |E_o|^2 \mathcal{I}_G \sin^2 \chi$$

and

$$\begin{aligned} \langle e_x^* e_y \rangle &= \langle e_x e_y^* \rangle^* \\ &= |E_o|^2 \mathcal{I}_G \sin \chi \cos \chi e^{j\phi} \end{aligned}$$

The time-averaged Stokes parameters are

$$\langle S \rangle = |E_o|^2 \mathcal{I}_G \begin{pmatrix} 1 \\ \cos 2\chi \\ \sin 2\chi \cos \phi \\ \sin 2\chi \sin \phi \end{pmatrix} \quad (1.5.19)$$

and thus $\mathcal{D} = 1$. This derivation shows that line broadening due to modulation does not in itself alter the degree of polarization of the light. The light can be pseudo-depolarized, however. Contrary to a discrete spectrum, for a continuous spectrum $\mathcal{D} \rightarrow 0$ monotonically with increasing bandwidth-delay from the depolarizing element.

1.5.2 Incoherently Depolarized Waves

Incoherently depolarized waves are comprised of individual components having time-varying polarimetry parameters. Light from the sun or noise from an optical amplifier are examples of completely depolarized light. An exposed air-gap polarization-dependent delay line, used to generate differential-group delay, can have a time-dependent retardance with a fixed ellipsometric orientation. The DOP of this source depends on the orientation of the input state.

A Narrowband Incoherent Wave

An narrowband incoherent wave is one where the projection angle χ and/or the phase slip ϕ changes with time. The field amplitude may also change in time, but that impacts only the wave intensity rather than the polarization state. The time scale for $\chi(t)$ and $\phi(t)$ change is assumed to be significantly shorter than the integration time of the $\mathcal{D}$ measurement. Moreover, it is understood that $\phi(t)$ of a single wave is synonymous with frequency shift, which makes the wave technically narrowband rather than monochromatic; it is assumed that $\phi(t)$ changes slowly enough so that the line broadening is inconsequential. A narrowband, incoherent-wave spectrum is written as

$$\mathbf{E}(\omega) = E_o \delta(\omega - \omega_o) \hat{p}(B) \quad (1.5.20)$$

where B denotes a spectral bandwidth which is consistent with the integration time. The time-domain field is

$$\mathbf{E}(t) = E_o e^{j\omega_o t} \begin{pmatrix} \cos \chi(t) \\ \sin \chi(t) e^{j\phi(t)} \end{pmatrix} \quad (1.5.21)$$

The corresponding Stokes parameters are

$$S(t) = |E_o|^2 \begin{pmatrix} 1 \\ \cos 2\chi(t) \\ \sin 2\chi(t) \cos \phi(t) \\ \sin 2\chi(t) \sin \phi(t) \end{pmatrix} \quad (1.5.22)$$

Consider an exposed air-gap polarization-dependent delay line with a stable input polarization. The input polarization beam splitter projects the input light onto two orthogonal axes and delays one with respect to the other. For a stable input polarization, the projection is fixed in time: $\chi(t) = \chi_o$. The exposed delay arm, however, imparts a time-varying retardance. In this case, the time-averaged Stokes parameters are

$$\langle S \rangle = |E_o|^2 \begin{pmatrix} 1 \\ \cos 2\chi_o \\ 0 \\ 0 \end{pmatrix}$$

The degree of polarization is therefore

$$\mathcal{D} = |\cos 2\chi_o| \quad (1.5.23)$$

$\mathcal{D}$ can attain values $0 \leq \mathcal{D} \leq 1$. When $\chi_o = 0^\circ$ all light travels in one arm or the other. Therefore $\mathcal{D} = 1$ as no relative phase shift is experienced. Alternatively, when $\chi_o = 45^\circ$, the light is equally split between the two arms and $\mathcal{D} = 0$. One should be careful about the stability of air-gap polarization controllers.

Separately, consider the more general case where both χ and ϕ change in time. In this case $\langle S \rangle = [1 \ 0 \ 0 \ 0]^T$ and $\mathcal{D} = 0$ over suitably long integration periods.

Multiple Narrowband Incoherent Waves

As an extension of (1.5.21), a ray bundle composed of multiple narrowband incoherent waves is written as

$$\mathbf{E}(t) = e^{j\omega t} \sum_n E_{on} \begin{pmatrix} \cos \tilde{\chi}_n \\ \sin \tilde{\chi}_n e^{j\tilde{\phi}_n} \end{pmatrix} \quad (1.5.24)$$

where $\tilde{\chi}$ and $\tilde{\phi}$ denote random variables χ and ϕ in time. The distributions of $\tilde{\chi}$ and $\tilde{\phi}$ are uniform for each wave in the ray bundle. The compound polarimetric parameters depend on time, too, and the averages are found as follows. Consider first the S_1 term, where

$$\begin{aligned} S_1 &= e_x^* e_x - e_y^* e_y \\ &= \sum_m E_{om}^* \cos \tilde{\chi}_m \sum_n E_{on} \cos \tilde{\chi}_n \\ &\quad - \sum_m E_{om}^* \sin \tilde{\chi}_m e^{-j\tilde{\phi}_m} \sum_n E_{on} \sin \tilde{\chi}_n e^{j\tilde{\phi}_n} \end{aligned} \quad (1.5.25)$$

Now, since $\tilde{\chi}_m$ and $\tilde{\chi}_n$ are uncorrelated, only diagonal components of the product-of-sums are non-zero after time averaging. For any pair of indices,

$$\langle \cos \tilde{\chi}_m \cos \tilde{\chi}_n \rangle = \frac{1}{2} \delta_{m,n}$$

where $\delta_{m,n}$ is the Kronecker delta function defined by $\delta_{m,n} = 1$ if $m = n$ and $\delta_{m,n} = 0$ otherwise. The time averages over the sums are therefore

$$\left\langle \sum_m \cos \tilde{\chi}_m \sum_n \cos \tilde{\chi}_n \right\rangle = \frac{N}{2}$$

and

$$\left\langle \sum_m \sin \tilde{\chi}_m e^{-j\tilde{\phi}_m} \sum_n \sin \tilde{\chi}_n e^{j\tilde{\phi}_n} \right\rangle = \frac{N}{2}$$

where the time-average is “long enough” and the absence of the weighting coefficients is irrelevant in the limit. Therefore,

$$\langle S_1 \rangle \rightarrow 0$$

Now consider

$$\begin{aligned} \langle S_2 \rangle &= \langle e_x^* e_y + e_y^* e_x \rangle \\ &= \left\langle \sum_m \cos \tilde{\chi}_m \sum_n \sin \tilde{\chi}_n e^{j\tilde{\phi}_n} \right\rangle + \left\langle \sum_m \sin \tilde{\chi}_m e^{-j\tilde{\phi}_m} \sum_n \cos \tilde{\chi}_n \right\rangle \end{aligned}$$

Unlike (1.5.25), the time averages for both on- and off-diagonal components of $\langle S_2 \rangle$ are zero. Consequently,

$$\langle S_2 \rangle \rightarrow 0, \text{ and } \langle S_3 \rangle \rightarrow 0$$

The only non-vanishing Stokes parameter is S_0 , the total intensity. The time-average intensity I_{incoh} for an incoherently depolarized ray bundle is

$$I_{\text{incoh}} = \langle S_0 \rangle = \sum_n |E_{on}|^2 \quad (1.5.26)$$

and the degree of polarization is $\mathcal{D} = 0$.

A significant extension of the preceding derivation is that multiple incoherent waves need not be narrowband but can be discretely or continuously polychromatic. Relocation of the $\exp(j\omega t)$ term of (1.5.24) within the summations does not change the vanishing time-average nature of $\langle S_{1,2,3} \rangle$. However, polychromatic wave addition can relax the distribution property constraints of $\tilde{\chi}$ and $\tilde{\phi}$ to achieve $\mathcal{D} = 0$.

1.5.3 Pseudo-Depolarized Waves

Pseudo-depolarized waves are waves that start fully polarized and are then depolarized by passage through a birefringent crystal. This configuration is called a Lyot depolarizer. The depolarizer imparts a frequency-dependent polarization on the components of the input light. Unlike natural polarization where each light component uniformly covers the Poincaré sphere, pseudo-depolarized light retains a well-defined pointing direction for each polarization component; these directions vary with frequency.

Consider a single-crystal depolarizer oriented at 45° to a horizontally polarized input state. Denote $\tau = \Delta n L / c$, where Δn is the birefringence, L is the length, and c is the speed of light. The output polarization state is

$$\frac{e^{-j\omega\tau/2}}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{j\omega\tau} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-j\omega\tau/2} & \\ & e^{j\omega\tau/2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (1.5.27)$$

It is readily verified that $S_1 = 0$. The non-vanishing Stokes parameters are

$$S_2 = \cos \omega\tau, \quad S_3 = \sin \omega\tau$$

These parameters are time invariant, but the pointing direction of the Stokes vector changes with frequency. For this example, an arc along a line of longitude on the Poincaré sphere is traced, the subtended arc angle being $\omega\tau$.

More generally, consider the Jones matrix in (1.5.27) operating on a polarized narrowband wave having a continuous spectrum (1.5.16). The spectrum has a modified polarimetric parameter due to the $\exp(j\omega\tau)$ term. The time-domain field components are

$$\begin{aligned}
e_x(t) &= E_o \cos \chi \int_{\Delta\omega} G(\omega) e^{j\omega t} d\omega \\
e_y(t) &= E_o \sin \chi e^{j\phi} \int_{\Delta\omega} G(\omega) e^{j\omega\tau} e^{j\omega t} d\omega
\end{aligned}$$

Following the time-averaging procedure of (1.5.17), the off-diagonal components of J are

$$\begin{aligned}
\langle e_x^* e_y \rangle &= \langle e_x e_y^* \rangle^* \\
&= |E_o|^2 \sin \chi \cos \chi e^{j\phi} \mathcal{I}_G(\tau)
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{I}_G(\tau) &= \int_{\Delta\omega} |G(\omega)|^2 e^{j\omega\tau} d\omega \\
&= \int_{\Delta\omega} |G(\omega)|^2 \cos(\omega\tau) d\omega + j \int_{\Delta\omega} |G(\omega)|^2 \sin(\omega\tau) d\omega
\end{aligned} \tag{1.5.29}$$

The diagonal components of J are

$$\begin{aligned}
\langle e_x^* e_x \rangle &= |E_o|^2 \mathcal{I}_G(0) \cos^2 \chi \\
\langle e_y^* e_y \rangle &= |E_o|^2 \mathcal{I}_G(0) \sin^2 \chi
\end{aligned}$$

Taking these factors into account, the Stokes parameters for a pseudo-depolarized narrowband wave are

$$\langle S \rangle = |E_o|^2 \begin{pmatrix} \mathcal{I}_G(0) \\ \mathcal{I}_G(0) \cos 2\chi \\ |\mathcal{I}_G(\tau)| \sin 2\chi \cos(\phi + \angle \mathcal{I}_G(\tau)) \\ |\mathcal{I}_G(\tau)| \sin 2\chi \sin(\phi + \angle \mathcal{I}_G(\tau)) \end{pmatrix} \tag{1.5.30}$$

Since $|G(\omega)|^2$ is always positive, the sine and cosine integrands in (1.5.29) are the only sources able to decrease $\mathcal{I}_G(\tau)$, see Fig. 1.9. In the limit that $\tau \rightarrow 0$, the oscillatory terms are nearly stationary and $\mathcal{I}_G(\tau) \rightarrow \mathcal{I}_{G,\max}$. Conversely, when there is enough birefringent delay such that $\tau \gg \Delta\omega^{-1}$, the oscillatory terms vary rapidly, resulting in $\mathcal{I}_G(\tau) \rightarrow 0$. For a continuous spectrum, the DOP decreases monotonically with increasing delay-bandwidth product.

It is interesting to note that $\tau \gg \Delta\omega^{-1}$ is a necessary but not sufficient condition for a single-stage Lyot depolarizer to drive $\mathcal{D} \rightarrow 0$. If the input polarization is aligned to an eigenaxis of the crystal then there is no dispersion of the polarization vector over frequency. The DOP remains unity. The DOP is minimized when the input polarization is equally split between axes of the crystal. For this reason, two or more stages are generally used in a Lyot depolarizer.

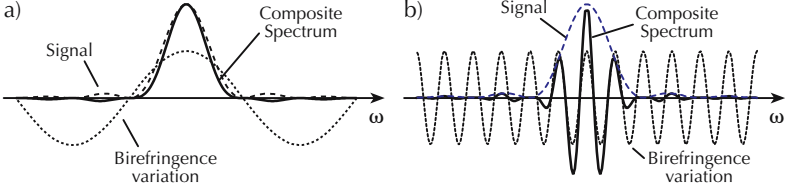


Fig. 1.9. Single-stage Lyot depolarizer impact on a continuous narrowband spectrum. a) Delay τ smaller than inverse signal bandwidth yields slow birefringence variation; the depolarizer has small effect on the integral $\mathcal{I}_G(\tau)$. b) Delay much larger than inverse signal bandwidth; $\mathcal{I}_G(\tau)$ is significantly smaller in this case. As τ increases, $\mathcal{D} \rightarrow 0$ monotonically.

In contrast to the continuous-spectrum case, consider a discrete spectrum described by

$$G(\omega) = \sum_n g_n \delta(\omega - \omega_n)$$

where amplitudes g_n decrease away from ω_o . In this case $\mathcal{I}_G(\tau)$ converts to

$$\mathcal{I}_G(\tau) = \sum_n g_n^2 \exp(j\omega_n \tau) \quad (1.5.31)$$

In contrast with the continuous wave, the integral $\mathcal{I}_G(\tau)$ does not monotonically decrease. Rather, the sum oscillates with a decreasing envelope as τ increases. The components of (1.5.31) are phasors (see Fig. 1.8), and the angle between adjacent phasors is determined by τ . As the phasors fan out for increasing τ eventually all even phasors point along $+1$ and all odd phasors point along -1 . The sum is zero if the spectrum is symmetric. Subsequent doubling of τ points all phasors along $+1$. Such oscillation persists until the birefringence raps around within the linewidth of an individual spectral component.

1.5.4 A Heterogeneous Ray Bundle: Coherent and Incoherent Waves

The preceding sections have studied the DOP for coherent and incoherent ray bundles separately. Signals in a practical system such as a fiber-optic communication link are generally comprised of both coherent and incoherent terms. Coherent light comes from the laser source and incoherent light comes from both the noise of optical amplifiers and depolarization due to polarization-mode dispersion. The degree of polarization for such a heterogeneous mixture is

$$\mathcal{D} = \frac{\sqrt{\langle S_{1-\text{coh}} + S_{1-\text{incoh}} \rangle^2 + \langle S_{2-\text{coh}} + S_{2-\text{incoh}} \rangle^2 + \langle S_{3-\text{coh}} + S_{3-\text{incoh}} \rangle^2}}{\langle S_{0-\text{coh}} + S_{0-\text{incoh}} \rangle}$$

Since the incoherent components have vanishing time-averaged Stokes parameters other than S_0 , only the coherent terms in the numerator survive. When there is no pseudo-depolarization in the system, the expression for the DOP is

$$\mathcal{D} = \frac{I_{\text{coh}}}{I_{\text{coh}} + I_{\text{incoh}}} \quad (1.5.32)$$

but when the spectrum is pseudo-depolarized, cf. (1.5.15), the DOP expression is

$$\mathcal{D} \leq \frac{I_{\text{coh}}}{I_{\text{coh}} + I_{\text{incoh}}} \quad (1.5.33)$$

For instance, when $I_{\text{incoh}} = 0$, pseudo-depolarization can drive the DOP to $\mathcal{D} = 0$. One generally finds expression (1.5.32) in the literature, but the very real effect of polarization mode dispersion in fiber-optic systems leads to the more general expression (1.5.33).

Table 1.1. Polarization States in Equivalent Representations

Polarization state	Jones vector	Stokes vector	Coherency matrix
Linear $\hat{x}$	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
Linear $\hat{y}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Linear at $\pm 45^\circ$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
Right-hand circular	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ j \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & -j \\ j & 1 \end{pmatrix}$
Left-hand circular	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -j \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & j \\ -j & 1 \end{pmatrix}$
Elliptical	$\begin{pmatrix} \cos \chi \\ \sin \chi e^{j\phi} \end{pmatrix}$	$\begin{pmatrix} 1 \\ \cos 2\varepsilon \cos 2\alpha \\ \cos 2\varepsilon \cos 2\alpha \\ \sin 2\varepsilon \end{pmatrix}$	$\begin{pmatrix} c^2 & e^{-j\phi} sc \\ e^{j\phi} sc & s^2 \end{pmatrix}$ $c = \cos \chi$ $s = \sin \chi$
Unpolarized	none	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
All vectors are normalized to a Jones vector of unit length.			

References

1. H. A. Haus, *Waves and Fields in Optoelectronics*. Englewood Cliffs, New Jersey: Prentice–Hall, 1984.
2. H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*. Englewood Cliffs, New Jersey: Prentice–Hall, 1989.
3. B. L. Heffner, “Automated measurement of polarization mode dispersion using Jones matrix eigenanalysis,” *IEEE Photonics Technology Letters*, vol. 4, no. 9, pp. 1066–1068, 1992.
4. S. Huard, *Polarization of Light*. New York: John Wiley & Sons, 1997.
5. R. Jones, “A new calculus for the treatment of optical systems, Part I. description and discussion of the calculus,” *Journal of the Optical Society of America*, vol. 31, no. 7, pp. 488–493, July 1941.
6. ———, “A new calculus for the treatment of optical systems, Part VI. experimental determination of the matrix,” *Journal of the Optical Society of America*, vol. 37, pp. 110–112, 1947.
7. J. A. Kong, *Electromagnetic Wave Theory*. New York: John Wiley & Sons, 1989.
8. P. Mohr and B. Taylor, “CODATA recommended values of the fundamental physical constants,” *Reviews of Modern Physics*, vol. 72, no. 2, pp. 351–495, 2000.
9. K. B. Rochford, *Encyclopedia of Physical Science and Technology*, 3rd ed. San Diego: Academic Press, 2002, ch. Polarization and Polarimetry, pp. 521–538.
10. G. Strang, *Linear Algebra and its Applications*, 3rd ed. New York: Harcourt Brace Jovanovich College Publishers, 1988.

The Spin-Vector Calculus of Polarization

Spin-vector calculus is a powerful tool for representing linear, unitary transformations in Stokes space. Spin-vector calculus attains a high degree of abstraction because rules for vector operations in Stokes space are expressed in vector form; there is no *a priori* reference to an underlying coordinate system. Absence of the underlying coordinate system allows for an elegant, compact calculus well suited for polarization studies.

Spin-vector calculus is well known in quantum mechanics, especially relating to quantized angular momentum. Aso, Frigo, Gisin, and Gordon and Kogelnik have greatly assisted the optical engineering community by adopting this calculus to telecommunications applications [1, 3–5]. The purpose of this chapter is to bring together a complete description of the calculus as found in a variety of disparate sources [2, 3, 5–8], and to tailor the presentation with a vocabulary familiar to the electrical engineer. Tables 2.2 and 2.3 located at the end of the chapter offer a summary of the principal relations.

2.1 Motivation

The purpose of this calculus is to build a geometric interpretation of polarization transformations. The geometric interpretation of polarization states was already developed in §1.4. The Jones matrix, while a direct consequence of Maxwell's equations when light travels through a medium, is a complex-valued 2×2 matrix. This is hard to visualize. The Mueller matrix, however, can be visualized as rotations and length-changes in Stokes space. The spin-vector formalism makes a bilateral connection between the Jones and Mueller matrices.

Of all the possible Jones matrices, two classes predominate in polarization optics: the unitary matrix and the Hermitian matrix. The unitary matrix preserves lengths and imparts a rotation in Stokes space. A retardation plate is described as a unitary matrix. The Hermitian matrix comes from a measurement, such as that of a polarization state. Since all measured values must be

real quantities, the eigenvalues of a Hermitian matrix are real. The projection induced by a polarizer is described as a Hermitian matrix.

Based on the characteristic form (1.4.27) on page 19 of the Mueller matrix for a unitary matrix, defined by $UU^\dagger = I$, one can write

$$\mathbf{J}_U \longrightarrow \mathbf{M}_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & \mathbf{R} & & \\ 0 & & & \end{pmatrix} \quad (2.1.1)$$

where $\mathbf{R}$ is a 3×3 rotation matrix having real-valued entries. Since the polarization transformation through multiple media is described as the product of Jones matrices, one would expect a one-to-one correspondence between multiple unitary matrices and multiple rotation matrices. This would lead to

$$\mathbf{J}_{U_2 U_1} \longrightarrow \mathbf{M}_{U_2 U_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & \mathbf{R}_2 \mathbf{R}_1 & & \\ 0 & & & \end{pmatrix} \quad (2.1.2)$$

This is indeed the case. Moreover, the Mueller matrix representing passage of light through any number of retardation plates always keeps the form of (2.1.1). Rotation matrix $\mathbf{R}$ is therefore a group closed under rotation.

Taking the abstraction one step further, any rotation has an axis of rotation and an angle through which the system rotates. Instead of describing the rotation $\mathbf{R}$ as a 3×3 matrix, it is more general to describe the rotation as a vector quantity: $\mathbf{R} = f(\hat{r}, \varphi)$, where $\hat{r}$ is the rotation axis in Stokes space and φ is the angle of rotation. The vector $\hat{r}$ need not be resolved onto an orthonormal basis to give $\hat{r} = \hat{x} r_x + \hat{y} r_y + \hat{z} r_z$; this operation may be postponed indefinitely. This is in contrast to writing $\mathbf{R}$ as a 3×3 matrix where the underlying orthonormal basis is explicit. Accordingly, $\hat{r}$ exists as a vector in vector space and can undergo operations such as rotation, inner product, and cross product with respect to other vectors.

In parallel to the unitary-matrix case, the Mueller matrix that corresponds to a Hermitian matrix, defined by $H = H^\dagger$, one can write

$$\mathbf{J}_H \longrightarrow \mathbf{M}_H = \begin{pmatrix} & & & \\ & \tilde{\mathbf{H}} & & \\ & & & \end{pmatrix} \quad (2.1.3)$$

This indeed is a tautology. As with the unitary matrices, products of Hermitian matrices in Jones space result in products of Mueller matrices in Stokes space. That is,

$$\mathbf{J}_{H_2 H_1} \longrightarrow \mathbf{M}_{H_2 H_1} = \mathbf{M}_{H_2} \mathbf{M}_{H_1} \quad (2.1.4)$$

All Hermitian operations are closed within the 4×4 Mueller matrix.

As it appears, products of unitary-corresponding Mueller matrices change only entries in the lower-right-hand 3×3 sub-matrix. Inclusion of even a single Hermitian-corresponding Mueller matrix scatters those nine elements into all sixteen matrix positions. This is a non-reversible process. There is, however, a remarkable exception. A traceless Hermitian matrix $\underline{H}$, defined by $\text{Tr}\underline{H} = 0$, has a corresponding Mueller matrix of the form

$$\mathbf{J}_{\underline{H}} \longrightarrow \mathbf{M}_{\underline{H}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathbf{V} & \\ 0 & & & \end{pmatrix} \quad (2.1.5)$$

where $\mathbf{V}$ is a Stokes-space vector having a length and pointing direction. (Note that a rotation operator has unit length, two angles that determine the vector direction, and one angle of rotation. A Stokes vector has a length and two angles that determine the vector direction. Both rotation operator and Stokes vector have three parameters). Arbitrary products of unitary matrix U and traceless Hermitian matrix $\underline{H}$ form an extended closed group in which entries change only in the lower right-hand 3×3 sub-matrix of $\mathbf{M}$.

Throughout this chapter and the chapters on polarization-mode dispersion, one looks for zero trace of Hermitian matrices. If this property is established, then a calculus that includes lossless rotations of vectors can be applied to the system. This calculus is called spin-vector calculus, and is the topic of the present chapter.

2.2 Vectors, Length, and Direction

Physical systems can often be described by the state the system is in at a particular time and position. The span of all possible states for a given system is called a state space. Any particular state represents all the information that one can know about the system at that time and position. Interaction between a physical system and external influences, such as transmission through media or applied force, can change the state. So, there are two categories of study: the description of state, and the transformation of state.

A state that describes wave motion can be represented by a vector with complex scalar entries. The dimensionality of the state vector is determined by the number of states that are invariant to an external influence. That is, the dimension of a state vector equals the number of eigenstates of the system. For polarization, the dimensionality is two. The important properties of a vector space are direction, length, and relative angles. These metrics will form a common theme throughout the following development.

2.2.1 Bra and Ket Vectors

Bra and ket spaces are two equivalent vector spaces that describe the same state space. Bra and ket spaces, or “bracket” space, is a formulation developed

by P. A. M. Dirac and used extensively in quantum mechanics. Bras and kets are vectors with dimension equal to the state dimension. When a bra space and ket space describe the same state vector, the bra and ket are duals of one another. For a state vector $\vec{a}$, the ket is written $|a\rangle$ and the bra is written $\langle a|$. The entries in bra and ket vectors are complex scalar numbers. A ket vector suitable for polarization studies is

$$|a\rangle = \begin{pmatrix} a_x \\ a_y \end{pmatrix} \quad (2.2.1)$$

where a_x and a_y are the components along an orthogonal basis. The entries are complex and accordingly there are four independent parameters contained in (2.2.1). Since the entries are complex, they have magnitude and phase:

$$|a\rangle = \begin{pmatrix} |a_x|e^{j\phi_x} \\ |a_y|e^{j\phi_y} \end{pmatrix} = e^{j\theta} \begin{pmatrix} |a_x| \\ |a_y|e^{j\phi} \end{pmatrix} \quad (2.2.2)$$

where θ is a common phase and ϕ is the phase difference of the second row. In the following the explicit magnitude symbols $|\cdot|$ will be dropped and the intent of magnitude or complex number should be clear from the context. Bra vector $\langle a|$ is said to be the dual of $|a\rangle$ because they are not equal but they describe the same state:

$$|a\rangle \xleftrightarrow{\text{dual}} \langle a|$$

The bra vector $\langle a|$ corresponding to $|a\rangle$ is

$$\langle a| = (a_x^* \ a_y^*) \quad (2.2.3)$$

for every $|a\rangle$. The bra vector is the adjoint ($\dagger$), or complex-conjugate transpose, of the corresponding ket vector:

$$\langle a| = (|a\rangle)^\dagger \quad (2.2.4)$$

Bra and ket vectors obey algebraic additive properties of identity, addition, commutation, and associativity. Identity and addition rules for kets are

$$\begin{array}{ll} \text{identity} & |a\rangle + |0\rangle = |a\rangle \\ \text{addition} & |a\rangle + |b\rangle = |\gamma\rangle \end{array}$$

where $|0\rangle$ is the null ket. Commutation and associativity are straightforward to prove using the matrix representation. A bra or ket vector can also be multiplied by a scalar quantity c :

$$c|a\rangle = |a\rangle c \quad (2.2.5)$$

Physically, the multiplication of a state vector by a scalar does not change the state and therefore the two commute. Operations that have no meaning are

the multiplication of multiple ket vectors or bra vectors. For example, $|b\rangle|a\rangle$ is meaningless.

Finally, it should be understood that state vectors $\langle a|$ and $|a\rangle$ are a more general representation than column and row vectors (2.2.1) and (2.2.3). A state vector is a coordinate-free abstraction that has the properties of length and direction; a row or column vector is a representation of a state vector given a choice of an underlying coordinate system.

2.2.2 Length and Inner Products

Bra and ket vectors have properties of length, phase, and pointing direction. The length of a real-valued vector is a scalar quantity and is determined by the dot product: $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$. For complex-valued bra-ket vectors, the inner product is used to find length of a vector and is determined by multiplying its bra representation $\langle a|$ with its ket representation $|a\rangle$: $\|a\|^2 = \langle a|a\rangle$, where $\|\cdot\|$ is the norm of the vector.

More generally, one wants to measure the length of one vector as projected onto another. The inner product of two different vectors is the product of the bra form of one vector and the ket form of the other: $\langle b|a\rangle$. For real-valued vectors it is clear that $\mathbf{b} \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{b}$. However, for bra-ket vectors, having complex entries, the order of multiplication dictates the sign of the resulting phase. That is,

$$\langle b|a\rangle = |\langle b|a\rangle| e^{j\gamma} \quad (2.2.6a)$$

$$\langle a|b\rangle = |\langle b|a\rangle| e^{-j\gamma} \quad (2.2.6b)$$

The two inner products are related by the complex conjugate:

$$\langle b|a\rangle = (\langle a|b\rangle)^* \quad (2.2.7)$$

The inner product of a bra and ket is a complex-valued scalar. Based on (2.2.7) it is clear that the inner product of a vector onto itself yields a real number, and since the inner product is a measure of length, the real number is positive definite: $\langle a|a\rangle = \text{real number} \geq 0$. Only the null ket has length zero. Any finite ket has a length greater than zero. Throughout the body of the text, polarization vectors are taken to be unit vectors unless otherwise stated. A unit vector has a direction, phase, and unity length. Any vector can be converted to a unit vector by division by its norm:

$$|\tilde{a}\rangle = \frac{1}{\sqrt{\langle a|a\rangle}} |a\rangle \quad (2.2.8)$$

so that

$$\langle \tilde{a}|\tilde{a}\rangle = 1 \quad (2.2.9)$$

In the following the tilde over the vectors will be dropped.

Two vectors are defined as orthogonal to one another when the inner product vanishes:

$$\langle b | a \rangle = 0 \quad (2.2.10)$$

This is an essential inner product used regularly.

When two polarization vectors are resolved onto a common coordinate system,

$$\langle b | a \rangle = b_x^* a_x + b_y^* a_y \quad (2.2.11)$$

Finally, the inner product in matrix representation of a normalized vector is the sum of the component magnitudes squared:

$$\langle a | a \rangle = |a_x|^2 + |a_y|^2 = 1 \quad (2.2.12)$$

2.2.3 Projectors and Outer Products

The inner product measures the length of a vector or the projection of one vector onto another. The result is a complex scalar quantity. In contrast, the outer product retains a vector nature while also producing length by projection. There are two outer product types to study: the projector, having the form $|p\rangle\langle p|$; and the outer product $|p\rangle\langle q|$. The form $|p\rangle\langle q|$ is called a dyadic pair because the vector pair has neither a dot nor cross product between them. In quantum mechanics the projector $|p\rangle\langle p|$ is called the density operator for the state.

Consider a projector that operates on ket $|a\rangle$:

$$|p\rangle\langle p|a\rangle = |p\rangle(\langle p|a\rangle) = c|p\rangle \quad (2.2.13)$$

The quantity $c = \langle p|a\rangle$ is just a complex scalar and commutes with the ket. Operating on $|a\rangle$ the projector measures the length of $|a\rangle$ on $|p\rangle$ and produces a new vector $|p\rangle$.

The effect of the projector is to point along the $|p\rangle$ direction where the length of $|p\rangle$ is scaled by $\langle p|a\rangle$. Projectors work equally well on bras, e.g.

$$\langle a|p\rangle\langle p| = c^*\langle p| \quad (2.2.14)$$

so in fact it should be clear that

$$\langle a|p\rangle\langle p| = (|p\rangle\langle p|a\rangle)^\dagger \quad (2.2.15)$$

The adjoint operator connects the bra and ket forms.

The behavior of the outer product $|p\rangle\langle q|$ is similar to the projector but for the fact that the projection vector and resultant pointing direction differ. The resultant pointing direction depends whether the outer product operates on a ket or a bra. Acting on a ket, the outer product yields

$$|p\rangle\langle q|a\rangle = (\langle q|a\rangle)|p\rangle \quad (2.2.16)$$

whereas acting on a bra of the same vector, the outer product yields

$$\langle a | p \rangle \langle q | = (\langle a | p \rangle) \langle q | \quad (2.2.17)$$

The resultant pointing direction and projected length depends on whether the outer product operates on a bra or ket vector.

In the study of polarization, the outer product is a 2×2 matrix with complex entries:

$$|b\rangle\langle a| = \begin{pmatrix} b_x a_x^* & b_x a_y^* \\ b_y a_x^* & b_y a_y^* \end{pmatrix} \quad (2.2.18)$$

The determinant is

$$\det(|b\rangle\langle a|) = 0 \quad (2.2.19)$$

and therefore the projector is non-invertible. The determinant of an outer product of any dimension is likewise zero. That means the action of $|b\rangle\langle a|$ on a ket is irreversible, which is reasonable because the original direction of the ket is lost. So, while all outer products are operators not all operators are outer products. Operators that are linear combinations of projectors are reversible under the right construction.

In summary, the outer product follows these rules:

equivalence	$(b\rangle\langle a)^\dagger = a\rangle\langle b $
associative	$(b\rangle\langle a) \gamma\rangle = \langle b (\langle a \gamma \rangle)$
trace	$\text{Tr}(b\rangle\langle a) = \langle a b \rangle$
irreversible	$\det(b\rangle\langle a) = 0$

where Tr stands for the trace operation. The trace connects the outer product to the inner product.

2.2.4 Orthonormal Basis

An orthonormal basis is a complete set of orthogonal unit-length axes on which any vector in the space can be resolved. Consider a vector space with N dimensions and orthogonal unit vectors $(|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle)$. The orthogonality requires

$$\langle a_m | a_n \rangle = \delta_{m,n} \quad (2.2.20)$$

where $\delta_{m,n}$ is the Kronecker delta function. Only a vector projected onto itself yields a non-vanishing inner product. When the set is complete, the outer products are closed, where closure is defined as

$$\sum_n |a_n\rangle\langle a_n| = I \quad (2.2.21)$$

When a basis set, or group, is closed, any operation to a member of the group results in another member within the group. Together, (2.2.20–2.2.21) are the two conditions that define an orthonormal basis.

Given an orthonormal basis, any arbitrary vector can be resolved onto the basis using (2.2.21). An arbitrary ket $|s\rangle$ is resolved as

$$|s\rangle = \left(\sum_n |a_n\rangle \langle a_n| \right) |s\rangle = \sum_n c_n |a_n\rangle \quad (2.2.22)$$

where the complex coefficients are given by $c_n = \langle a_n | s \rangle$. The inner product $\langle s | s \rangle$ is the sum of the absolute-value squares of the coefficients c_n :

$$\langle s | s \rangle = \sum_{a'} |c_{a'}|^2 \quad (2.2.23)$$

When $|s\rangle$ is normalized $\sum_{a'} |c_{a'}|^2 = 1$.

2.3 General Vector Transformations

Interaction between a physical system and external influences can change the state of a system. Left unperturbed, a state persists indefinitely. Operators embody the action of external influences and are distinct from the state of the system itself. The bra and ket vectors of the preceding section are two equivalent spaces that describe the same state space. Operators also have two distinct and equivalent spaces that describe the same state transformation. While there is no special notation to represent a “ket” operator or a “bra” operator, equivalence between operator spaces is maintained under

$$X |a\rangle \xleftrightarrow{\text{dual}} \langle a | X^\dagger \quad (2.3.1)$$

$X^\dagger$ is said to be the adjoint operator of X . Care should be taken because the action of $X |a\rangle$ is not the same as $\langle a | X$; these two results are different.

2.3.1 Operator Relations

Operators always act on kets from the left and bras from the right, e.g. $X |a\rangle$ or $\langle a | X$. The expressions $|a\rangle X$ and $X \langle a |$ are undefined. An operator multiplying a ket produces a new ket, and an operator multiplying a bra produces a new bra. In general, an operator changes the state of the system,

$$X |a\rangle = c |b\rangle \quad (2.3.2)$$

where c is a scaling factor induced solely by X . Operators are said to be equal if

$$X |a\rangle = Y |a\rangle \Rightarrow X = Y \quad (2.3.3)$$

Operators obey the following arithmetic properties of addition:

commutative	$X + Y = Y + X$
associative	$X + (Y + Z) = (X + Y) + Z$
distributive	$X(a\rangle + b\rangle) = X a\rangle + X b\rangle$

Operators in general do not commute under multiplication. That is

$$XY \neq YX \quad (2.3.4)$$

In matrix form, only when X and Y are diagonal matrices does $XY = YX$. Other multiplicative properties are

identity	$I a\rangle = a\rangle$
associative	$X(YZ) = (XY)Z$
distributive	$X(Y a\rangle) = XY a\rangle$

All of the above arithmetic properties apply equally well to bra vectors.

The effect X has on state $\vec{a}$ is measured by

$$\text{expectation value of } X \text{ on } \vec{a} = \frac{\langle a|X|a\rangle}{\langle a|a\rangle} \quad (2.3.5)$$

In general, an inner product that encloses an operator gives a complex number:

$$\langle b|(X|a\rangle) = \langle b|X|a\rangle = \text{complex number} \quad (2.3.6)$$

Consider dual constructions, first where $X|a\rangle$ is left-multiplied by $\langle b|$, and second where the dual $\langle a|X^\dagger$ is right-multiplied by $|b\rangle$:

$$\langle b|X|a\rangle = (\langle a|X^\dagger|b\rangle)^* \quad (2.3.7)$$

These two cases are duals of one another and are therefore complex conjugates.

In the study of polarization, operators are represented as 2×2 complex-valued Jones matrices:

$$X = \begin{pmatrix} a e^{j\alpha} & b e^{j\beta} \\ c e^{j\gamma} & d e^{j\eta} \end{pmatrix} \quad (2.3.8)$$

There are eight independent variables contained in the operator. If $\det X \neq 0$, then X is invertible and the action of X can be undone.

The properties of operators are summarized as follows:

operator duality	$X a\rangle \xleftrightarrow{\text{dual}} \langle a X^\dagger$
change of state	$X a\rangle = c b\rangle$ $\langle a X^\dagger = c^* \langle b $
inner product with operator	$\langle b X a\rangle = \text{complex number}$
conjugate relation	$\langle b X a\rangle = \langle a X^\dagger b\rangle^*$
conjugate transpose	$(XY)^\dagger = Y^\dagger X^\dagger$

Just as an arbitrary ket can be resolved onto an orthonormal basis, an arbitrary operator X can be resolved onto a set of projection operators formed on the orthonormal basis. Applying the closure relation (2.2.21) yields

$$\begin{aligned} X &= \left(\sum_m |a_m\rangle \langle a_m| \right) X \left(\sum_n |a_n\rangle \langle a_n| \right) \\ &= \sum_n \sum_m |a_m\rangle \langle a_m| X |a_n\rangle \langle a_n| \end{aligned} \quad (2.3.9)$$

The indexing symmetry of (2.3.9) looks like a matrix with $\langle a_m|X|a_n\rangle$ as the (m, n) entry. For polarization, the matrix is 2×2 and looks like

$$\sum_n \sum_m |a_m\rangle \langle a_m| X |a_n\rangle \langle a_n| \mapsto \begin{pmatrix} \langle a_1|X|a_1\rangle & \langle a_1|X|a_2\rangle \\ \langle a_2|X|a_1\rangle & \langle a_2|X|a_2\rangle \end{pmatrix} \quad (2.3.10)$$

The resolved form of X in (2.3.10) will become particularly simple in discussion of Hermitian and unitary matrices.

2.4 Eigenstates, Hermitian and Unitary Operators

Many physical systems exhibit particular states that are not transformed by interaction with the system. These invariant states are called eigenstates of the system. In spin-vector calculus, operators embody the influence of a phenomena. The eigenvectors of an operator are the eigenstates of the system. When an operator X acts on its own eigenstate $\vec{a}$,

$$X |a_1\rangle = a_1 |a_1\rangle \quad (2.4.1a)$$

$$\langle a_1|X^\dagger = a_1^* \langle a_1| \quad (2.4.1b)$$

the state of the system is unaltered but for a scaling factor a_1 . The scale factor is the eigenvalue of X associated with eigenstate $\vec{a}_1$. Each eigenvector has an associated eigenvalue, and a well-conditioned matrix has as many eigenvectors as rows in the matrix or, equivalently, dimensions in the state space.

The eigenvectors of Hermitian and unitary operators are orthogonal when the associated eigenvalues are distinct. The eigenvalues of a Hermitian operator are real-valued scalars, and the eigenvalues of a unitary operator are complex exponential scalars. A Hermitian or unitary operator X having N eigenkets $(|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle)$ and associated eigenvalues $(a_1, a_2, \dots, a_N)$ produces the series of inner products

$$\langle a_m|X^\dagger X|a_n\rangle = |a_m|^2 \delta_{m,n} \quad (2.4.2)$$

The operator $X^\dagger X$ scales each axis by a different amount, but does not rotate nor create reflection of the original basis. The eigenvalues of operator X are related to the determinant and trace by

$$\det(X) = a_1 a_2 \cdots a_N \quad (2.4.3a)$$

$$\text{Tr}(X) = a_1 + a_2 + \cdots + a_N \quad (2.4.3b)$$

Since the eigenvalues of a Hermitian matrix are real, its determinant and trace are real.

2.4.1 Hermitian Operators

The defining property of a Hermitian operator is

$$H^\dagger = H \quad (2.4.4)$$

The associated Hermitian matrix in polarization studies has only four independent variables: three amplitudes and one phase. This contrasts with the general Jones matrix (2.3.8) which has eight.

The eigenvectors of H form a complete orthonormal basis and the eigenvalues are real. That the eigenvalues are real is proved from the following difference:

$$\begin{aligned} \langle a_n | (H^\dagger - H) | a_m \rangle &= (a_n^* - a_m) \langle a_n | a_m \rangle \\ &= 0 \end{aligned} \quad (2.4.5)$$

Non-trivial solutions are found when neither vector is null. The eigenvectors may be the same or different. Consider first when the eigenvectors are the same. Since $\langle a_n | a_n \rangle \neq 0$, $(a_n^* - a_n) = 0$ and the eigenvalue is real. Consider when the eigenvectors are different. Unless $a_m = a_n$, in which case the eigenvectors are not linearly independent, it must be the case that $\langle a_n | a_m \rangle = 0$. All eigenvalues are therefore real. Hermitian operators H scale its own basis set:

$$\langle a_m | H^\dagger H | a_n \rangle = a_m^2 \delta_{m,n} \quad (2.4.6)$$

When $\det(H) \neq 0$, H is invertible and the action of H on the state of a system is reversible.

The expansion of H onto its own basis generates a diagonal eigenvalue matrix. Under construction (2.3.9) the expansion yields

$$\begin{aligned} H &= \sum_n \sum_m |a_m\rangle \langle a_m | H | a_n \rangle \langle a_n | \\ &= \sum_m a_m |a_m\rangle \langle a_m| \end{aligned} \quad (2.4.7)$$

where $\langle a_m | H | a_n \rangle = a_n \delta_{m,n}$. The orthonormal expansion is written in matrix form as $H = S \Lambda S^{-1}$, where S is a square matrix whose columns are the eigenvectors of H and Λ is a diagonal matrix whose entries are the associated eigenvalues. Schematically,

$$S = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{pmatrix}, \text{ and } \Lambda = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_N \end{pmatrix} \quad (2.4.8)$$

where $|a_n\rangle = v_n^T$.

2.4.2 Unitary Operators

The defining property of a unitary operator is

$$T^\dagger T = I \quad (2.4.9)$$

Acting on its orthogonal eigenvectors $|a_n\rangle$, the unitary operator preserves the unity basis length:

$$\langle a_m | T^\dagger T | a_n \rangle = \delta_{m,n} \quad (2.4.10)$$

Taking the determinant of both sides of (2.4.9) gives $\det(T^\dagger T) = 1$. Since the determinant of a product is the product of the determinants and the adjoint operator preserves the norm, the determinant of T must be

$$\det(T) = e^{j\theta} \quad (2.4.11)$$

Since the the determinant is the product of eigenvalues, the eigenvalues of T must themselves be complex exponentials and, accounting for (2.4.10), they must have unity magnitude. Therefore T acting on an eigenvector yields

$$T |a_n\rangle = e^{-j\alpha_n} |a_n\rangle \quad (2.4.12)$$

The eigenvalues of T lie on the unit circle in the complex plane.

A special form of T exists where the determinant is unity. This special form is denoted U and is characterized by $\det(U) = +1$. To transform from T to U , the common phase factor $\beta = \exp(j\theta/N)$ must be extracted from each eigenvalue of T , where N is the dimensionality of the operator. The T and U forms are thereby related:

$$T = e^{j\beta} U \quad (2.4.13)$$

It should be noted that when $\det(U) = -1$, a reflection is present along an odd number of axes in the basis set of U .

The eigenvalue equation for U is

$$U |a_n\rangle = e^{-j\phi_n} |a_n\rangle \quad (2.4.14)$$

U expands on its own basis set in the same way H expands (2.4.7):

$$U = \sum_m e^{-j\phi_m} |a_m\rangle \langle a_m| \quad (2.4.15)$$

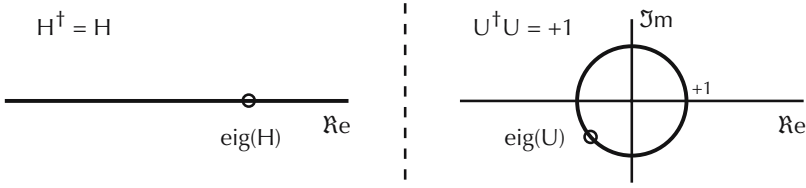


Fig. 2.1. Eigenvalue loci of H and U . Left: eigenvalues of H lie on the real number line. Right: eigenvalues of U lie on the unit circle in the complex plane.

This orthonormal expansion has the matrix analogue of $U = S \exp(-j\Lambda)S^{-1}$, where the diagonal matrix is

$$\exp(-j\Lambda) = \begin{pmatrix} e^{-j\phi_1} & & & \\ & e^{-j\phi_2} & & \\ & & \ddots & \\ & & & e^{-j\phi_N} \end{pmatrix} \quad (2.4.16)$$

and S is the same form as (2.4.8).

2.4.3 Connection between Hermitian and Unitary Matrices

The connection between Hermitian and unitary operators is quite intimate. Figure 2.1 illustrates the eigenvalue domains for H and U . The eigenvalues of H lie on the real number line, while those of U lie on the unit circle in the complex plane. Multiplying the eigenvalues of H by $-j$ and taking the exponential, one can construct the eigenvalues of U . Note that the eigenvalues of U are cyclic, so only the real number line modulo 2π is significant.

Based on the operator expansions of the preceding sections, one has

$$H = S\Lambda_H S^{-1} \quad (2.4.17a)$$

$$U = R e^{-j\Lambda_U} R^{-1} \quad (2.4.17b)$$

Since in general $\exp(S\Lambda S^{-1}) = S \exp(\Lambda) S^{-1}$, the H and U operators may be connected as

$$U = e^{-jH} \implies S e^{-j\Lambda_U} S^{-1} = S e^{-j\Lambda_H} S^{-1} \quad (2.4.18)$$

For every Hermitian operator H there is an associated unitary operator U that shares the same basis set and has eigenvalues related through the complex exponential.

2.4.4 Similarity Transforms

Frequently one has Hermitian operator H and orthonormal basis $|p_n\rangle$ that are not aligned. That is, the eigenvectors $|a_n\rangle$ of H are not parallel to vectors $|p_n\rangle$. Expansion of H into $|p_n\rangle$ using the expansion expression (2.3.10)

generates a matrix that is not diagonal. However, the expansion matrix can be diagonalized by rotating basis $|p_n\rangle$ into $|a_n\rangle$. The unitary matrix does this operation. Taking advantage of $U^\dagger U = 1$, one can write

$$\begin{aligned}\langle p|H|p\rangle &= \langle p|U^\dagger U H U^\dagger U|p\rangle \\ &= \langle a|U H U^\dagger|a\rangle \\ &= \langle a|H_T|a\rangle\end{aligned}\tag{2.4.19}$$

Since (2.4.19) holds for any choice of initial basis $|p_n\rangle$, the operators

$$H_T = U H U^\dagger\tag{2.4.20}$$

must be equal. Equation (2.4.20) is known as a similarity transform. Both the determinant and trace of H are independent of basis; that is

$$\begin{aligned}\det(H_T) &= \det(U) \det(H) \det(U^\dagger) \\ &= \det(H)\end{aligned}\tag{2.4.21}$$

and

$$\text{Tr}(H_T) = \text{Tr}(U H U^\dagger)\tag{2.4.22}$$

The trace is always preserved under a similarity transform.

2.4.5 Construction of General Unitary Matrix

The characteristic property $T^\dagger T = 1$ of unitary matrices restricts the eight independent variables generally available for a polarization operator (2.3.8). Derivation of the restrictions generates a general form of the unitary matrix.

First consider U , where $\det(U) = +1$ and . Substitution of X (2.3.8) for U in $U^\dagger U = 1$ generates the following requirements:

$$\begin{aligned}|a|^2 + |b|^2 &= 1, \quad |a|^2 = |d|^2, \quad |b|^2 = |c|^2, \\ ac e^{-j(\alpha-\gamma)} + bd e^{-j(\beta-\eta)} &= 0\end{aligned}\tag{2.4.23}$$

The determinant requirement generates

$$ad e^{j(\alpha+\eta)} - bc e^{j(\beta+\gamma)} = 1\tag{2.4.24}$$

The amplitude restrictions in (2.4.23) are satisfied by $a = \cos \kappa$ and $b = \sin \kappa$. There remains, however, a sign degeneracy in that $c = \pm b$ and $d = \pm a$. This degeneracy is insignificant in that any choice flows through the restriction criteria and produces the same matrix form.

Combination of the last equation in (2.4.23) and (2.4.24) generates two restrictions on the phase:

$$\begin{aligned} e^{j(\gamma+\beta)} &= -e^{j(\alpha+\eta)} \\ e^{j\alpha} &= e^{-j\eta} \\ e^{j\gamma} &= -e^{-j\beta} \end{aligned} \quad (2.4.25)$$

There are only two independent phases. Combining all of the above restrictions, the general matrix form of U is written

$$U = \begin{pmatrix} e^{j\alpha} \cos \kappa & -e^{j\beta} \sin \kappa \\ e^{-j\beta} \sin \kappa & e^{-j\alpha} \cos \kappa \end{pmatrix} \quad (2.4.26)$$

There are three independent variables in U : one amplitude and two phases. The fourth independent variable has been suppressed because of the arbitrary selection $\det(U) = +1$. The unitary matrix T includes the common phase:

$$T = e^{j\phi} \begin{pmatrix} e^{j\alpha} \cos \kappa & -e^{j\beta} \sin \kappa \\ e^{-j\beta} \sin \kappa & e^{-j\alpha} \cos \kappa \end{pmatrix} \quad (2.4.27)$$

where there are now four independent variables: one amplitude and three phases.

The Cayley-Klein form of U , using complex entries a and b , is

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \quad (2.4.28)$$

The inverse of the unitary matrix is $U^{-1}(a, b) = U(a^*, -b)$.

2.4.6 Group Properties of SU(2)

For polarization studies, unitary operators are 2×2 square matrices with complex entries. The group defined by multiplication operations of 2×2 unitary matrices is called $U(2)$, and the subgroup of unitary matrices where $\det(U) = +1$ is called $SU(2)$, “S” for special. The group properties for multiplication are

Identity	$UI = U$
Closure	$U_1 U_2 = U_3$
Inverse	$U^{-1}U = I$
Associativity	$(U_1 U_2)U_3 = U_1(U_2 U_3)$

where in all cases $U_{1,2,3} \in SU(2)$. $SU(2)$ is closed under these four operations.

2.5 Vectors Cast in Jones and Stokes Spaces

Thus far, state spaces and operators have been presented without restriction on their dimensionality. The properties of these vectors and matrices have been studied in general with passing observations about polarization-specific points of interest. At this stage the scope of presentation will concentrate on the study of polarization so that a formal connection between Jones and Stokes space can be established. The tools developed in the preceding sections are essential to make the bilateral connections that follow.

Recall from (1.4.5) on page 13 that a polarization vector is written as

$$|s\rangle = E_o e^{j\theta} \begin{pmatrix} \cos \chi \\ \sin \chi e^{j\phi} \end{pmatrix} \quad (2.5.1)$$

where E_o is real. There are two polar angles in (2.5.1): χ and ϕ . The common phase $\exp(j\theta)$ is lost on conversion to Stokes space.

There are seven measurements necessary to determine the polarization ellipse uniquely. The first measurement is for the overall intensity and the remaining measurements project the ellipse onto six different reference axes. The formal construction of a projection matrix is necessary at this point.

Consider points along two orthogonal axes and their projection onto a line L inclined by angle θ that passes through the origin. As illustrated in Fig. 2.2, the coordinate $(1, 0)$ is projected to point a on line L . The coordinates of a as measured along the two orthogonal axes are $(\cos^2 \theta, \sin \theta \cos \theta)$. After a similar analysis for the coordinate $(0, 1)$, one can construct the projection matrix $\mathcal{P}$:

$$\mathcal{P} = \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} \quad (2.5.2)$$

It is clear that $\det(\mathcal{P}) = 0$; $\mathcal{P}$ is non-invertable and its action is irreversible. There is loss of information after projection. Moreover, $\mathcal{P}^2 = \mathcal{P}$, so once the projection is taken, subsequent projections along the same line L do not change the result.

2.5.1 Complete Measurement of the Polarization Ellipse

There are seven measurements necessary for complete determination of the polarization ellipse. The first measurement is one of total intensity, the remaining six measurements are projections. The projections are defined in pairs and the difference values are associated with the Stokes coordinates. The result of an intensity of the polarization ellipse is the inner product

$$\langle s | s \rangle = \langle s | \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} | s \rangle \quad (2.5.3)$$

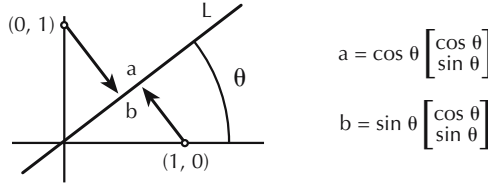


Fig. 2.2. Projection of unit coordinates $(1, 0)$ and $(0, 1)$ onto line L , which is inclined by angle θ and passes through the origin. The projected coordinates are tabulated on the right. A second projection of a and b onto L does not change the coordinates of a and b . The projection operator is non-invertible.

Without loss of generality, $\langle s | s \rangle = 1$ in the following.

The first projection pair is $\theta = 0$ and $\theta = \pi/2$. The projection measure comes from the inner products

$$P_0 = \langle s | \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} | s \rangle, \quad \text{and} \quad P_{\pi/2} = \langle s | \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} | s \rangle \quad (2.5.4)$$

The Stokes parameter s_1 is defined as

$$s_1 = P_0 - P_{\pi/2} \quad (2.5.5)$$

Substitution of (2.5.4) into (2.5.5) makes

$$s_1 = \langle s | \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} | s \rangle \quad (2.5.6)$$

The second projection pair is $\theta = \pm\pi/4$. These projections produce

$$P_{+\pi/4} = \frac{1}{2} \langle s | \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} | s \rangle, \quad \text{and} \quad P_{-\pi/4} = \frac{1}{2} \langle s | \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} | s \rangle \quad (2.5.7)$$

The Stokes parameter s_2 is defined as

$$s_2 = P_{+\pi/4} - P_{-\pi/4} \quad (2.5.8)$$

which makes

$$s_2 = \langle s | \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} | s \rangle \quad (2.5.9)$$

The last projection requires the measurement of the ellipse circularity. By convention, right-hand circular polarization rotates in the counter-clockwise (ccw) direction when observed along the $-\hat{z}$ direction (looking into the light). The right-hand circular polarization vector is

$$|s\rangle_R = \begin{pmatrix} 1 \\ j \end{pmatrix} \quad (2.5.10)$$

The ccw vector needs mapping to the $\theta = 0$ axis; a unitary transform does the rotation. The right- and left-hand projections are calculated via

$$P_R = \langle s | U^\dagger \mathcal{P}_0 U | s \rangle \quad (2.5.11a)$$

$$P_L = \langle s | U^\dagger \mathcal{P}_{\pi/2} U | s \rangle \quad (2.5.11b)$$

The unitary matrix

$$U = \frac{1}{2} \begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix} \quad (2.5.12)$$

maps right-hand circular polarization to the $\theta = 0$ axis:

$$\frac{1}{2} \begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix} \begin{pmatrix} 1 \\ j \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2.5.13)$$

Substituting (2.5.2) and (2.5.12) into (2.5.11) produces

$$P_R = \frac{1}{4} \langle s | \begin{pmatrix} 1 & -j \\ j & 1 \end{pmatrix} | s \rangle, \text{ and } P_L = \frac{1}{4} \langle s | \begin{pmatrix} 1 & j \\ -j & 1 \end{pmatrix} | s \rangle \quad (2.5.14)$$

The Stokes parameter s_3 is defined as

$$s_3 = P_R - P_L \quad (2.5.15)$$

which makes

$$s_3 = \langle s | \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} | s \rangle \quad (2.5.16)$$

From these seven measurements one can transform from a ket in Jones space to three Stokes coordinates that lie on the unit sphere:

$$|s\rangle \implies \hat{s} \quad (2.5.17)$$

where the vector $\hat{s}$ is a column vector defined by $\hat{s} = (s_1, s_2, s_3)^T$.

This completes the measurement of the polarization ellipse. From these measurements the polarimetric angles χ and ϕ are uniquely determined. These measurements combined with the definition of the Stokes parameters generate the Pauli spin matrices. This is the topic of the next section.

2.5.2 Pauli Spin Matrices

The Pauli spin matrices connect Jones to Stokes spaces through the projection measurements of the preceding section. The identity Pauli matrix is

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.5.18)$$

The Pauli spin matrices are¹

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} \quad (2.5.19)$$

The spin matrices are both Hermitian and unitary:

$$\sigma_k^\dagger = \sigma_k \quad \text{and} \quad \sigma_k^\dagger \sigma_k = I \quad (2.5.20)$$

The determinants of the spin matrices are -1 and the traces zero:

$$\det(\sigma_k) = -1 \quad \text{and} \quad \text{Tr}(\sigma_k) = 0 \quad (2.5.21)$$

A spin matrix multiplied by itself yields

$$\sigma_k \sigma_k = I \quad (2.5.22)$$

and multiplied by other matrices gives

$$\sigma_i \sigma_j = -\sigma_j \sigma_i = j \sigma_k \quad (2.5.23)$$

where the indices of the multiplication table (i, j, k) are cyclic permutations of $(1, 2, 3)$.

Each Stokes coordinate of a polarization state $|s\rangle$ is calculated by inserting the associated Pauli matrix into the inner product $\langle s | \cdot | s \rangle$. The individual Stokes coordinates are

$$s_k = \langle s | \sigma_k | s \rangle \quad (2.5.24)$$

This is shorthand for the projection-difference measurements of (2.5.6, 2.5.9, 2.5.16). Since the spin matrices are Hermitian, the Stokes coordinates s_k are real, signed quantities. Moreover, since $\det(\sigma_k) = -1$ and the Jones vector $|s\rangle$ is assumed to be normalized, s_k is bounded by $-1 \leq s_k \leq +1$. The proof that the norm of $\hat{s}$ is unity, $|\hat{s}| = 1$, is shown below.

2.5.3 The Pauli Spin Vector and the Bilateral Connection Between Jones and Stokes Vectors

The Pauli spin vector condenses further the notation of (2.5.24). The spin vector is defined as

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} \quad (2.5.25)$$

¹ In physics texts the z direction is denoted by the σ_1 spin matrix while here it is denoted by σ_3 . Historically, the Pauli spin matrices describe electron spin, which is either up or down in the “ z ” direction. In polarization optics, one usually thinks of a horizontal polarization state aligned to the “ x ” axis.

where $\vec{\sigma}$ is a vector of matrices. The vector of Stokes coordinates $\hat{s}$ is derived from the Jones vector $|s\rangle$ using the spin vector:

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \langle s | \sigma_1 | s \rangle \\ \langle s | \sigma_2 | s \rangle \\ \langle s | \sigma_3 | s \rangle \end{pmatrix} \quad (2.5.26)$$

More concisely,

$$\hat{s} = \langle s | \vec{\sigma} | s \rangle \quad (2.5.27)$$

This is the most compact way to map Jones vectors to Stokes vectors.

The reciprocal connection is made through an eigenvalue equation whose parameters are the Stokes vector $\hat{s}$ and the spin vector. First, observe that the spin vector behaves both as a 3×1 vector and as a 2×2 matrix, depending on the context. Above shows the spin vector acting as a 3×1 vector. Alternatively, the dot product of $\hat{s}$ with the spin vector yields

$$\begin{aligned} \hat{s} \cdot \vec{\sigma} &= s_1 \sigma_1 + s_2 \sigma_2 + s_3 \sigma_3 \\ &= \begin{pmatrix} s_1 & s_2 - js_3 \\ s_2 + js_3 & -s_1 \end{pmatrix} \end{aligned} \quad (2.5.28)$$

$\hat{s} \cdot \vec{\sigma}$ in this case is a 2×2 Jones matrix and, since the coefficients s_k are real, $\hat{s} \cdot \vec{\sigma}$ is Hermitian: $(\hat{s} \cdot \vec{\sigma})^\dagger = (\hat{s} \cdot \vec{\sigma})$.

Next, recall from §2.2.3 that the trace operation connects the projector with its inner product: $\text{Tr}(|s\rangle\langle s|) = \langle s | s \rangle$. Since the trace of each Pauli matrix is zero it is also true that $\text{Tr}(\hat{s} \cdot \vec{\sigma}) = 0$. For a normalized state vector such that $\langle s | s \rangle = 1$, one can construct the projector for ket $|s\rangle$ in terms of the spin vector:

$$|s\rangle\langle s| = \frac{1}{2} (I + \hat{s} \cdot \vec{\sigma}) \quad (2.5.29)$$

Subsequent multiplication on the right by $|s\rangle$ generates the eigenvalue equation

$$\hat{s} \cdot \vec{\sigma} |s\rangle = |s\rangle \quad (2.5.30)$$

This is the most compact way to map Stokes vectors to Jones vectors. The eigenvector of $\hat{s} \cdot \vec{\sigma}$ associated with eigenvalue +1 generates the Jones vector $|s\rangle$ from Stokes vector $\hat{s}$.

2.5.4 Spin-Vector Identities

Vector operations that include spin-vectors do not yield to the same intuition one is accustomed to with “normal” vectors. For example, while one is quite familiar with $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{a}) = 0$, since $\mathbf{a}$ is orthogonal to $\mathbf{b} \times \mathbf{a}$, the spin-vector analogue produces $\vec{\sigma} \cdot (\vec{a} \times \vec{\sigma}) = -2j(\vec{a} \cdot \vec{\sigma})$. The difference comes from

the cyclic multiplication table for spin-vectors (2.5.22–2.5.23), where the sign of a product is determined by the order in which the spin-vectors appear.

The purpose of the following identity tabulation is to provide reductions in the order k of $(\vec{\sigma})^k$. For the following identities, $\vec{a}$ and $\vec{b}$ are real-valued 3×1 vectors and $\vec{\sigma}$ is the spin vector. Real vectors $\vec{a}$ and $\vec{b}$ are not interchangeable with the spin vector $\vec{\sigma}$.

Identities of order $(\vec{\sigma})^0$ and $(\vec{\sigma})$:

$$\vec{a} \cdot \vec{a} = a^2 \quad (2.5.31)$$

$$\vec{a} \cdot \vec{\sigma} = \vec{\sigma} \cdot \vec{a} \quad (2.5.32)$$

$$\vec{a}(\vec{a} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{\sigma})\vec{a} \quad (2.5.33)$$

Identities of order $(\vec{\sigma})^2$:

$$\vec{\sigma} \cdot \vec{\sigma} = 3I \quad (2.5.34)$$

$$\vec{\sigma}(\vec{a} \cdot \vec{\sigma}) = \vec{a}I + j\vec{a} \times \vec{\sigma} \quad (2.5.35)$$

$$(\vec{a} \cdot \vec{\sigma})\vec{\sigma} = \vec{a}I - j\vec{a} \times \vec{\sigma} \quad (2.5.36)$$

$$(\vec{a} \cdot \vec{\sigma})(\vec{a} \cdot \vec{\sigma}) = a^2 I \quad (2.5.37)$$

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b})I + (j\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (2.5.38)$$

$$[(\vec{a} \cdot \vec{\sigma}), \vec{\sigma}] = -2j\vec{a} \times \vec{\sigma} \quad (2.5.39)$$

$$\{(\vec{a} \cdot \vec{\sigma}), \vec{\sigma}\} = 2\vec{a} I \quad (2.5.40)$$

$$\left[(\vec{a} \cdot \vec{\sigma}), (\vec{b} \cdot \vec{\sigma}) \right] = 2(j\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (2.5.41)$$

$$\left\{ (\vec{a} \cdot \vec{\sigma}), (\vec{b} \cdot \vec{\sigma}) \right\} = 2(\vec{a} \cdot \vec{b}) I \quad (2.5.42)$$

$$\vec{\sigma} \cdot (j\vec{a} \times \vec{\sigma}) = 2(\vec{a} \cdot \vec{\sigma}) \quad (2.5.43)$$

$$(j\vec{a} \times \vec{\sigma}) \cdot \vec{\sigma} = -2(\vec{a} \cdot \vec{\sigma}) \quad (2.5.44)$$

$$(j\vec{a} \times \vec{\sigma})(\vec{a} \cdot \vec{\sigma}) = a^2 \vec{\sigma} - \vec{a}(\vec{a} \cdot \vec{\sigma}) \quad (2.5.45)$$

$$(\vec{a} \cdot \vec{\sigma})(j\vec{a} \times \vec{\sigma}) = \vec{a}(\vec{a} \cdot \vec{\sigma}) - a^2 \vec{\sigma} \quad (2.5.46)$$

where $[A, B] = AB - BA$ is the commutator and $\{A, B\} = AB + BA$ is the anti-commutator.

Identities of order $(\vec{\sigma})^3$:

$$\vec{\sigma} \cdot ((\vec{a} \cdot \vec{\sigma})\vec{\sigma}) = (\vec{\sigma}(\vec{a} \cdot \vec{\sigma})) \cdot \vec{\sigma} = -(\vec{a} \cdot \vec{\sigma}) \quad (2.5.47)$$

$$\vec{\sigma} \cdot (\vec{\sigma}(\vec{a} \cdot \vec{\sigma})) = ((\vec{a} \cdot \vec{\sigma})\vec{\sigma}) \cdot \vec{\sigma} = 3(\vec{a} \cdot \vec{\sigma}) \quad (2.5.48)$$

$$(\vec{a} \cdot \vec{\sigma})\vec{\sigma}(\vec{a} \cdot \vec{\sigma}) = 2\vec{a}(\vec{a} \cdot \vec{\sigma}) - a^2 \vec{\sigma} \quad (2.5.49)$$

Identity of order $(\vec{\sigma})^n$:

$$(\vec{a} \cdot \vec{\sigma})^n = \begin{cases} a^n & n \text{ even} \\ a^{n-1}(\hat{a} \cdot \vec{\sigma}) & n \text{ odd} \end{cases} \quad (2.5.50)$$

Finally, there are identities that relate to inner products taken with various forms of the spin vector. These identities are as follows:

$$\langle s | \vec{a} \cdot \vec{\sigma} | s \rangle = \vec{a} \cdot \langle s | \vec{\sigma} | s \rangle = \vec{a} \cdot \hat{s} \quad (2.5.51)$$

$$\langle s | \vec{a} \times \vec{\sigma} | s \rangle = \vec{a} \times \langle s | \vec{\sigma} | s \rangle = \vec{a} \times \hat{s} \quad (2.5.52)$$

$$\langle s | R\vec{\sigma} | s \rangle = R\langle s | \vec{\sigma} | s \rangle = R\hat{s} \quad (2.5.53)$$

where $\vec{a}$ and $\vec{b}$ are arbitrary vectors in Stokes space, a is the length of $\vec{a}$, and R is a 3x3 matrix. Identity (2.5.53) is always a source of confusion, so it is repeated explicitly:

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} \langle s | \sigma_1 | s \rangle \\ \langle s | \sigma_2 | s \rangle \\ \langle s | \sigma_3 | s \rangle \end{pmatrix} = \begin{pmatrix} \langle s | r_{11}\sigma_1 + r_{12}\sigma_2 + r_{13}\sigma_3 | s \rangle \\ \langle s | r_{21}\sigma_1 + r_{22}\sigma_2 + r_{23}\sigma_3 | s \rangle \\ \langle s | r_{31}\sigma_1 + r_{32}\sigma_2 + r_{33}\sigma_3 | s \rangle \end{pmatrix}$$

where $R\hat{s}$ on the left and $\langle s | R\vec{\sigma} | s \rangle$ on the right.

2.5.5 Conservation of Length

Expressions (2.5.27) and (2.5.30) complete the bilateral connection between Jones and Stokes vectors. Length must be conserved, of course, and this is verified now.

The Stokes-vector length is derived from the product $\hat{s} \cdot \hat{s} = s_1^2 + s_2^2 + s_3^2$. Consider one coordinate alone,

$$s_k^2 = \langle s | \sigma_k | s \rangle \langle s | \sigma_k | s \rangle \quad (2.5.54)$$

Substitution of the projector $|s\rangle\langle s|$, (2.5.29), for the innermost term gives

$$s_k^2 = \frac{1}{2} \langle s | s \rangle + \frac{1}{2} \langle s | \sigma_k (\hat{s} \cdot \vec{\sigma}) \sigma_k | s \rangle \quad (2.5.55)$$

The sum of all three terms gives

$$s_1^2 + s_2^2 + s_3^2 = \frac{3}{2} \langle s | s \rangle + \frac{1}{2} \langle s | \vec{\sigma} \cdot ((\hat{s} \cdot \vec{\sigma})\vec{\sigma}) | s \rangle \quad (2.5.56)$$

The spin-vector identity (2.5.48) simplifies (2.5.56):

$$s_1^2 + s_2^2 + s_3^2 = \frac{3}{2} \langle s | s \rangle - \frac{1}{2} \langle s | \hat{s} \cdot \vec{\sigma} | s \rangle = \langle s | s \rangle \quad (2.5.57)$$

Thus, $\|\vec{s}\|^2 = \langle s | s \rangle$. Length is clearly preserved in this direction. For the reverse mapping, construction of (2.5.30) without the assumption $\langle s | s \rangle = 1$ produces

$$\vec{s} \cdot \vec{\sigma} |s\rangle = \langle s | s \rangle |s\rangle \longrightarrow \hat{s} \cdot \vec{\sigma} |s\rangle = |s\rangle \quad (2.5.58)$$

That length is conserved over the bilateral connections is thus established.

There is, however, one piece of information that is lost in the mapping from Jones to Stokes. Since Stokes coordinates are derived from intensity measurements, the common phase of the Jones vector is lost. Transformation from Stokes back to Jones does not reintroduce this phase. Any Jones vector constructed from a Stokes vector is accurate to the “true” Jones vector to within an arbitrary common phase. Physically this just means that the absolute time it took for the light to travel from its source to the observer cannot be determined from Stokes measurements.

2.5.6 Orthogonal Polarization States

For every polarization state $|s_+\rangle$ there is a unique polarization state $|s_-\rangle$ such that $\langle s_- | s_+ \rangle = 0$. These states $|s_+\rangle$ and $|s_-\rangle$ are orthogonal. Given $|s_+\rangle$ how does one can construct the orthogonal state $|s_-\rangle$ and its Stokes equivalent?

From (2.5.30) on page 56 one writes

$$\langle s_- | s_+ \rangle = \left(\langle s_- | (\hat{s}_- \cdot \vec{\sigma})^\dagger \right) \left((\hat{s}_+ \cdot \vec{\sigma}) | s_+ \rangle \right) = 0 \quad (2.5.59)$$

Since $(\hat{s}_- \cdot \vec{\sigma})$ is Hermitian, (2.5.59) is rewritten using spin-vector identity (2.5.38) as

$$\langle s_- | s_+ \rangle = (\hat{s}_- \cdot \hat{s}_+) \langle s_- | s_+ \rangle + j \langle s_- | (\hat{s}_- \times \hat{s}_+) \cdot \vec{\sigma} | s_+ \rangle \quad (2.5.60)$$

As $\langle s_- | s_+ \rangle = 0$, (2.5.60) requires that $(\hat{s}_- \times \hat{s}_+) = 0$. There are two orientations that produce $(\hat{s}_- \times \hat{s}_+) = 0$: $\hat{s}_- \cdot \hat{s}_+ = \pm 1$. If $\hat{s}_- \cdot \hat{s}_+ = +1$, then $\langle s_- | s_+ \rangle = 1$, contradicting the orthogonality of the two states. Therefore it must be the case that

$$\hat{s}_- \cdot \hat{s}_+ = -1 \quad (2.5.61)$$

The Stokes coordinates for any two orthogonal polarization states are on opposite sides of the Poincaré sphere: $\hat{s}_- = -\hat{s}_+$. Specifically, a chord that connects any two orthogonal states crosses through the origin of the sphere. The polarimetric parameters χ and ϕ are related through the Stokes vectors as

$$\begin{pmatrix} \cos 2\chi_- \\ \sin 2\chi_- \cos \phi_- \\ \sin 2\chi_- \sin \phi_- \end{pmatrix} = - \begin{pmatrix} \cos 2\chi_+ \\ \sin 2\chi_+ \cos \phi_+ \\ \sin 2\chi_+ \sin \phi_+ \end{pmatrix} \quad (2.5.62)$$

A sufficient requirement to satisfy (2.5.62) is

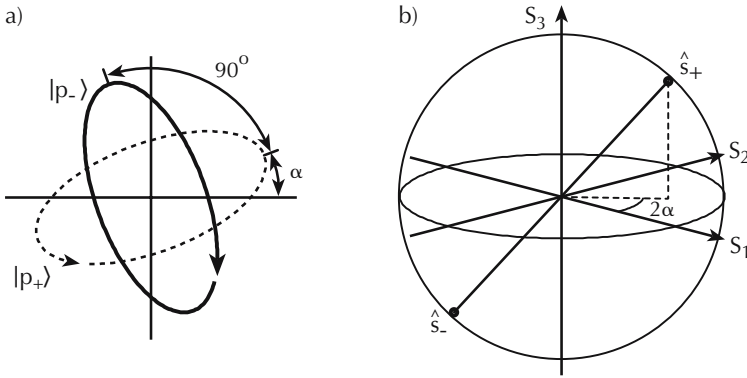


Fig. 2.3. Orthogonal polarization states in Jones and Stokes space. a) The handedness of the polarization ellipse is reversed and the major axis is rotated by $\pi/2$. b) Points on opposite sides of the Poincaré sphere are orthogonal.

$$2\chi_- = 2\chi_+ + \pi \quad (2.5.63a)$$

$$\phi_- = \phi_+ \quad (2.5.63b)$$

Equations (2.5.61) and (2.5.63) show that orthogonal polarization states have opposite handedness and perpendicular orientations of the respective elliptical major axes. The Jones and Stokes representation of orthogonal polarization pairs is illustrated in Fig. 2.3.

2.5.7 Non-Orthogonal Polarization States

The inner product magnitude between two polarization states may be calculated either in Jones or Stokes space. Consider two Jones vectors $|p\rangle$ and $|q\rangle$ that are not normalized, and recall that $\text{Tr}(|p\rangle\langle q|) = \langle p|q\rangle$. In a manner similar to (2.5.29), the inner product between the two Jones vectors is written

$$|p\rangle\langle p| = \frac{1}{2} (I + \vec{p} \cdot \vec{\sigma}) \langle p|p\rangle \quad (2.5.64)$$

Multiplication on the right by $|q\rangle$ and on the left by $\langle q|$, and some rearrangement, makes

$$\frac{|\langle p|q\rangle|^2}{\langle p|p\rangle\langle q|q\rangle} = \frac{1}{2} (1 + \vec{p} \cdot \vec{q}) \quad (2.5.65)$$

When $|p\rangle$ and $|q\rangle$ are normalized, the identity reduces to

$$|\langle p|q\rangle|^2 = \frac{1}{2} (1 + \hat{p} \cdot \hat{q}) \quad (2.5.66)$$

The magnitude of the inner product in Jones space is derived directly from the Stokes vectors using this equation. What cannot be discerned, however, is

the phase of the inner product. To recover the phase, $\langle p | q \rangle$ must be calculated explicitly in Jones space. To construct the Jones vectors, one must either solve the eigenvalue equation (2.5.30) or make the Jones vector (1.4.19) on page 18 for both $|p\rangle$ and $|q\rangle$

2.5.8 Pauli Spin Operators

A general operator can be constructed from the identity matrix and spin-vector by the form

$$\begin{aligned} A &= a_0 I + \vec{a} \cdot \vec{\sigma} \\ &= a_0 I + a_1 \sigma_1 + a_2 \sigma_2 + a_3 \sigma_3 \\ &= \begin{pmatrix} a_0 + a_1 & a_2 - ja_3 \\ a_2 + ja_3 & a_0 - a_1 \end{pmatrix} \end{aligned} \quad (2.5.67)$$

where all a_k are complex numbers. This matrix has the eight requisite independent variables necessary for a general Jones matrix. The entries in A are isolated by the trace:

$$a_0 = \frac{1}{2} \text{Tr}(A) \quad \text{and} \quad \vec{a} = \frac{1}{2} \text{Tr}(A \vec{\sigma}) \quad (2.5.68)$$

A Hermitian operator is a special case of A :

$$H = a_0 I + \vec{a} \cdot \vec{\sigma}, \quad a_k \text{ real} \quad (2.5.69)$$

The determinant is $\det(H) = a_0^2 - (a_1^2 + a_2^2 + a_3^2)$. Moreover, when the trace of H is zero, the Hermitian matrix equals a spin-vector form

$$H_{\text{Tr}=0} = \vec{a} \cdot \vec{\sigma} \quad (2.5.70)$$

Throughout much of this text, Hermitian operators with zero trace, and operators that preserve that trace, are associated with the spin-vector form.

The general operator A can be decomposed into Hermitian and skew-Hermitian matrices

$$A = H_r + jH_i \quad (2.5.71)$$

where the operator $K = (jH_i)$ is skew-Hermitian: $K^\dagger = -K$. The eigenvalues of skew-Hermitian matrices are purely imaginary. The matrices H_r and H_i contain the real and imaginary parts of A , respectively. The decomposition is taken further by separating the finite-trace component from the traceless components. Writing the complex number a_0 as $a_0 = a'_0 + ja''_0$ and identifying each traceless Hermitian matrix with a spin-vector form, one has

$$A = \vec{a}_r \cdot \vec{\sigma} + j\vec{a}_i \cdot \vec{\sigma} + (a'_0 + ja''_0) I \quad (2.5.72)$$

This is the most general spin-vector form of an arbitrary operator A . In practice, the decomposition matrices are derived from A as follows:

$$H_r = \frac{1}{2} (A + A^\dagger), \quad \text{and} \quad H_i = -\frac{j}{2} (A - A^\dagger) \quad (2.5.73)$$

The matrices H_r and H_i are made traceless by calculating $a_0 = \frac{1}{2} \text{Tr}(H)$ for the real and imaginary components. The real-valued Stokes vectors $\vec{a}_r$ and σ_i can be read from the matrices $\vec{a}_{r,i} \cdot \vec{\sigma} = H_{r,i} - a_{r,i,0} I$.

The Pauli spin operator $\mathcal{S}$ produces the most compact way to describe operators and concatenations of operators. The spin operator is a matrix exponential form of A and can describe Hermitian, unitary, and general operators. The matrix exponential is written

$$\mathcal{S} = \exp(A/2) \quad (2.5.74)$$

The exponential is evaluated using its Taylor expansion. For instance,

$$\begin{aligned} \exp(M) &= I + M + \frac{1}{2!} M^2 + \frac{1}{3!} M^3 + \frac{1}{4!} M^4 + \dots \\ &= \left(I + \frac{1}{2!} M^2 + \frac{1}{4!} M^4 + \dots \right) + \left(M + \frac{1}{3!} M^3 + \frac{1}{5!} M^5 + \dots \right) \end{aligned} \quad (2.5.75)$$

For a Hermitian matrix, $M = \alpha_o I + (\vec{\alpha} \cdot \vec{\sigma})$ where α_k are real. The n^{th} -order spin-vector identity (2.5.50) gives the necessary reductions for $(\vec{\alpha} \cdot \vec{\sigma})^n$, which gives

$$\exp(\vec{\alpha} \cdot \vec{\sigma}/2) = I \cosh(\alpha/2) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha/2) \quad (2.5.76)$$

where $\vec{\alpha} = \alpha \hat{\alpha}$. The Pauli spin operator for a Hermitian operator is thus

$$\begin{aligned} \mathcal{S}_H &= \exp(\alpha_0/2) \exp(\vec{\alpha} \cdot \vec{\sigma}/2) \\ &= e^{\alpha_0/2} \left[I \cosh(\alpha/2) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha/2) \right] \end{aligned} \quad (2.5.77)$$

One can interpret this operator as that for a partial polarizer: the common loss is $\alpha_0/2$ (which is a negative quantity for loss), the maximum and minimum differential losses are $1 \pm \tanh \alpha/2$, and the Stokes direction of partial polarization is $\hat{\alpha}$.

The Pauli spin operator for a unitary matrix is constructed by recalling the connection between Hermitian and unitary operators (2.4.18) on page 49. For coefficients β_k real, the unitary form of M is $M = -j(\beta_o I + (\vec{\beta} \cdot \vec{\sigma}))$. The equivalent to (2.5.76) is

$$\exp(-j \vec{\beta} \cdot \vec{\sigma}/2) = I \cos(\beta/2) - j(\hat{\beta} \cdot \vec{\sigma}) \sin(\alpha/2) \quad (2.5.78)$$

where $\vec{\beta} = \beta \hat{\beta}$. The Pauli spin operator for a unitary operator is thus

$$\begin{aligned} \mathcal{S}_T &= \exp(-j \beta_0/2) \exp\left(-j \vec{\beta} \cdot \vec{\sigma}/2\right) \\ &= e^{-j \beta_0/2} \left[I \cos(\beta/2) - j(\hat{\beta} \cdot \vec{\sigma}) \sin(\beta/2) \right] \end{aligned} \quad (2.5.79)$$

One interprets this operator as a retardation plate: $-\beta_0/2$ is the common phase, the full retardance is β , and the axis of retardation is $\hat{\beta}$. When the common phase is removed the unitary matrix U is recovered.

In the particular case where the axes of polarization and retardance coincide, the compound effect is expressed as the spin vector $(-j\vec{\beta} + \vec{\alpha}) \cdot \vec{\sigma}$. Following the form of (2.5.76) the Pauli spin operator expands to matrix form as

$$\begin{aligned} &\exp((-j \beta_0 + \alpha_0)/2) \exp\left((-j\vec{\beta} + \vec{\alpha}) \cdot \vec{\sigma}/2\right) = \\ &e^{(-j\beta_0 + \alpha_0)/2} \left[I \cosh((-j\beta + \alpha)/2) + (\hat{r} \cdot \vec{\sigma}) \sinh((-j\beta + \alpha)/2) \right] \end{aligned} \quad (2.5.80)$$

where $\hat{r}$ is the axis of polarization and retardance.

Two relevant properties of matrix exponentials are

$$\frac{\partial}{\partial t} e^{At} = A e^{At} = e^{At} A \quad (2.5.81)$$

$$\frac{\partial}{\partial t} e^{At} e^{Bt} e^{Ct} = A e^{At} e^{Bt} e^{Ct} + e^{At} B e^{Bt} e^{Ct} + e^{At} e^{Bt} e^{Ct} C \quad (2.5.82)$$

while two non-intuitive results are

$$\frac{\partial}{\partial A} e^{At} \neq t e^{At} \quad (2.5.83)$$

$$e^{At} e^{Bt} \neq e^{(A+B)t} \quad (2.5.84)$$

unless, for the second equation, A and B commute.

When one writes a compound operation such as

$$H_1 U H_2 = \exp((\vec{\alpha}_1 \cdot \vec{\sigma})/2) \exp(-j(\vec{\beta} \cdot \vec{\sigma})/2) \exp((\vec{\alpha}_2 \cdot \vec{\sigma})/2) \quad (2.5.85)$$

one is only figuratively expressing the series of operations on a polarization state. The evaluation of $H_1 U H_2$ requires substitution of the matrix form for each operator.

2.6 Equivalent Unitary Transformations

The preceding sections have established the bilateral connection between Jones and Stokes vectors, and have constructed Hermitian, unitary, and general Jones operators that act on Jones vectors. The connection between a Hermitian matrix and an equivalent Stokes matrix is made with the Mueller

matrix, (1.4.22) on page 18 and figuratively (2.1.3) on page 38. Mueller matrices operate on 4×1 Stokes vectors to create new, transformed 4×1 Stokes vectors.

The connection between a unitary matrix and an equivalent Stokes matrix is also made with the Mueller matrix, but as indicated by (2.1.1) on page 38, only the lower right 3×3 sub-matrix $\mathbf{R}$ is relevant. Sub-matrix $\mathbf{R}$ maps spherical coordinates (S_1, S_2, S_3) into new spherical coordinates (S'_1, S'_2, S'_3) without change of length. Therefore one expects the existence of a rotation operator R corresponding to matrix $\mathbf{R}$ that performs rotations on the Poincaré sphere. The operator R does indeed exist and its derivation and properties are so central to the description of retardance that this entire section is devoted to its understanding.

Operators U and R are equivalent representation of the same transformation cast in two different vector spaces. The operators are called isomorphic because they have similar, but not equal, effects. The isomorphism is two-to-one since, as well be seen, there are two Jones operations that have the same effect as every one Stokes operation.

Consider equivalent vectors $|s\rangle$ and $\hat{s}$ at the input of a system and equivalent vectors $|t\rangle$ and $\hat{t}$ at the output. In Jones space a unitary transformation T , corresponding to the underlying Maxwell's equations in anisotropic media, links the input and output. In Stokes space the rotation operator R links the input and output. The parallel transformations are

$$|t\rangle = T|s\rangle \xleftrightarrow{\text{dual}} \hat{t} = R\hat{s} \quad (2.6.1)$$

Expansion of the Stokes vectors on the right side of (2.6.1) into their corresponding inner products gives the relation between R and T

$$\langle t|\vec{\sigma}|t\rangle = R\langle s|\vec{\sigma}|s\rangle \quad (2.6.2)$$

Replacing $|t\rangle$ with $T|s\rangle$ and applying identity (2.5.53) gives

$$\langle s|T^\dagger\vec{\sigma}T|s\rangle = \langle s|R\vec{\sigma}|s\rangle \quad (2.6.3)$$

Since (2.6.3) holds for any $|s\rangle$, the embedded operators must be equal. Therefore

$$R\vec{\sigma} = U^\dagger\vec{\sigma}U \quad (2.6.4)$$

where the common phase of T commutes with $\vec{\sigma}$ and is eliminated. Equation (2.6.4) has an unusual form; the interpretation is

$$\left(\begin{array}{c} \\ 3 \times 3 \\ \end{array} \right) \left(\begin{array}{c} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{array} \right) = \left(\begin{array}{c} U^\dagger\sigma_1U \\ U^\dagger\sigma_2U \\ U^\dagger\sigma_3U \end{array} \right) \quad (2.6.5)$$

That R is unitary is derived by multiplying (2.6.4) by its adjoint:

$$\vec{\sigma}^\dagger R^\dagger R \vec{\sigma} = U^\dagger \vec{\sigma}^\dagger U U^\dagger \vec{\sigma} U \quad (2.6.6)$$

The product $\vec{\sigma}^\dagger \vec{\sigma} = 3I$, so only if $R^\dagger R = I$ does (2.6.6) hold. Therefore,

$$R^\dagger R = R R^\dagger = I \quad (2.6.7)$$

Since R is unitary it embodies a pure rotation; there is no scaling or translation; $\hat{t}$ is related to $\hat{s}$ by a rotation in Stokes space. The group properties of R and their correspondence to the group properties of U are listed in the following section.

There are two ways to derive an explicit expression for R , either through the matrix form of U , generating a matrix form of R ; or through the Pauli spin operator form of U , generating a vector form of R . The vector form is more “lightweight” and powerful than the matrix form because successive operations in Stokes space are evaluated purely in vector form without an *a priori* choice of orthonormal basis. The vector form is also a template with which to construct any matrix R without matrix multiplication.

2.6.1 Group Properties of SU(2) and O(3)

The group defined by multiplication operators of R is called O(3), where O stands for orthogonal and (3) for the three rotational dimensions of R . The group properties for multiplication mirror those for SU(2), cf. §2.4.6:

Identity	$UI = U$	$RI = R$
Closure	$U_1 U_2 = U_3$	$R_1 R_2 = R_3$
Inverse	$U^{-1} U = I$	$R^{-1} R = I$
Associativity	$(U_1 U_2) U_3 = U_1 (U_2 U_3) \quad (R_1 R_2) R_3 = R_1 (R_2 R_3)$	

where in all cases $U_{1,2,3} \in \text{SU}(2)$ and $R_{1,2,3} \in \text{O}(3)$.

To confirm the entries in the above table, note that multiplication of successive U and R operators form a one-to-one correspondence. For example,

$$|t\rangle = T_2 T_1 |s\rangle \xleftarrow{\text{dual}} \hat{t} = R_2 R_1 \hat{s} \quad (2.6.8)$$

Operator equality is generated through the expression

$$R_2 R_1 \vec{\sigma} = U_1^\dagger U_2^\dagger \vec{\sigma} U_2 U_1 \quad (2.6.9)$$

Making the correspondences $R' = R_2 R_1$ and $U' = U_2 U_1$, then (2.6.9) is equivalent to (2.6.4). Given that $U_2 U_1 \in \text{SU}(2)$ and $U' \longleftrightarrow R'$, one infers closure on O(3).

There are twice as many entries in the group SU(2) as in O(3). For every U_o , the corresponding R_o is $R_o \vec{\sigma} = U_o^\dagger \vec{\sigma} U_o$. For every $-U_o$ the corresponding operator R_o is the same: $R_o \vec{\sigma} = (-U_o)^\dagger \vec{\sigma} (-U_o)$. The isomorphism between SU(2) and O(3) is two-to-one.

2.6.2 Matrix Entries of R in a Fixed Coordinate System

One way to generate R explicitly is through the matrix form of U . Earlier, the generation of matrix entries for the Mueller matrix from entries in the Jones matrix was given without derivation, (1.4.22) on page 18. That equation created a 4×4 matrix; only for a unitary matrix is a distinct 3×3 sub-matrix formed. The matrix entries for $\mathbf{M}$ are found as follows.

Start with the two relations $\hat{t}_j = \langle t | \sigma_j | t \rangle$ and $\hat{t}_j = \text{Tr}(M \sigma_j)$, equations (2.5.24) on page 55 and (2.5.68) on page 61, respectively. The index j is for $j = 0, 1, 2, 3$. Moreover, assume the system $|t\rangle = J|s\rangle$. Identification with $\langle t | t \rangle = \text{Tr}(|t\rangle\langle t|)$ gives

$$\begin{aligned} t_j &= \langle t | \sigma_j | t \rangle \\ &= \text{Tr}(|t\rangle\langle t| \sigma_j) \\ &= \text{Tr}(J|s\rangle\langle s| J^\dagger \sigma_j) \end{aligned} \quad (2.6.10)$$

Next, the outer product $|s\rangle\langle s|$ is replaced with something close to the spin-vector form (2.5.29) on page 56. Since the vector is not necessarily normalized, the expression $|s\rangle\langle s| = \frac{1}{2}\langle s | s \rangle (I + \hat{s} \cdot \vec{\sigma})$ will be used. Thus,

$$\begin{aligned} \text{Tr}(J|s\rangle\langle s| J^\dagger \sigma_j) &= \frac{1}{2} \langle s | s \rangle \text{Tr}(J(I + \hat{s} \cdot \vec{\sigma}) J^\dagger \sigma_j) \\ &= \frac{1}{2} \langle s | s \rangle \text{Tr}(J(\hat{s} \cdot \vec{\sigma}) J^\dagger \sigma_j) \\ &= \frac{1}{2} \sum_{k=0}^3 \text{Tr}(J \sigma_k J^\dagger \sigma_j) s_k \end{aligned} \quad (2.6.11)$$

where $\vec{\sigma}$ has been loosely indexed here to include σ_0 , $s_k = \langle s | s \rangle \hat{s}_k$, and the common phase in T commutes with $(\hat{s} \cdot \vec{\sigma})$ and is eliminated. The last equation has the matrix multiplication form

$$t_j = \sum_{k=0}^3 M_{j+1, k+1} s_k \quad (2.6.12)$$

Identification of the matrix entries gives the final expression

$$M_{j+1, k+1} = \frac{1}{2} \text{Tr}(J \sigma_k J^\dagger \sigma_j) \quad (2.6.13)$$

The specialized case of $J = U$ is of immediate interest. Substitution of U for J in (2.6.13) shows that for $k = 0, j \neq 0$, and for $j = 0, k \neq 0$, the matrix entries are identically zero. This is because $\text{Tr}(\sigma_j) = 0$. Other than the $M_{1,1}$ matrix entry, which is unity, only the sub-matrix R survives the trace. Explicitly, the matrix entries for R given a matrix form of U are

$$R_{j,k} = \frac{1}{2} \text{Tr}(U \sigma_k U^\dagger \sigma_j) \quad (2.6.14)$$

Table 2.1. Elementary Rotations in Jones and Stokes Space

$\kappa = 0$	$U = \begin{pmatrix} e^{j\alpha} & \\ & e^{-j\alpha} \end{pmatrix}$	$R_1 = \begin{pmatrix} 1 & & \\ & \cos 2\alpha & \sin 2\alpha \\ & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}$
$\alpha = 0$ $\beta = \pi/2$	$U = \begin{pmatrix} \cos \kappa & -j \sin \kappa \\ -j \sin \kappa & \cos \kappa \end{pmatrix}$	$R_2 = \begin{pmatrix} \cos 2\kappa & \sin 2\kappa & \\ & 1 & \\ -\sin 2\kappa & \cos 2\kappa & \end{pmatrix}$
$\alpha = \beta = 0$	$U = \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix}$	$R_3 = \begin{pmatrix} \cos 2\kappa & -\sin 2\kappa & \\ \sin 2\kappa & \cos 2\kappa & \\ & & 1 \end{pmatrix}$

Using the general form of U given in (2.4.26) on page 51, R is resolved as

$$R = \begin{pmatrix} \cos 2\kappa & -\cos(\alpha - \beta) \sin 2\kappa & -\sin(\alpha - \beta) \sin 2\kappa \\ \cos(\alpha + \beta) \sin 2\kappa & \cos 2\alpha \cos^2 \kappa - \cos 2\beta \sin^2 \kappa & \sin 2\alpha \cos^2 \kappa + \sin 2\beta \sin^2 \kappa \\ -\sin(\alpha + \beta) \sin 2\kappa & -\sin 2\alpha \cos^2 \kappa + \sin 2\beta \sin^2 \kappa & \cos 2\alpha \cos^2 \kappa + \cos 2\beta \sin^2 \kappa \end{pmatrix} \quad (2.6.15)$$

Calculation of the determinant gives

$$\det(R) = 1 \quad (2.6.16)$$

which verifies that R is invertible, unitary, and contains no reflections.

Table 2.1 gives the three elementary rotations about the Stokes axes (S_1, S_2, S_3) and their original unitary form. Notice that all angles which appear in the unitary matrices are doubled in the corresponding Stokes matrices. Stokes angles are twice that of physical, or “laboratory” angles. This is also why the Jones to Stokes isomorphism is two-to-one: rotation of π in Jones space is invariant, and so is rotation by 2π in Stokes space.

2.6.3 Vector Expression of R in a Local Coordinate System

The vector form of R abstracts away any notion of an underlying, fixed coordinate system. Rather, each operation R has its own local coordinate system based on the eigenvectors and spin direction of R . The vectorial form of R gives the highest level of geometric interpretation to transformation mechanics in Stokes space.

The vector expression for R is derived from the vector form of U . The operator U is resolved into its eigenvector-based projectors using (2.4.15) on page 48; the resolution for a two-dimensional system gives

$$U = e^{-j\varphi/2} |r_+\rangle\langle r_+| + e^{j\varphi/2} |r_-\rangle\langle r_-| \quad (2.6.17)$$

where the projectors are equated to the spin-vector form via

$$|r_\pm\rangle\langle r_\pm| = \frac{1}{2} (I \pm \hat{r} \cdot \vec{\sigma}) \quad (2.6.18)$$

Note that $\hat{r}_+ = -\hat{r}_-$ are orthogonal Stokes coordinates. In the following, $\hat{r}$ will denote the eigenvector with the positive subscript. Substitution of the spin-vector form of the projector into U gives

$$U = I \cos(\varphi/2) - j(\hat{r} \cdot \vec{\sigma}) \sin(\varphi/2) \quad (2.6.19)$$

Equation (2.6.19) is now in familiar form and can be mapped to the exponential equivalent as

$$U = e^{-j(\varphi/2)(\hat{r} \cdot \vec{\sigma})} \quad (2.6.20)$$

where $(\varphi/2)(\hat{r} \cdot \vec{\sigma})$ is the Hermitian operator associated with U .

Substitution of (2.6.19) into the equivalence relation (2.6.4), and applying the spin-vector identities (2.5.35), (2.5.36), and (2.5.49) produces

$$R \vec{\sigma} = \cos \varphi \vec{\sigma} + (1 - \cos \varphi) \hat{r}(\hat{r} \cdot \vec{\sigma}) + \sin \varphi (\hat{r} \times \vec{\sigma}) \quad (2.6.21)$$

Since each term on the left- and right-hand sides of (2.6.21) operates on $\vec{\sigma}$, one can extract the embedded relation for R

$$R = \cos \varphi I + (1 - \cos \varphi)(\hat{r}\hat{r}\cdot) + \sin \varphi(\hat{r}\times) \quad (2.6.22)$$

Grouping like terms gives

$$R = (\hat{r}\hat{r}\cdot) + \sin \varphi(\hat{r}\times) + \cos \varphi(I - (\hat{r}\hat{r}\cdot)) \quad (2.6.23)$$

Recalling the vector identity $\mathbf{a} \times \mathbf{a} \times \mathbf{c} = \mathbf{b}(\mathbf{a} \cdot \mathbf{a}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$, the last term on the right-hand side is identified as

$$(\hat{r}\times)(\hat{r}\times) = (\hat{r}\hat{r}\cdot) - I \quad (2.6.24)$$

The final vectorial form of R is then

$$R = (\hat{r}\hat{r}\cdot) + \sin \varphi(\hat{r}\times) - \cos \varphi(\hat{r}\times)(\hat{r}\times) \quad (2.6.25)$$

Equation (2.6.25) is a beautifully compact expression for the action any unitary operator has on a polarization vector. The vector $\hat{r}$ points in the direction of the positive eigenvector of U . The vector operators $\{(\hat{r}\hat{r}\cdot), (\hat{r}\times), (\hat{r}\times)(\hat{r}\times)\}$ form a local orthonormal basis. The local basis requires a vector about which

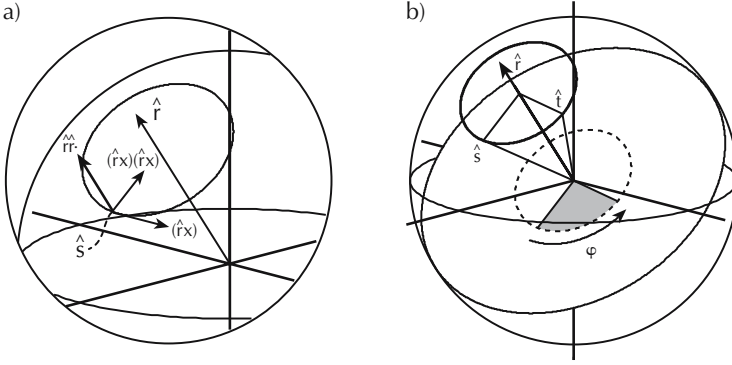


Fig. 2.4. Vector components of rotation operator R . a) The local orthonormal basis $\{(\hat{r}\hat{r}\cdot), (\hat{r}\times), (\hat{r}\times)(\hat{r}\times)\}$ as resolved on $\hat{s}$. b) Transformation to $\hat{t}$ from $\hat{s}$ via precession about $\hat{r}$, travelling through precession angle φ .

the basis can be fully resolved; for instance, operation on state $\hat{s}$ generates the basis $(\hat{r}, \hat{r} \times \hat{s}, \hat{r} \times \hat{r} \times \hat{s})$. In the absence of being fully resolved, the local basis has immutable properties that are independent of the resolving vector.

Figure 2.4(a) illustrates the local basis resolved by $\hat{s}$. Vector $\hat{s}$ in relation to $\hat{r}$ defines a precession circle, the circle about which $\hat{s}$ travels. Local axis $(\hat{r}\hat{r}\cdot)$ always points parallel to $\hat{r}$. The local axes $(\hat{r}\times), (\hat{r}\times)(\hat{r}\times)$ define the plane of the precession circle and are perpendicular to $\hat{r}$. The local axis $(\hat{r}\times)$ is tangent to the precession circle and $(\hat{r}\times)(\hat{r}\times)$ points to the origin of the precession circle. The particular pointing directions of $(\hat{r}\times)$ and $(\hat{r}\times)(\hat{r}\times)$, while always in the precession plane, are determined only after determination of $\hat{s}$. Figure 2.4(b) illustrates transformation to state $\hat{t}$ from $\hat{s}$ about $\hat{r}$. The precession angle is φ and the precession direction follows the right-hand rule.

Since the motion of precession is so central in the description of polarization transformation mechanics, Fig. 2.5 is included to describe precession in a local coordinate system. Consider the input state $\hat{s}$ and the precession axis $\hat{r}$. The precession axis can be the birefringent axis of a dielectric medium or the principal-state-of-polarization axis used to describe polarization-mode dispersion. In any case, the angle γ separates the two vectors. The motion of precession is to turn $\hat{s}$ about $\hat{r}$ in a circle while keeping the angle γ fixed. This is the same motion a gyroscope exhibits under gravitational influence. The angle subtended by projections of states $\hat{s}$ and $\hat{t}$ onto the base circle is the precession angle. The differential equation of motion can be deduced from R in local-coordinate form. Consider state $\hat{s}$ that undergoes a small change in angle $\delta\varphi$. The motion is

$$\hat{s} + \delta\hat{s} = R_{\delta\varphi}\hat{s} \quad (2.6.26)$$

Taking R in the form (2.6.22) and simplifying for small angles,

$$\hat{s} + \delta\hat{s} = I\hat{s} + \delta\varphi \hat{r} \times \hat{s} \quad (2.6.27)$$

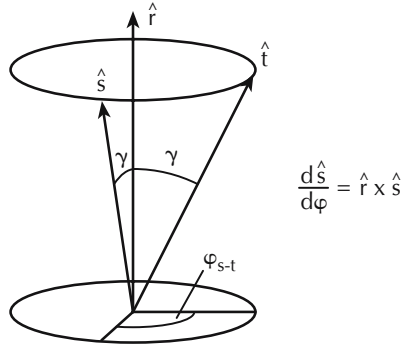


Fig. 2.5. Precessional motion of $\hat{s}$ about $\hat{r}$, passing through state $\hat{t}$. Angle γ remains fixed, while angle φ , as projected onto the base, is the degree of precession.

which is rewritten in differential form as

$$\frac{d\hat{s}}{d\varphi} = \hat{r} \times \hat{s} \quad (2.6.28)$$

The $\hat{r} \times \hat{s}$ term dictates that $\hat{s}$ moves perpendicular to $\hat{r}$.

Equation 2.6.22 can be used as a template to construct a matrix representation for R given any $\hat{r}$ and φ . The matrix representations for $\hat{r}\hat{r}\cdot$ and $\hat{r}\times$ are

$$(\hat{r}\hat{r}\cdot) = \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}, \quad (\hat{r}\times) = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} \quad (2.6.29)$$

The first matrix is a projector, as verified by $\det(\hat{r}\hat{r}\cdot) = 0$. The second matrix is derived from $\hat{r}\times = \hat{r} \times S_1 + \hat{r} \times S_2 + \hat{r} \times S_3$, where $S_{1,2,3}$ are the three Stokes axes.

2.6.4 Select Vector Identities

Figure 2.6 illustrates two identities that have a simple geometric interpretation. The identity

$$\hat{r} \cdot (R \vec{a}) = \hat{r} \cdot \vec{a} \quad (2.6.30)$$

is illustrated by Fig. 2.6(a). The rotation of $\vec{a}$ generated by R about $\hat{r}$ through angle φ does not change the angle between $\hat{r}$ and $\vec{a}$. Thus the dot product may be taken with or without the intermediate rotation. The identity

$$(R \vec{a}) \cdot (R \vec{b}) = \vec{a} \cdot \vec{b} \quad (2.6.31)$$

is illustrated by Fig. 2.6(b). The rotation due to R does not change the angle between vectors $\vec{a}$ and $\vec{b}$, so again the dot product may be taken with or without the intermediate rotation.

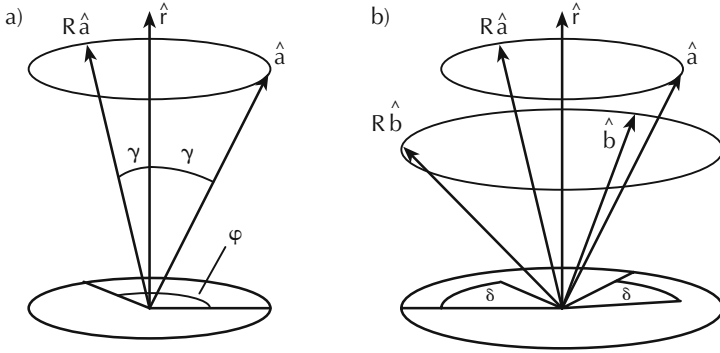


Fig. 2.6. Geometric representation of two rotational identities. a) $\hat{r} \cdot (R \vec{a}) = \hat{r} \cdot \vec{a}$. b) $(R \vec{a}) \cdot (R \vec{b}) = \vec{a} \cdot \vec{b}$.

2.6.5 Euler Rotations

The Euler rotations are an alternative method to construct the operator R in matrix form (2.6.15). While there are several ways to establish the connection, the fact that multiplication in Stokes space corresponds to multiplication in Jones space is simple enough to construct the operator U in the form (2.4.26), and to map the terms from Jones back to Stokes space. One can verify that

$$U = \begin{pmatrix} e^{ju} & \\ & e^{-ju} \end{pmatrix} \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix} \begin{pmatrix} e^{jv} & \\ & e^{-jv} \end{pmatrix} \quad (2.6.32)$$

where $u = (\alpha + \beta)/2$ and $v = (\alpha - \beta)/2$. Identification of the Jones operators in (2.6.32) with the Stokes operators in Table 2.1 gives the equivalent Stokes transformations

$$R = R_1(\alpha + \beta) R_3(2\kappa) R_1(\alpha - \beta) \quad (2.6.33)$$

These rotations generate the general rotation matrix in (2.6.15).

Another way to view the unitary operator is to diagonalize the matrix. The eigenvalues of U are complex exponentials that have unity magnitude and they are conjugates of one another: $U|r_{\pm}\rangle = \exp(\mp j\varphi/2)|r_{\pm}\rangle$. The operator U can be separated as

$$U = V\Lambda V^{\dagger} \quad (2.6.34)$$

where the matrix V is a 2×2 unitary matrix with the two eigenvectors $|r_{\pm}\rangle$ of U entered as columns of V , and where the matrix Λ is a diagonal 2×2 matrix with the eigenvalues of U on the diagonal. The corresponding Stokes operation is

$$R = R_E R_1(\varphi) R_E^{\dagger} \quad (2.6.35)$$

where R_E is the Euler rotation associated with V . Now, several observations can be made. If the state at the input is $|r_{\pm}\rangle$, then the action of U does not

change the state, only a phase is contributed. The state is invariant under U . For every $|r_{\pm}\rangle$ there are corresponding $\hat{r}_{\pm}$ vectors in Stokes space. The behavior of $R_{\text{E}}^{\dagger}$ is to rotate $\hat{r}_{\pm}$ to $\pm s_1$; $R_1(\varphi)$ then pirouettes the state about the s_1 axis, and R_{E} returns the state back to $\hat{r}_{\pm}$.

The decomposition of U as in (2.6.34) has much significance in relation to propagation through birefringent media. For example, the propagation constants for ordinary and extraordinary waves in a birefringent medium are $\beta_o = \omega n_o/c$ and $\beta_e = \omega n_e/c$. The eigenvalues of the propagation matrix are $\exp(\mp j(\beta_e - \beta_o)z/2)$. The polarization transformation in Stokes space is accordingly

$$R_{\Delta\beta} = R_{\text{E}}R_x(\Delta\beta z)R_{\text{E}}^{\dagger} \quad (2.6.36)$$

The inner matrix R_x creates precession about the s_1 axis in Stokes space while the Euler rotation and its adjoint transforms the eigenstates of the system onto the s_1 axis and then restores the pointing direction of the eigenstate. The precession about s_1 is transformed to precession about $\hat{r}$.

2.6.6 Some Relevant Transformation Applications

Four examples are presented here to give some illustrative detail on how to use the Stokes transformation operator R cast in local-coordinate form. First, differential precession rules for a single homogeneous birefringent section are written. Then, the polarization evolution through a concatenation of two misaligned birefringent sections, as a function of length, is illustrated. Third, the shortest distance between two polarization states is found. Finally, uniform and biased polarization scattering examples are given.

For polarization studies, the axis $\hat{r}$ is the extraordinary axis of a birefringent medium. For a birefringent crystal the birefringent axis lies in the equatorial plane in Stokes space. An input state $\hat{s}$ precesses about $\hat{r}$ as the state propagates through the medium. The retardation of a birefringent plate is just the angle φ through which the state processes: $\varphi = \omega\Delta nL/c$, albeit any two rotations of φ that are the same modulo 2π yield identical Stokes transformations. The angle φ is called the birefringent phase.

The differential equation of motion for a state of polarization as it evolves through a homogeneous birefringent material is the same for differentials in either position or frequency. The retardation as a function of length z and radial frequency ω is

$$\varphi = \frac{\omega\Delta n z}{c} \quad (2.6.37)$$

Moreover, the birefringent axis Ω is parallel to $\hat{r}$. Differentiation of (2.6.37) with respect to z and substitution into (2.6.28) gives

$$\frac{d\hat{s}}{dz} = \Omega \times \hat{s} \quad (2.6.38)$$

where $\Omega = (\omega\Delta n/c)\hat{r}$. Equally possible is differentiation with respect to ω , which gives

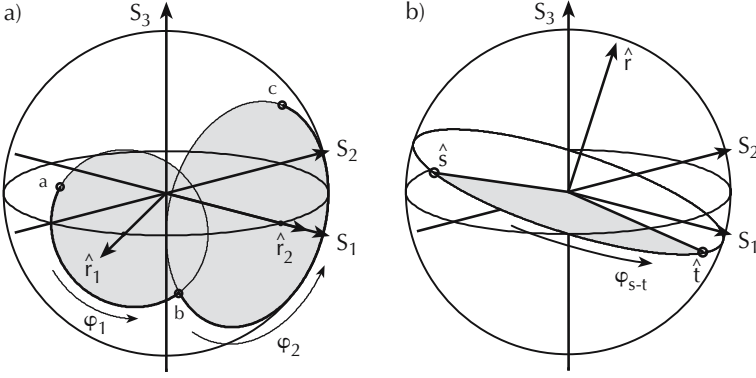


Fig. 2.7. a) Polarization evolution through two misaligned birefringent sections. Input state (a) precesses about $\hat{r}_1$ to state (b). That state enters the second stage, precesses about $\hat{r}_2$, and leaves as state (c). b) Construction of great circle through states $\hat{s}$ and $\hat{t}$. The normal to the circle is $\hat{r}$.

$$\frac{d\hat{s}}{d\omega} = \Omega \times \hat{s} \quad (2.6.39)$$

The differential precession rule for a single homogeneous birefringent section is the same whether the position or frequency changes. This simplicity is quickly broken when two or more homogeneous sections are concatenated. Birefringent concatenation is in the category of polarization-mode dispersion.

The polarization state evolution through two misaligned birefringent sections as a function of length can be evaluated using (2.6.25) in the following way. Since the media are misaligned, their birefringent axes are not parallel; that is, $\hat{r}_1 \neq \hat{r}_2$. Figure 2.7(a) illustrates the polarization evolution through the sections. The input state, arbitrarily selected, is located at position (a). That state precesses about $\hat{r}_1$ through angle φ_1 , dictated by the length and birefringence of the section, as well as the input frequency. The output polarization from the first section is located at position (b). That state then enters the second section which transforms it about $\hat{r}_2$. The polarization state now traces a second circle that is different from the first. The output state is eventually located at position (c). The aggregate polarization transformation is calculated by the concatenation of R_2 and R_1 . The compounded polarization transformation is

$$R_2 R_1 = \left((\hat{r}_2 \hat{r}_2 \cdot) + \sin \varphi_2 (\hat{r}_2 \times) - \cos \varphi_2 (\hat{r}_2 \times) (\hat{r}_2 \times) \right) \quad (2.6.40)$$

$$\left((\hat{r}_1 \hat{r}_1 \cdot) + \sin \varphi_1 (\hat{r}_1 \times) - \cos \varphi_1 (\hat{r}_1 \times) (\hat{r}_1 \times) \right) \quad (2.6.41)$$

Each transformation is denoted by a unique index and the concatenation is written right-to-left as is usual for matrix multiplication. Sometimes the vector products offer simplifications that can reduce the complexity of the overall motion.

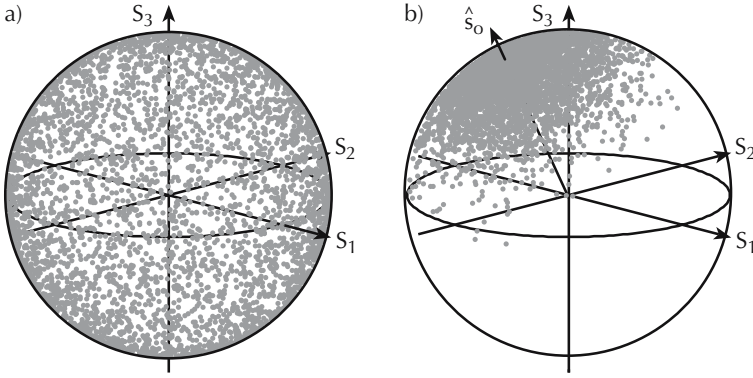


Fig. 2.8. Uniform and biased scattering through operator R . a) Uniform scattering. $\hat{r}$ points in any direction with equal likelihood. φ is uniformly distributed. b) Biased scattering, $a = 0.05$. $\hat{R}$ is constructed along s_3 and oriented toward $\hat{s}_o$.

As another application, the shortest distance between points $\hat{s}$ and $\hat{t}$ on a unit sphere is along a great circle. The axis normal to the great circle is evidently

$$\hat{r} = \frac{\hat{s} \times \hat{t}}{|\hat{s} \times \hat{t}|} \quad (2.6.42)$$

and the rotation angle between the two points is

$$\cos \varphi_{s-t} = \hat{s} \cdot \hat{t} \quad (2.6.43)$$

Figure 2.7(b) illustrates the motion. $\hat{r}$ is derived from the cross of $\hat{s}$ and $\hat{t}$. Angle φ_{s-t} rotates $\hat{s}$ through to $\hat{t}$.

Uniform and biased polarization scattering is useful in connection with polarization-mode dispersion fiber-modelling calculations. The scattering process occurs between any two adjacent birefringent sections and it intended to model the relative alignments of the respective birefringent axes. A uniform scattering process sends the polarization state at the output of one section, $\hat{s}_o$, to any point on the Poincaré sphere with equal probability. That state, $\hat{s}_i$, is then input to the next birefringent section. The biased scattering process weights the scattering along a predetermined direction, often the direction of $\hat{s}_o$. For either uniform or biased scattering an operator R needs to be constructed.

There are two variables contained in R , (2.6.25): pointing direction $\hat{r}$ and precession angle φ . Direction $\hat{r}$ itself has two independent variables, the polar angles of declination and azimuth. Combined, R has three independent variables. The random process is derived using the unit deviate $\tilde{u}$. To have $\hat{r}$ point in any direction on the unit sphere with equal likelihood, the azimuth angle ϕ and position along the s_3 axis are both uniformly distributed. Also, the precession angle is uniformly distributed to generate precessions with equal

likelihood. The random variable expressions are

$$\tilde{r}_3 = 2\tilde{u} - 1 \quad (2.6.44a)$$

$$\tilde{\phi} = (2\tilde{u} - 1)\pi \quad (2.6.44b)$$

$$\tilde{\varphi} = (2\tilde{u} - 1)\pi \quad (2.6.44c)$$

Relating $\tilde{r}_3$ to the polar angle as $\tilde{r}_3 = \cos\theta$, the remaining coordinates are

$$\tilde{r}_1 = \sin\theta \cos\tilde{\phi} \quad (2.6.45a)$$

$$\tilde{r}_2 = \sin\theta \sin\tilde{\phi} \quad (2.6.45b)$$

The random variable $\tilde{R}$ can now be constructed. Since uniform scattering is completely symmetric on the unit sphere, the pointing direction $\tilde{r}$ does not need to be oriented toward $\hat{s}_o$. Figure 2.8(a) illustrates an output state scattered on the Poincaré sphere.

Biased scattering is used to enhance the likelihood of rare events, rare events being those where multiple birefringent axes are preferentially aligned, misaligned, or some other construction. To preferentially align the birefringent sections, $\hat{r}$ should be biased to point toward $\hat{s}_o$. A simple way to generate the bias is first to bias $\tilde{r}$ towards the s_3 direction using the following formula:

$$\tilde{r}_3 = 2\tilde{u}^{1/a} - 1 \quad (2.6.46)$$

For $a \leq 1$ the bias is toward $+s_3$ and for $a \geq 1$ the bias is toward $-s_3$. The scattering operator R is now biased toward s_3 and is denoted R_3 . Before R_3 can be applied to $\hat{s}_o$ the former needs to be rotated into the latter. Following the previous example of the shortest distance between two points in Stokes space, a deterministic operator R_{3-s_o} is constructed to perform the required rotation. Operator R_{3-s_o} needs to be calculated only once. Figure 2.8(b) illustrates an output state scattered on the Poincaré sphere and biased toward $\hat{s}_o$.

Table 2.2. Jones and Stokes Equivalent Expressions

Jones expressions	Stokes expressions
$ s\rangle = \begin{pmatrix} s_x e^{j\phi_x} \\ s_y e^{j\phi_y} \end{pmatrix}$	$\hat{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}$
$ t\rangle = T s\rangle$	$\hat{t} = R \hat{s}$
$ t\rangle = T_2 T_1 s\rangle$	$\hat{t} = R_2 R_1 \hat{s}$
$TT^\dagger = T^\dagger T = I$	$RR^\dagger = R^\dagger R = I$
$ s\rangle = \hat{s} \cdot \vec{\sigma} s\rangle$	$\hat{s} = \langle s \vec{\sigma} s \rangle$
$ s\rangle \langle s = \frac{1}{2} (I + \hat{s} \cdot \vec{\sigma})$	$\hat{s} = \frac{1}{2} \text{Tr}(s\rangle \langle s \vec{\sigma})$
$U \vec{\sigma} U^\dagger$	$R \vec{\sigma}$
$U = I \cos(\varphi/2) - j(\hat{r} \cdot \vec{\sigma}) \sin(\varphi/2)$	$R = (\hat{r} \hat{r} \cdot) + \sin \varphi (\hat{r} \times) - \cos \varphi (\hat{r} \times)(\hat{r} \times)$
$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$	$R_{j,k} = \frac{1}{2} \text{Tr}(U \sigma_k U^\dagger \sigma_j)$
$U_1 = \begin{pmatrix} e^{j\alpha} & \\ & e^{-j\alpha} \end{pmatrix}$	$R_1 = \begin{pmatrix} 1 & & \\ & \cos 2\alpha & \sin 2\alpha \\ & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}$
$U_2 = \begin{pmatrix} \cos \kappa & -j \sin \kappa \\ -j \sin \kappa & \cos \kappa \end{pmatrix}$	$R_2 = \begin{pmatrix} \cos 2\kappa & \sin 2\kappa & \\ & 1 & \\ -\sin 2\kappa & \cos 2\kappa & \end{pmatrix}$
$U_3 = \begin{pmatrix} \cos \kappa & -\sin \kappa \\ \sin \kappa & \cos \kappa \end{pmatrix}$	$R_3 = \begin{pmatrix} \cos 2\kappa & -\sin 2\kappa & \\ \sin 2\kappa & \cos 2\kappa & \\ & & 1 \end{pmatrix}$
$\frac{d s\rangle}{d\varphi} = -j/2(\hat{r} \cdot \vec{\sigma}) s\rangle$	$\frac{d\hat{s}}{d\varphi} = \hat{r} \times \hat{s}$

Table 2.3. Spin-Vector Expressions

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}$$

$$\vec{\sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}$$

$$a_0 I + \vec{a} \cdot \vec{\sigma} = \begin{pmatrix} a_0 + a_1 & a_2 - j a_3 \\ a_2 + j a_3 & a_0 - a_1 \end{pmatrix}$$

$$R = (\hat{r}\hat{r}\cdot) + \sin \varphi (\hat{r}\times) + \cos \varphi (I - (\hat{r}\hat{r}\cdot))$$

$$\hat{r}\hat{r}\cdot = \begin{pmatrix} r_1 r_1 & r_1 r_2 & r_1 r_3 \\ r_2 r_1 & r_2 r_2 & r_2 r_3 \\ r_3 r_1 & r_3 r_2 & r_3 r_3 \end{pmatrix}, \quad \hat{r}\times = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix}$$

$$H = \exp(\alpha_0/2) \exp(\vec{\alpha} \cdot \vec{\sigma}/2) = e^{\alpha_0/2} [I \cosh(\alpha/2) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha/2)]$$

$$T = \exp(-j \beta_0/2) \exp(-j \vec{\beta} \cdot \vec{\sigma}/2) = e^{-j \beta_0/2} [I \cos(\beta/2) - j(\hat{\beta} \cdot \vec{\sigma}) \sin(\beta/2)]$$

References

1. O. Aso, I. Ohshima, and H. Ogoshi, “Unitary-conserving construction of the Jones matrix and its applications to polarization-mode dispersion analysis,” *Journal of the Optical Society of America A*, vol. 14, no. 8, pp. 1988–2005, Aug. 1997.
2. D. M. Brink and G. R. Satchler, *Angular Momentum*, 3rd ed. Oxford: Oxford Science Publications, 1999.
3. N. Frigo, “A generalized geometric representation of coupled mode theory,” *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 11, pp. 2131–2140, 1986.
4. N. Gisin and B. Huttner, “Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers,” *Optics Communications*, vol. 142, pp. 119–125, Oct. 1997.
5. J. P. Gordon and H. Kogelnik, “PMD fundamentals: Polarization mode dispersion in optical fibers,” *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: <http://www.pnas.org>
6. M. Rose, *Elementary Theory of Angular Momentum*. New York: Dover Publications, 1995.
7. J. J. Sakurai, *Modern Quantum Mechanics*. New York: Addison–Wesley, 1985.
8. G. Strang, *Linear Algebra and its Applications*, 3rd ed. New York: Harcourt Brace Jovanovich College Publishers, 1988.

Interaction of Light and Dielectric Media

Optical components and waveguiding fiber are designed to control the interaction between light and media. The regime of interest in this text is material transparency, to first order. Transparent glasses, crystals, and garnets are the building blocks on which passive optical components and optical fiber are made. Moreover, the interactions addressed in this text are optically linear in that the material response is assumed linear with field intensity. This is not actually the case since, for example, optical fiber has a prominent Kerr effect, but linearity will do for the studies to follow. An equally broad topic is the interaction between light and semiconductors such as diode lasers and optical detectors, but such interactions are not covered here.

The main purpose of this chapter is to introduce elementary classical descriptions for the constitutive relations of isotropic, anisotropic, gyrotropic, and optically active materials, and to detail how these constitutive relations, when included in Maxwell's equations, change the wavefront, power flow, and polarization of light. Glasses and birefringent crystals, principal examples of isotropic and anisotropic materials, are well characterized by classical wave-electron interaction. Faraday rotation induced by diamagnetic materials, a particular example of a gyrotropic material, yields to classical analysis, but a quantum-mechanical model is necessary to describe rotation in ferrimagnetic garnets. Finally, a constitutive relation of optical activity based on a classical description can be sketched to give flavor, but a detailed dipole-dipole interaction model is really necessary.

There are two levels of treatment for the interaction of light and media. First a constitutive relation must be found that dictates how the incident field effects the dipole moments (and possibly free electrons) of the material, and how these dipole moments in turn effect the field. Second, Maxwell's equations are solved based on the inclusion of the constitutive relation. The solution of Maxwell's equations, in this context, is completely classical and the rigorous treatment of the kDB system is provided.

3.1 Introduction of Media Terms into Maxwell's Equations

The modifications to Maxwell's equations to include media terms are treated first. Overall, the equations must show how light behaves within a bulk, homogeneous medium and at the interface into and out of the region. In vacuum an electromagnetic wave is described by electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$. Within a material, however, the wave is described by the electric and magnetic flux densities $\mathbf{D}$ and $\mathbf{B}$. The flux densities incorporate the original field and the media reaction. The following sections then classify possible interaction types and recast Maxwell's equations for solution within a material.

When an electromagnetic wave propagates through a material, the electric and magnetic fields couple to bound and free electrons. A material with free electrons is conductive and any currents that are generated by the field absorb energy from that field and impart loss. A material without free electrons is dielectric. The travelling field forces the bound electrons to oscillate and they, in turn, radiate. In a pure dielectric without impurities, the induced radiation field is not diffuse but instead co- and counter-propagates with the incident field, so, when combined, the total field has a shortened wavelength and slowed energy flow.

There are two interactions to account for when light propagates within a dielectric medium: the radiation field from electronic dipoles that are stimulated by the incident field; and the radiation field from magnetic dipoles, or current loops, that are likewise stimulated. The electric field can directly couple to the electric dipoles and the magnetic field can directly couple to the magnetic dipoles. There can also be cross-coupling where the electric field induces magnetic dipole radiation and the magnetic field induces electric dipole radiation, as is the case in a chiral medium.

First the electric-dipole radiation. Including paired charge density ρ_p and unpaired (or free) charge density ρ_u , Gauss' law for the electric field is

$$\nabla \cdot \varepsilon_0 \mathbf{E} = \rho_p + \rho_u \quad (3.1.1)$$

A dielectric, non-conductive medium has only paired charges since every electron is bound to an atom; the unpaired charge density is zero, $\rho_u = 0$. Without free electrons there is no current, so $\mathbf{J} = 0$ as well. For incident field energies below the ionization energy, the field stimulates the electrons to oscillate. To first order atomic nuclei do not move, so the electrons move closer and further away from the nucleus to generate a oscillating dipole. The oscillation frequency matches the frequency of the incident field.

The electric dipole moment of the oscillator is $\mathbf{p} = -e\mathbf{r}$, where $-e$ is the electron charge and $\mathbf{r}$ is the vector distance between electron and nucleus. Vector $\mathbf{r}$ points in the direction of positive charge. The polarization density vector $\mathbf{P}$ of the media is the product of the individual dipole moments and the number of dipoles N per unit volume, or

$$\mathbf{P} = -N\mathbf{e}\mathbf{r} \quad (3.1.2)$$

As the electrons oscillate, the electric charge changes position. A differential volume dV containing a charge density N has at any time a charge of $q = -e \mathbf{r} \cdot d\mathbf{a}$ exterior to its volume, where $d\mathbf{a}$ is a differential surface-area element. The total net charge Q remaining on the interior comes from integrating across the surface,

$$Q = - \oint_S (N\mathbf{e}\mathbf{r}) \cdot d\mathbf{a}$$

Rewriting Q in terms of dipole density $\mathbf{P}$, Gauss' integral law shows that

$$\oint_S \mathbf{P} \cdot d\mathbf{a} = \int_V \nabla \cdot \mathbf{P} dV$$

The net charge within volume V is then

$$Q = - \int_V \nabla \cdot \mathbf{P} dV$$

But by definition, the net charge Q within the same volume due to the paired charge density ρ_p is

$$Q = \int_V \rho_p dV$$

Comparing these two expressions for charge density, the paired charge density is related to the divergence of the polarization density by

$$\rho_p = -\nabla \cdot \mathbf{P} \quad (3.1.3)$$

Substitution back into Gauss' law gives

$$\nabla \cdot (\varepsilon_o \mathbf{E} + \mathbf{P}) = 0 \quad (3.1.4)$$

So now it is clear how to define the electric-flux density $\mathbf{D}$:

$$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P}, \quad \text{and} \quad \nabla \cdot \mathbf{D} = 0 \quad (3.1.5)$$

where $\mathbf{D}$ has units of (C/m²) and the electric-flux density has zero divergence.

Next comes magnetic-dipole radiation. Gauss' law for the magnetic field is

$$\nabla \cdot \mu_o \mathbf{H} = 0 \quad (3.1.6)$$

The divergence of the field is zero because there is no magnetic monopole. Magnetic fields and the currents that generate them must close on themselves. The elementary model for a unit magnet is a current loop having current i circulating along a perfect conductor having radius R enclosing area $\mathbf{a}$, where

the direction of $\mathbf{a}$ is normal to the surface element. The magnetic dipole moment $\mathbf{m}$ can be identified as $\mathbf{m} = i\mathbf{a}$. In analogy with the electric polarization density, the magnetization density $\mathbf{M}$ is defined as

$$\mathbf{M} = N\mathbf{m}$$

Now, in the far-field, the scalar potentials of an electric dipole and magnetic dipole have the same form, provided that the magnetic dipole is identified as $\rho_m = \mu_o m$ [8]. So, in analogy with (3.1.3), the magnetic density is defined as

$$\rho_m = -\nabla \cdot \mu_o \mathbf{M} \quad (3.1.7)$$

Introducing of ρ_m into Gauss' magnetic law gives

$$\nabla \cdot (\mu_o \mathbf{H} + \mu_o \mathbf{M}) = 0 \quad (3.1.8)$$

In parallel with the electric-flux density $\mathbf{D}$, the magnetic-flux density $\mathbf{B}$ is defined by

$$\mathbf{B} = \mu_o \mathbf{H} + \mu_o \mathbf{M}, \text{ and } \nabla \cdot \mathbf{B} = 0 \quad (3.1.9)$$

where $\mathbf{B}$ has units of (V-s/m²), or the derived unit of (Tesla).

All materials at the very least exhibit the magnetic effects of nuclear and electronic spin. In general the resultant magnetic fields close on themselves on an atomic scale and therefore average out on a macro-scale. Moreover, a magnetic moment has an angular moment and any reorientation of that moment has a delay with respect to a sinusoidal driving field. At optical frequencies the induced magnetization, e.g. of a ferrite, is essentially zero. Bianisotropic materials exhibit a helical molecular structure or crystalline structure which can combine to couple the electric and magnetic fields via dipole and current excitation. These materials cause optical activity and are of growing importance in telecommunications since liquid crystal structures have a helical component.

Having defined the electric and magnetic flux densities $\mathbf{D}$ and $\mathbf{B}$, they must be incorporated into Faraday's and Ampère's laws. Faraday's law says that the curl of $\mathbf{E}$ is generated by the time variation of a field $\mathbf{F}$:

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{F}$$

Taking the divergence of both sides yields

$$\nabla \cdot (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t} \nabla \cdot \mathbf{F} = 0$$

Since the divergence of a solenoidal field is zero, the divergence of $\mathbf{F}$ must be constant: $\nabla \cdot \mathbf{F} = c_m$. Since $\nabla \cdot \mathbf{D} = 0$, the flux density $\mathbf{D}$ is a suitable field to include in Faraday's law. The same analysis holds for Ampère's law. Therefore one can restate Maxwell's equations in terms of $\mathbf{D}$ and $\mathbf{B}$:

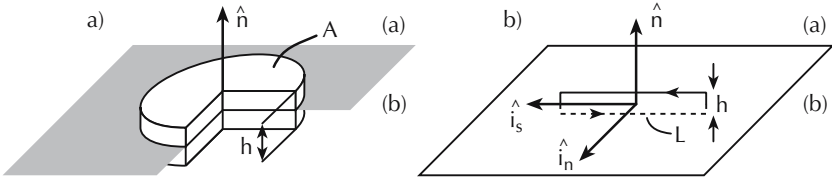


Fig. 3.1. Analysis of electric and magnetic flux density continuity across a boundary. a) Electric flux density continuity determined by a “pillbox” across the surface. b) Magnetic flux density continuity determined by a loop through the surface.

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{B} \quad (3.1.10a)$$

$$\nabla \times \mathbf{H} = \frac{\partial}{\partial t} \mathbf{D} \quad (3.1.10b)$$

The electric and magnetic flux densities are now fully incorporated into Maxwell's equations. In order to use these equations, constitutive relations that relate the polarization $\mathbf{P}$ and magnetization $\mu_o \mathbf{M}$ to the electric and magnetic fields are required. Constitutive relations determine these interactions.

Normal to an interface the electric and magnetic flux densities are continuous. Tangent to a surface the electric and magnetic fields are continuous. These continuity conditions are derived as follows.

The continuity condition for fluxes are derived from Gauss' law. Consider a small volume V in the shape of a pillbox that intersects a smooth, charge-free interface between two homogeneous regions denoted by (a) and (b) (Fig. 3.1(a)). The volume has height h and area normal to the interface A . Also, a unit vector $\hat{n}$ points from region (b) to (a). Integration of $\nabla \cdot \mathbf{D} = 0$ over the volume and taking the limit to zero height gives

$$\begin{aligned} \lim_{h \rightarrow 0} \int_V \nabla \cdot \mathbf{D} \, dV &= \lim_{h \rightarrow 0} \oint_S \mathbf{D} \cdot d\mathbf{a} \\ &= \lim_{h \rightarrow 0} \left(\hat{n} \cdot (\mathbf{D}^{(a)} - \mathbf{D}^{(b)}) A + h \int_C \mathbf{D} \cdot d\mathbf{s} \right) \\ &= 0 \end{aligned}$$

where S is the surface enclosed by volume V , C is the contour around area element A , and Stokes integral law is used to transform from volume to surface integrals. The electric flux density normal to the interface is therefore continuous:

$$\hat{n} \cdot (\mathbf{D}^{(a)} - \mathbf{D}^{(b)}) = 0 \quad (3.1.11)$$

The normal component of the electric field, however, is not continuous. Just consider the step between vacuum and a dielectric:

$$\hat{n} \cdot (\epsilon_o \mathbf{E}^{(a)} - (\epsilon_o \mathbf{E}^{(b)} + \mathbf{P}^{(b)})) = 0$$

Table 3.1. Inclusion of Electric and Magnetic Media Terms in Maxwell's Equations

$\nabla \cdot \varepsilon_o \mathbf{E} = -\nabla \cdot \mathbf{P}$	$\nabla \cdot \mu_o \mathbf{H} = -\nabla \cdot \mu_o \mathbf{M}$
$\mathbf{P} = -N\mathbf{e}r$	$\mathbf{M} = N\mathbf{m}$
$\nabla \cdot (\varepsilon_o \mathbf{E} + \mathbf{P}) = 0$	$\nabla \cdot (\mu_o \mathbf{H} + \mu_o \mathbf{M}) = 0$
$\nabla \times \mathbf{H} = \frac{\partial}{\partial t} (\varepsilon_o \mathbf{E} + \mathbf{P})$	$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} (\mu_o \mathbf{H} + \mu_o \mathbf{M})$
$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P}$	$\mathbf{B} = \mu_o \mathbf{H} + \mu_o \mathbf{M}$
$\mathbf{D} : (\text{C/m}^2)$	$\mathbf{B} : (\text{Vs/m}^2)$
$\hat{\mathbf{n}} \cdot (\mathbf{D}^{(b)} - \mathbf{D}^{(a)}) = 0$	$\hat{\mathbf{n}} \cdot (\mathbf{B}^{(b)} - \mathbf{B}^{(a)}) = 0$
$\hat{\mathbf{n}} \times (\mathbf{E}^{(b)} - \mathbf{E}^{(a)}) = 0$	$\hat{\mathbf{n}} \times (\mathbf{H}^{(b)} - \mathbf{H}^{(a)}) = 0$

The amplitude of $\mathbf{E}^{(b)}$ must take into account the strength of $\mathbf{P}^{(b)}$ for this relation to hold.

Using Gauss' law for the magnetic flux gives, via the same analysis

$$\hat{\mathbf{n}} \cdot (\mathbf{B}^{(a)} - \mathbf{B}^{(b)}) = 0 \quad (3.1.12)$$

The normal component of the magnetic flux density is continuous across an interface.

The tangential continuity conditions are derived from Faraday's and Ampère's laws. Consider a small loop that intersects the same interface (Fig. 3.1(b)). The area of the loop is A and its normal $\hat{\mathbf{i}}_n$ is parallel to the interface. Integration of (3.1.10a) and taking the limit of zero height yields

$$\lim_{h \rightarrow 0} \int_S \nabla \times \mathbf{E} \cdot d\mathbf{a} = - \lim_{h \rightarrow 0} \frac{\partial}{\partial t} \int_S \mathbf{B} \cdot d\mathbf{a}$$

Separate evaluation of the integrals gives

$$\lim_{h \rightarrow 0} \oint_C \mathbf{E} \cdot d\mathbf{s} = \lim_{h \rightarrow 0} \hat{\mathbf{i}}_s \cdot (\mathbf{E}^{(a)} - \mathbf{E}^{(b)}) L$$

and

$$- \lim_{h \rightarrow 0} \frac{\partial}{\partial t} \int_S \mathbf{B} \cdot d\mathbf{a} = - \lim_{h \rightarrow 0} \frac{\partial}{\partial t} \hat{\mathbf{i}}_n \mathbf{B} h L$$

Using the identities $\hat{\mathbf{i}}_s = \hat{\mathbf{i}}_n \times \hat{\mathbf{n}}$ and $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$, the integrals produce the continuity law for the tangential electric field:

$$\hat{\mathbf{n}} \times (\mathbf{E}^{(a)} - \mathbf{E}^{(b)}) = 0 \quad (3.1.13)$$

The same analysis of Ampère's law gives the continuity condition for the tangential magnetic field:

$$\hat{\mathbf{n}} \times (\mathbf{H}^{(a)} - \mathbf{H}^{(b)}) = 0 \quad (3.1.14)$$

Table 3.1 summarizes the results of this section.

3.2 Constitutive Relation Tensors

The preceding section introduced polarization $\mathbf{P}$ and magnetization $\mu_o\mathbf{M}$ terms into Maxwell's equations. These terms are derived from the radiation fields of bound electrons within a dielectric medium. The radiation fields are themselves excited by incident electromagnetic waves; the combined incident and radiated fields inextricably co- and counter-propagate throughout the medium. The combined electric and magnetic fields are denoted by $\mathbf{D}$ and $\mathbf{B}$ and are called the electric and magnetic flux densities, respectively. Maxwell's equations including the flux densities are

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial}{\partial t} \mathbf{B}(\mathbf{r}, t) \quad (3.2.1a)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\partial}{\partial t} \mathbf{D}(\mathbf{r}, t) \quad (3.2.1b)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, t) = 0 \quad (3.2.1c)$$

$$\nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 \quad (3.2.1d)$$

A specific medium is modelled by the relation between the field terms $\mathbf{E}$ and $\mathbf{H}$ and the flux density terms $\mathbf{D}$ and $\mathbf{B}$. The most general form of these constitutive relations is

$$\begin{pmatrix} c\mathbf{D} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{M} & \mathbf{Q} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ c\mathbf{B} \end{pmatrix} \quad (3.2.2)$$

where c is the speed of light and $\mathbf{P}$, $\mathbf{L}$, $\mathbf{M}$, and $\mathbf{Q}$ are 3×3 matrix tensors. The form of (3.2.2) is preferred for its invariance to relativistic transformation [11]. The constitutive tensors are generally frequency dependent, and when cast in time-harmonic form, are generally complex quantities.

Materials are classified according to the matrix entries of (3.2.2). When the cross-coupling terms $\mathbf{L}$ and $\mathbf{M}$ are non-zero, the medium is called bianisotropic. Optically active materials are bianisotropic. When the cross-coupling terms are zero ($\mathbf{L} = \mathbf{M} = 0$) then the medium is anisotropic. Within anisotropic materials the electric field excites only the electric flux, and the magnetic field excites only the magnetic flux. These excitations are generally not spatially uniform and depend on the eigenvectors of $\mathbf{P}$ and $\mathbf{Q}$. Isotropy is a special

case of anisotropy in that $\mathbf{P}$ and $\mathbf{Q}$ are diagonal tensors with all entries equal. Physically, excitation of dipole moments is spatially uniform for isotropic materials.

The materials considered in this text are lossless to first order. Losslessness imposes certain symmetry conditions on the constitutive tensors. These conditions are determined by Poynting's conservation theorem. The Poynting's theorem derived from time-harmonic versions of (3.2.1a,b) is

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}^*) = j\omega (\mathbf{E} \cdot \mathbf{D}^* - \mathbf{H}^* \cdot \mathbf{B})$$

Losslessness requires that the divergence of time-averaged power flow vanishes:

$$\langle \nabla \cdot \mathbf{S} \rangle = \frac{1}{2} \Re \left[j\omega (\mathbf{E} \cdot \mathbf{D}^* - \mathbf{H}^* \cdot \mathbf{B}) \right] = 0 \quad (3.2.3)$$

To evaluate the impact of this constraint on the constitutive relations, expression (3.2.2) is re-ordered to put the fields on one side and the flux densities on the other:

$$\begin{pmatrix} \mathbf{D} \\ \mathbf{B} \end{pmatrix} = \mathbf{C}_{EH} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \mathbf{C}_{DB} \begin{pmatrix} \mathbf{D} \\ \mathbf{B} \end{pmatrix} \quad (3.2.4)$$

where

$$\mathbf{C}_{EH} = \begin{pmatrix} \varepsilon & \xi \\ \zeta & \mu \end{pmatrix}, \quad \text{and} \quad \mathbf{C}_{DB} = \begin{pmatrix} \kappa & \xi^i \\ \zeta^i & \nu \end{pmatrix} \quad (3.2.5)$$

The entries of $\mathbf{C}_{EH}$ and $\mathbf{C}_{DB}$ are, generally, 3×3 tensors.

Fluxes $\mathbf{D}$ and $\mathbf{B}$ in (3.2.3) can now be expressed in terms as $\mathbf{E}$ and $\mathbf{H}$. With the identity $\mathbf{E} \cdot \varepsilon^* \cdot \mathbf{E}^* = \mathbf{E}^* \cdot \varepsilon^\dagger \cdot \mathbf{E}$ and a similar one for $\mathbf{H}$ and μ , the lossless condition imposes the following constraints on the tensors of $\mathbf{C}_{EH}$ and $\mathbf{C}_{DB}$:

$$\varepsilon = \varepsilon^\dagger, \quad \mu = \mu^\dagger, \quad \xi = \zeta^\dagger \quad (3.2.6)$$

and

$$\kappa = \kappa^\dagger, \quad \nu = \nu^\dagger, \quad \xi^i = \zeta^{i\dagger} \quad (3.2.7)$$

Losslessness forces the on-diagonal tensors to be Hermitian and the off-diagonal tensors to possess symmetry. Accordingly, transmission through any such media retains a real-valued dispersion relation and is therefore lossless.

Tensors ξ and ζ are non-zero for bianisotropic media and identically zero for anisotropic media. Lossless isotropic materials have scalar (ε, μ) , rather than tensor (ε, μ) values. In anisotropic and isotropic materials, ε and μ are readily identified as the electric permittivity and magnetic permeability tensors. Likewise, κ and ν are identified as the impermittivity and impermeability tensors, where

$$\kappa = \varepsilon^{-1}, \quad \text{and} \quad \nu = \mu^{-1} \quad (3.2.8)$$

For anisotropic and isotropic materials, the fields and flux densities are decoupled:

$$\mathbf{D} = \boldsymbol{\varepsilon} \cdot \mathbf{E} \quad (3.2.9a)$$

$$\mathbf{B} = \boldsymbol{\mu} \cdot \mathbf{H} \quad (3.2.9b)$$

Only when $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ are scalars are the fluxes necessarily aligned to the fields. Otherwise, fields and fluxes align along the eigenvectors of their respective tensors.

Finally, Poynting's theorem restated to include explicit time dependence is

$$\nabla \cdot \mathbf{S} + \frac{\partial W}{\partial t} = 0 \quad (3.2.10)$$

where the total stored energy is

$$W = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}) \quad (3.2.11)$$

In the time-domain picture, losslessness requires an instantaneous response of the polarization and magnetization densities to the applied electro-magnetic field; a phase lag of $\mathbf{D}$ to $\mathbf{E}$, or $\mathbf{B}$ to $\mathbf{H}$, generates loss in the medium.

3.3 The *kDB* System

Section §3.1 showed that within a medium the incident fields and induced dipole radiation propagate together as electric and magnetic flux densities. These flux densities are the natural terms in which to describe the electro-magnetic behavior. There is a second reason why these are the natural terms, and that is because the *k*-vector always lies perpendicular to plane-wave solutions of $\mathbf{D}$ and $\mathbf{B}$. One could instead consider using the fields $\mathbf{E}$ and $\mathbf{H}$ as descriptors within a medium since the Poynting vector always lies perpendicular to the fields: $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. The choice of which system to use is made by looking for mathematical simplicity, not on physical grounds. Since the *k*-vector always lies perpendicular to $\mathbf{D}$ and $\mathbf{B}$, one of three vectors drops out of Maxwell's equations. Thus the so-called *kDB* system is introduced. Once the fluxes are everywhere determined, the fields are calculated from the inverse constitutive relations and, ultimately, the Poynting vector is determined.

The *kDB* system is constructed to take advantage of that natural basis defined by the triplet $(\mathbf{k}, \mathbf{D}, \mathbf{B})$. Plane-wave solutions separate the oscillatory term $\exp\{j(\omega t - \mathbf{k} \cdot \mathbf{r})\}$ from a slowly varying and (piecewise) spatially independent flux envelope. Maxwell's equations (3.2.1) are thus recast as

$$\mathbf{k} \times \mathbf{E} = \omega \mathbf{B} \quad (3.3.1a)$$

$$\mathbf{k} \times \mathbf{H} = -\omega \mathbf{D} \quad (3.3.1b)$$

$$\mathbf{k} \cdot \mathbf{D} = 0 \quad (3.3.1c)$$

$$\mathbf{k} \cdot \mathbf{B} = 0 \quad (3.3.1d)$$

The natural coordinates for these flux vectors and k -vector are $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$. In particular, $\hat{e}_3$ always points along the k -vector ($\mathbf{k} = \hat{e}_3 k$) and the $\mathbf{D}$ and $\mathbf{B}$ vectors line in the DB plane defined by $(\hat{e}_1, \hat{e}_2)$ normal to $\hat{e}_3$ (Fig. 3.2). The result is that $D_3 = B_3 = 0$ when resolved on the kDB coordinate system. This provides the promised simplification.

Typically the constitutive tensors are written along their eigenvectors. The permittivity tensor for a birefringent crystal, for example, is generally written with only diagonal entries. However, the flux densities can propagate in an arbitrary direction. To reconcile the two reference frames, the eigenvector frame of the tensors, denoted by coordinate system (x, y, z) , is rotated into the natural frame of the fluxes and k -vector. This transformation is done with the rotation operator T , which defines a first rotation about z and a second rotation about $\hat{e}_1$, the latter being the local x axis. Thus

$$\begin{aligned} T &= R_x(\theta)R_z(\phi) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

For vectors A and A_k resolved in the crystal and flux coordinates, respectively, the forward and inverse vector transformations are

$$A_k = TA, \quad \text{and} \quad A = T^{-1}A_k$$

where

$$T = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \cos \theta \sin \phi & \cos \theta \cos \phi & -\sin \theta \\ \sin \theta \sin \phi & \sin \theta \cos \phi & \cos \theta \end{pmatrix} \quad (3.3.2)$$

and $T^{-1} = T^T$. The operators T and T^{-1} are unitary: $TT^{-1} = I$.

When acting on matrices rather than vectors, the transformation operator T imparts a similarity transform on the coordinate system, cf. §2.4.4. Given the kDB constitutive relation

$$\begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \begin{pmatrix} \kappa & \xi^i \\ \zeta^i & \nu \end{pmatrix} \begin{pmatrix} \mathbf{D} \\ \mathbf{B} \end{pmatrix} \quad (3.3.3)$$

resolved on an underlying (x, y, z) coordinate system, the constitutive relation is rotated into the flux frame according to

$$\begin{pmatrix} \mathbf{E}_k \\ \mathbf{H}_k \end{pmatrix} = \begin{pmatrix} T \cdot \kappa \cdot T^{-1} & T \cdot \xi^i \cdot T^{-1} \\ T \cdot \zeta^i \cdot T^{-1} & T \cdot \nu \cdot T^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{D}_k \\ \mathbf{B}_k \end{pmatrix} \quad (3.3.4a)$$

$$= \begin{pmatrix} \kappa_k & \xi_k^i \\ \zeta_k^i & \nu_k \end{pmatrix} \begin{pmatrix} \mathbf{D}_k \\ \mathbf{B}_k \end{pmatrix} \quad (3.3.4b)$$

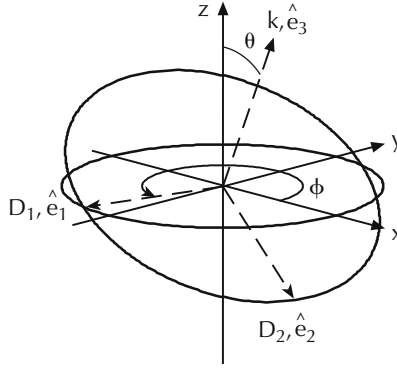


Fig. 3.2. The kDB coordinate system is written relative to a (x, y, z) coordinate system that is typically aligned to the eigenvectors of the constitutive tensors. First, a right-hand rotation about $\hat{z}$ by ϕ , then a right-hand rotation about the $\hat{e}_1$ axis by θ . Axis $\hat{e}_3$ is aligned to the k -vector, and $\hat{e}_1$ always lies in the (x, y) plane.

Each constitutive relation tensor is transformed by the coordinate-system change, but since T is unitary, there is no dilation or compression of the tensors, only a pure rotation.

Now, since $\mathbf{k} = k \hat{e}_3$ and $D_3 = B_3 = 0$, the fields $\mathbf{E}$ and $\mathbf{H}$ can be eliminated in (3.3.1a,b) to solve for $D_{1,2}$ and $B_{1,2}$. The resulting coupled equations are

$$\begin{pmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = - \begin{pmatrix} \xi_{11}^i & \xi_{12}^i - u \\ \xi_{21}^i + u & \xi_{22}^i \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \quad (3.3.5a)$$

$$\begin{pmatrix} \nu_{11} & \nu_{12} \\ \nu_{21} & \nu_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = - \begin{pmatrix} \zeta_{11}^i & \zeta_{12}^i + u \\ \zeta_{21}^i - u & \zeta_{22}^i \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \quad (3.3.5b)$$

where the phase velocity is defined as

$$u = \omega/k \quad (3.3.6)$$

As written, the tensor components κ_{ij} , ν_{ij} , ξ_{ij}^i , and ζ_{ij}^i are resolved in the kDB system. Depending on the complexity of the material, (3.3.5a,b) can be solved simultaneously, or either $\mathbf{D}_k$ or $\mathbf{B}_k$ can be eliminated and the resulting equation can be solved.

Equations (3.3.5a,b) are of the utmost importance to the description of electromagnetic propagation within dielectric media. These equations will be used in the following to determine the eigenmodes of propagation within a medium, the phase and group velocities, the dispersion relations, and the direction of energy flow. The elegance of the kDB system is that all linear, lossless materials can be described by this one system of equations, the particulars depending only on the tensor entries.

3.4 The Lorentz Force

As outlined in the introduction to the chapter, constitutive relations for isotropic, birefringent, gyrotropic, and optically active materials are derived from classical electron-wave interaction in the following. The Lorentz force and Newton's second law are sufficient to model simple behavior of isotropic and anisotropic media. Anisotropic media, of course, includes birefringent and gyrotropic materials. The constitutive relation for optical activity is derived heuristically as any robust model requires more work than space allows.

In isotropic and anisotropic materials the electric field couples to the bound electrons through the Lorentz force. The Lorentz force $\mathbf{f}$ is defined by

$$\mathbf{f} = -e \left(\mathbf{E} + \frac{\partial \mathbf{r}}{\partial t} \times \mathbf{B} \right) \quad (3.4.1)$$

where $(-e)$ is the charge of the electron and $\mathbf{r}$ is the displacement of the electron from its neutral position. Since the electrons in lossless material are bound to the nucleus, the charge-attraction serves as a restoring force. For low-intensity light the electron displacement is purely linear, so one can take the restoring force K as constant with displacement $\mathbf{r}$. Newton's second law $\mathbf{f} = m\mathbf{a}$ generates the equation of motion for a bound electron driven by the Lorentz force:

$$m \frac{\partial^2 \mathbf{r}}{\partial t^2} = -K\mathbf{r} - e\mathbf{E} - e \frac{\partial \mathbf{r}}{\partial t} \times \mathbf{B} \quad (3.4.2)$$

In the absence of an externally applied magnetic field, the cross-product force $\mathbf{r} \times \mathbf{B}$ is vanishingly small compared to $\mathbf{E}$. Indeed, recast in time-harmonic form and substituting in Faraday's law gives $\mathbf{r} \times (\mathbf{k} \times \mathbf{E}) \rightarrow |r\omega/c| \mathbf{E}$. For $r \sim 0.1\text{\AA}$ and $\omega \sim 2\pi \times 200\text{ THz}$, $|rk| < 10^{-5}$. To turn the cross-product term "on" either an intense external magnetic field must be applied or the internal magnetization must be high, as is the case for materials with magnetic domains.

3.5 Isotropic Materials

The equation of motion for a bound electron in a linear isotropic material is

$$m \frac{\partial^2 \mathbf{r}}{\partial t^2} = -K\mathbf{r} - e\mathbf{E} - m\gamma \frac{\partial \mathbf{r}}{\partial t} \quad (3.5.1)$$

where γ is a resonant damping coefficient, and the restorative and drag forces on the electron are centro-symmetric. The magnetic-induced force present in (3.4.2) is neglected. When an incident field is a plane wave at frequency ω , (3.5.1) is rewritten in time-harmonic form:

$$(-m\omega^2 + jm\gamma\omega + K) \mathbf{r} = -e\mathbf{E} \quad (3.5.2)$$

The following subsection incorporates this equation of motion into Maxwell's equations to obtain a dispersion relationship and the functional forms for the material permittivity, refractive index, and absorption.

3.5.1 Permittivity of Isotropic Materials

Equation (3.1.2) relates the material polarization to the electron displacement. Multiplying (3.5.2) by Ne , the frequency dependence of the polarization density is

$$\mathbf{P} = \frac{Ne^2}{K - m\omega^2 + jm\gamma\omega} \mathbf{E} \quad (3.5.3)$$

The polarization $\mathbf{P}$ is a linear function of the applied electric field $\mathbf{E}$ and points in the same direction. The frequency-dependent coefficient is a resonant factor where the resonance frequency ω_o is defined as

$$\omega_o = \sqrt{K/m} \quad (3.5.4)$$

The resonance frequency is not a function of the field energy or intensity (to this order of approximation) and is therefore a fixed quantity that depends on the composition of the material.

The material susceptibility χ_e is the tensor that relates $\mathbf{E}$ to $\mathbf{P}$ for linear media. For isotropic media the tensor is a scalar χ_e . The field and polarization are thus related via

$$\mathbf{P} = \varepsilon_o \chi_e(\omega) \mathbf{E} \quad (3.5.5)$$

Identifying (3.5.5) with (3.5.3), the complex susceptibility of a simple isotropic material is

$$\chi_e(\omega) = \frac{1}{m\varepsilon_o} \frac{Ne^2}{\omega_o^2 - \omega^2 + j\gamma\omega} \quad (3.5.6)$$

The susceptibility tensor or scalar is fundamental because it embodies the homogeneous response of a material.

Susceptibility χ_e is used to define permittivity in the following way. The electric-flux density $\mathbf{D}$ is the combination of the incident field and its induced dipole polarization:

$$\mathbf{D} = \varepsilon_o \mathbf{E} + \mathbf{P} = \varepsilon(\omega) \mathbf{E} \quad (3.5.7)$$

where (3.5.5) is used to define the material permittivity ε accordingly:

$$\varepsilon(\omega) = \varepsilon_o (1 + \chi_e(\omega)) \quad (3.5.8)$$

When the susceptibility is a tensor so is the permittivity. Often the permittivity relative to vacuum is used to compare materials. The relative permittivity ε_r is defined as

$$\varepsilon_r = \varepsilon/\varepsilon_o \quad (3.5.9)$$

By analogy to permittivity, material permeability μ is defined as $\mathbf{B} = \mu \mathbf{H}$. However, for materials considered here there is little or no interaction between the field and the magnetic moments of the material. Therefore, μ is treated as a scalar, frequency-independent quantity.

With the above definitions of the material permittivity and permeability, the Helmholtz wave equation is (compare (1.1.6) on page 3)

$$\nabla^2 \mathbf{E} = \mu \varepsilon \frac{\partial^2}{\partial t^2} \mathbf{E} \quad (3.5.10)$$

Plane-wave solutions of the form

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_o \exp \left(j \left(\omega t - \tilde{\mathbf{k}} \cdot \mathbf{r} \right) \right)$$

when inserted into (3.5.10), generate the dispersion relation

$$\tilde{k} = \omega \sqrt{\mu \varepsilon}$$

Now, the permittivity ε is a complex scalar, so $\tilde{k}$ is complex as well. That means absorption as well as retardation are fundamental implications of the Lorentz equation of electron motion.

The frequency dependence of the loss and refractive index of an isotropic material is found as follows. Expansion of the permittivity and dispersion relation into real and imaginary parts,

$$\begin{aligned} \varepsilon &= \varepsilon' + j\varepsilon'' \\ (k + j\alpha) &= \frac{\omega}{c} (n + j\kappa) \end{aligned}$$

gives an expression for the real and imaginary parts of the relative permittivity:

$$(n + j\kappa)^2 = \varepsilon'_r + j\varepsilon''_r,$$

The real and imaginary parts of the relative permittivity are identified with the refractive index n and extinction coefficient κ as

$$\varepsilon'_r = n^2 - \kappa^2, \quad \text{and} \quad \varepsilon''_r = 2n\kappa \quad (3.5.11)$$

Finally, expanding the resonant form of the susceptibility χ_e into its real and imaginary parts and identifying with (3.5.11) gives the coupled equations for the refractive index and extinction coefficient as

$$n^2 - \kappa^2 = 1 + \frac{Ne^2}{m\varepsilon_o} \frac{\omega_o^2 - \omega^2}{(\omega_o^2 - \omega^2)^2 + (\gamma\omega)^2} \quad (3.5.12)$$

and

$$2n\kappa = -\frac{Ne^2}{m\varepsilon_o} \frac{\gamma\omega}{(\omega_o^2 - \omega^2)^2 + (\gamma\omega)^2} \quad (3.5.13)$$

In the transparent regime, the bound-electron resonance is far away from the frequency of the incident field. In this case, $\kappa \simeq 0$ and the dispersion relation is approximately

$$k = \frac{\omega}{c} n \quad (3.5.14)$$

The phase velocity of the plane wave as it propagates within the medium is $v_{\text{ph}} = c/n$. Inspection of (3.5.12) in light of $\kappa \simeq 0$ shows that the refractive

index n is greater than unity. This is generally the case, although exceptions are possible, such as in the x-ray region where multiple material resonances are below, rather than above, the excitation frequency. For the near-infrared transparent regime important to telecommunications, the refractive index is greater than one. Another implication of polarization of the material is the wavelength change within the medium. The wavelength in the material λ is related to the free-space wavelength λ_o as

$$\lambda = \lambda_o/n \quad (3.5.15)$$

That the wavelength is reduced is not surprising since the frequency of the light does not change whether the light is in vacuum or material, and $\lambda\omega = 2\pi v_{\text{ph}}$ always holds.

The phase velocity v_{ph} is the velocity of a monochromatic wavefront. However, the velocity of energy flow of a narrowband pulse is related to the frequency derivative of the wavenumber:

$$\frac{1}{v_g} = \frac{dk}{d\omega} \quad (3.5.16)$$

where v_g is the group velocity. As has been shown, even a simple dielectric material has permittivity dispersion, so the phase and group velocities generally differ. The group index n_g , defined by $v_g = c/n_g$, is

$$n_g = n - \lambda \frac{dn}{d\lambda} \quad (3.5.17)$$

Generally, material and waveguide dispersion has a negative index slope with wavelength. So, generally, the group index lies above the refractive index.

Now, the key simplification so far has been that a single resonance exists within the isotropic material. In general this is not the case. Multiple resonances can be related to various absorption bands of the atoms or molecules that constitute the material. A more robust model includes multiple resonances and weighs the contributions to the susceptibility according to the fraction of atoms that are associated with each resonance. Well into the transparent regime where the damping factor can be ignored, the refractive index square is modelled as

$$n^2(\omega) = 1 + \frac{Ne^2}{m\varepsilon_o} \sum_n \frac{f_n}{\omega_o^2 - \omega^2}$$

Often an additional pole at low frequency and another at high frequency are added based on phenomenological experience. The refractive index equation is then

$$n^2(\omega) = 1 + \frac{a_\infty}{\omega_o^2} + \frac{Ne^2}{m\varepsilon_o} \sum_n \frac{f_n}{\omega_o^2 - \omega^2} - \frac{a_o}{\omega^2}$$

Table 3.2. Atomic Lines for Abbe Number Definition

Wavelength (nm)	Symbol	Spectral line	Element
656.2725	C	Red hydrogen line	H
589.2938	d	Yellow sodium line	Na
587.5618	d	Yellow helium line	He
486.1327	F	Blue hydrogen line	H

Converting frequency ω to wavelength λ and introducing material-specific fitting coefficients, the Sellmeier equation for refractive index dispersion is

$$n^2(\lambda) = A + \sum_n \frac{B_n \lambda^2}{\lambda^2 - C_n} - D \lambda^2 \quad (3.5.18)$$

There are, in fact, any number of forms of this equation. While the measured refractive index data is absolute, the value of the coefficients in (3.5.18) depends on which equation is used to model the data. As an example of another form of the Sellmeier equation, the equation published by Schott Glass is

$$n^2(\lambda) - 1 = \frac{B_1 \lambda^2}{\lambda^2 - C_1} + \frac{B_2 \lambda^2}{\lambda^2 - C_2} + \frac{B_3 \lambda^2}{\lambda^2 - C_3} \quad (3.5.19)$$

The Abbe number is a measure of the refractive index dispersion throughout the visible region. The Abbe number generally works well for glasses rather than semiconductor or crystals because the amorphous, homogeneous nature of glass precludes strong resonances. The Abbe number v_d is defined on the d-line as

$$v_d = \frac{n_d - 1}{n_F - n_C} \quad (3.5.20)$$

where n_d , n_F , and n_C are the measured refractive indices on the d, F, and C lines, respectively. The wavelengths for these lines are defined in Table 3.2.

3.5.2 Propagation in Isotropic Materials

Even though the kDB system is intended for more complex propagation calculations, propagation within an isotropic material will be cast in this formalism for the sake of example. Within a material, one needs to find the directions of the k -vector and the Poynting vector, as well as the wavenumber. Also, the refraction properties into and out of the material needs to be determined; that is the topic of the following section.

In an isotropic material, tensors ξ^i and ξ^i in $\mathbf{C}_{DB}$ are zero, and κ and ν are scalars. The constitutive relations are

$$\begin{aligned} \mathbf{E} &= \kappa \mathbf{D} \\ \mathbf{H} &= \nu \mathbf{B} \end{aligned} \quad (3.5.21)$$

The underlying coordinate system is arbitrary. A rotation of the coordinate system into kDB coordinates leaves the constitutive relations unchanged since $T \cdot \kappa T^{-1} = \kappa$. Thus, substitution of (3.5.21) into the coupled kDB equations (3.3.5) gives

$$\kappa \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \quad (3.5.22a)$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \quad (3.5.22b)$$

Elimination of $\mathbf{B}$ generates the governing equation

$$\begin{pmatrix} u^2 - \nu\kappa & 0 \\ 0 & u^2 - \nu\kappa \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.5.23)$$

The electric field may be polarized along the $\hat{e}_1$ direction, the $\hat{e}_2$ direction, or a mixture of the two. For example, when the field is polarized along $\hat{e}_1$ then $D_1 \neq 0$ and $D_2 = 0$. In any case, the governing equation is satisfied when the phase velocity is

$$u = \sqrt{\nu\kappa} = \frac{1}{\sqrt{\mu\varepsilon}} \quad (3.5.24)$$

Substituting $u = \omega/k$ gives the dispersion relationship

$$k = \omega\sqrt{\mu\varepsilon} \quad (3.5.25)$$

from which the refractive index can be extracted:

$$n = \sqrt{\mu_r \varepsilon_r} \quad (3.5.26)$$

It is important to note that $|\mathbf{k}| = k$ and $k = \omega n/c$. Geometrically, the k -vector within the material can point in any direction, while the vector length is always k . The surface defined by all possible pointing directions of the k -vector is a sphere of radius $\omega n/c$:

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2} \quad (3.5.27)$$

Finally, the Poynting vector is determined by the fields. Since κ and ν are scalars, the fields and respective flux densities are aligned. Since by definition the k -vector is aligned to $\hat{e}_3$, the Poynting vector is

$$\begin{aligned} \mathbf{S} &= \mathbf{E}_k \times \mathbf{H}_k \\ &= \nu\kappa \mathbf{D}_k \times \mathbf{B}_k \\ &= \frac{1}{\mu\varepsilon} \hat{e}_3 \end{aligned} \quad (3.5.28)$$

So, $\mathbf{k} \parallel \mathbf{S}$ in an isotropic material.

3.5.3 Refraction at an Interface

There are two questions to be addressed when dealing with refraction at a smooth interface: the angle of refraction, and the transmission and reflection coefficients. The refraction angle at an interface of two dissimilar isotropic materials is determined by the relative refractive indices alone and is independent of the polarization of the incident wave. The transmission and reflection coefficients, however, do depend on the incident polarization.

Phase matching of two waves, one in a first material and the other in a second material, is the single condition that must be satisfied along an interface. Refraction results when the refractive indices of the two materials differ. The canonical example is illustrated in Fig. 3.3(a). Here the plane wave k -vector in material 1 is inclined from the interface normal by θ_1 . The wavelength within the material is $\lambda_1 = \lambda_o/n_1$ and the wave period as projected onto the interface is $\lambda_1/\sin\theta_1$. The same analysis applies to the second material, with θ_2 and n_2 in replacement. Phase matching at the interface requires

$$\frac{\lambda_1}{\sin\theta_1} = \frac{\lambda_2}{\sin\theta_2} \quad (3.5.29)$$

Snell's law is a restatement of (3.5.29) using the refractive index instead:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad (3.5.30)$$

Since the refractive index of either (isotropic) material is independent of the polarization of the wave, the refraction angle is independent of polarization.

A geometric construction of refraction and reflection is illustrated in Fig. 3.3(b). Recall that within an isotropic medium the k -surface is a sphere having radius $\omega n/c$. The interface between two media bisects each sphere, dividing them into hemispheres of different radii. In Fig. 3.3(b) the refractive index of the first material is less than the second. The top and bottom half-circles represent the k -surfaces as projected onto the page, and the $\mathbf{k}_i$, $\mathbf{k}_t$, and $\mathbf{k}_r$ vectors are illustrated, subscripts i , t , and r referring to incident, transmitted, and reflected waves, respectively. Phase matching at the interface requires that the projection lengths of all three k -vectors match on the interface. In order for the tip of $\mathbf{k}_t$ to reach the $\omega n_2/c$ circle, the direction of $\mathbf{k}_t$ must change with respect to $\mathbf{k}_i$. For the reflected wave, however, $|\mathbf{k}_i| = |\mathbf{k}_r|$ and therefore the angle of the reflected wave is the same, albeit mirror-imaged about the normal to the interface. Finally, since these are isotropic materials, the Poynting and k -vectors for all three waves are aligned and coincident.

3.5.4 Reflection and Transmission for TE Waves

Next the reflection and transmission coefficients are determined. These coefficients are different for waves with transverse-electric (TE) and transverse-magnetic (TM) polarizations. The TE and TM polarization states are distinguished because, in the first case, the electric field oscillates in the plane of

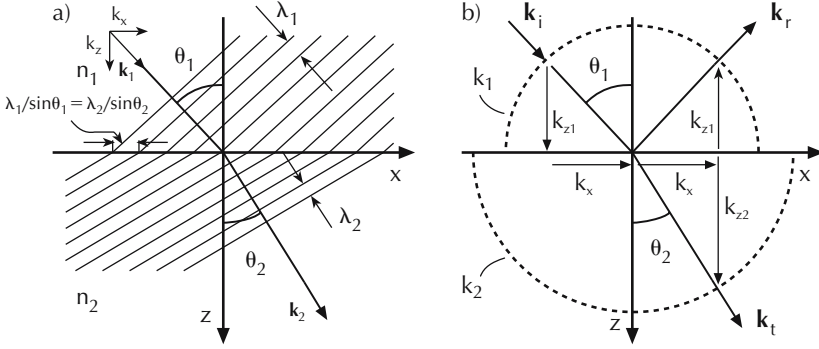


Fig. 3.3. Phase matching at a smooth dielectric interface. a) The free-space wavelength is reduced by the refractive index for waves within the dielectric. For $n_2 > n_1$ the wavelength in material 2 is shorter than that in material 1. Phase matching at the interface requires that the k -vector change direction at the interface. b) A k -vector diagram for refraction. The half-circle radii represent the respective refractive indices; the contours are circular in isotropic media. Phase matching requires the k_x vectors of both waves to match.

the interface while, in the second case, the magnetic field oscillates in that plane plane. The tangential continuity conditions (3.1.13-3.1.14) determine the relative field amplitudes in the two materials.

For TE plane waves, illustrated in Fig. 3.4(a), the total electric fields in the two materials are

$$\begin{aligned}\mathbf{E}_y^{(1)} &= \hat{y}E_o \left(e^{-jk_{z1}z} + \Gamma e^{jk_{z1}z} \right) e^{-jk_x x} \\ \mathbf{E}_y^{(2)} &= \hat{y}E_o T e^{-jk_{z2}z} e^{-jk_x x}\end{aligned}\quad (3.5.31)$$

where Γ and T are complex coefficients of the reflection and transmission amplitudes, respectively. Faraday's law determines the associated magnetic fields:

$$\mathbf{H} = \frac{1}{j\omega\mu_o} \left(\hat{x} \frac{\partial}{\partial z} E_y - \hat{z} \frac{\partial}{\partial x} E_y \right)$$

where $\mathbf{E}_y = \hat{y}E_y$. The magnetic field components are therefore

$$\begin{aligned}\mathbf{H}_x^{(1)} &= -\hat{x} \frac{k_{z1}}{\omega\mu_o} E_o \left(e^{-jk_{z1}z} - \Gamma e^{jk_{z1}z} \right) e^{-jk_x x} \\ \mathbf{H}_x^{(2)} &= -\hat{x} \frac{k_{z2}}{\omega\mu_o} E_y^{(2)}\end{aligned}$$

and

$$\mathbf{H}_z^{(1,2)} = \hat{z} \frac{k_x}{\omega\mu_o} E_y^{(1,2)}$$

Only the tangential $\mathbf{E}$ and $\mathbf{H}$ field components are required to satisfy continuity across the interface. Application of the boundary conditions

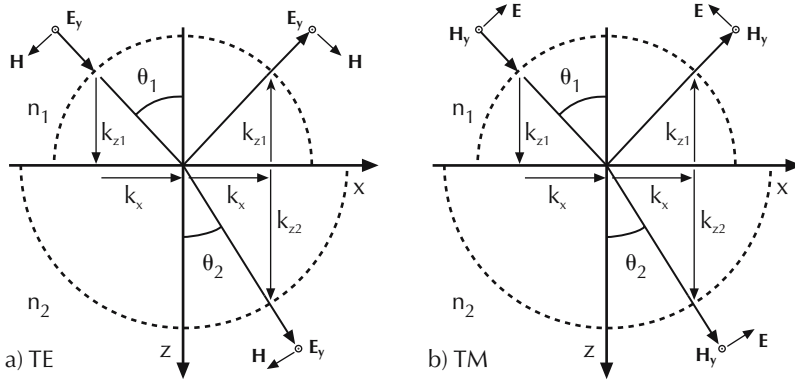


Fig. 3.4. Refraction diagrams for TE and TM waves. a) TE wave: the electric field lies tangential to the interface. b) TM wave: the magnetic field lies tangential to the interface.

$$\mathbf{E}_y^{(1)} = \mathbf{E}_y^{(2)} \quad \text{and} \quad \mathbf{H}_x^{(1)} = \mathbf{H}_x^{(2)}$$

on the interface where $x = 0$ gives

$$1 + \Gamma = T \quad (3.5.32a)$$

$$k_{z1} (1 - \Gamma) = k_{z2} T \quad (3.5.32b)$$

Combining (3.5.32a,b) gives the TE reflection coefficient in terms of material indices and ray angles:

$$\Gamma_{\text{TE}} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \quad (3.5.33)$$

or, in terms of n_1 and θ_1 alone,

$$\Gamma_{\text{TE}} = \frac{\sqrt{1 - \sin^2 \theta_1} - (n_2/n_1) \sqrt{1 - (n_1/n_2)^2 \sin^2 \theta_1}}{\sqrt{1 - \sin^2 \theta_1} + (n_2/n_1) \sqrt{1 - (n_1/n_2)^2 \sin^2 \theta_1}} \quad (3.5.34)$$

Now, if one considers a box drawn around the point of incidence, then all the power that flows into the box from the incident wave must flow out of the box via the reflected and transmitted waves. The time-averaged Poynting vectors are

$$\begin{aligned} \langle S_i \rangle &= \frac{1}{2\omega\mu_o} (\hat{x}k_x + \hat{z}k_{z1}) |E_o|^2 \\ \langle S_r \rangle &= \frac{1}{2\omega\mu_o} (\hat{x}k_x - \hat{z}k_{z1}) |\Gamma|^2 |E_o|^2 \\ \langle S_t \rangle &= \frac{1}{2\omega\mu_o} (\hat{x}k_x + \hat{z}k_{z2}) |T|^2 |E_o|^2 \end{aligned}$$

Power conservation requires that

$$|\langle S_i \rangle| = |\langle S_r \rangle| + |\langle S_t \rangle| \quad (3.5.35)$$

Substitution of the TE Poynting vectors gives

$$1 = |\Gamma|^2 + \frac{n_2}{n_1} |T|^2$$

Defining $R = |\Gamma|^2$ and $T = n_2/n_1 |T|^2$ generates the power conservation equation:

$$R + T = 1 \quad (3.5.36)$$

Clearly all that is reflected and transmitted must come from the incident power.

3.5.5 Reflection and Transmission for TM Waves

For TM plane waves, illustrated in Fig. 3.4(b), the magnetic field oscillates in the plane of the interface. In analogy to the TE wave solution, the total magnetic fields in the two materials are

$$\begin{aligned} \mathbf{H}_y^{(1)} &= \hat{y} H_o \left(e^{-jk_{z1}z} + \Gamma e^{jk_{z1}z} \right) e^{-jk_x x} \\ \mathbf{H}_y^{(2)} &= \hat{y} H_o T e^{-jk_{z2}z} e^{-jk_x x} \end{aligned} \quad (3.5.37)$$

where Γ and T are complex coefficients of the TM-wave reflection and transmission amplitudes, respectively. The associated electric fields are determined from Ampère's law,

$$\mathbf{E} = -\frac{1}{j\omega\epsilon} \left(\hat{x} \frac{\partial}{\partial z} H_y - \hat{z} \frac{\partial}{\partial x} H_y \right)$$

where $\mathbf{H}_y = \hat{y} H_y$. Solving for the fields and matching the tangential continuity condition across the interface yields

$$1 + \Gamma = T \quad (3.5.38a)$$

$$\frac{k_{z1}}{\epsilon_1} (1 - \Gamma) = \frac{k_{z2}}{\epsilon_2} T \quad (3.5.38b)$$

Combining (3.5.38a,b) gives the TM reflection coefficient in terms of material indices and ray angles:

$$\Gamma_{\text{TM}} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2} \quad (3.5.39)$$

or, in terms of n_1 and θ_1 ,

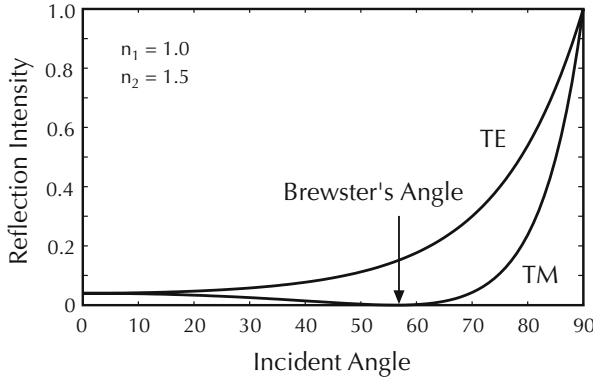


Fig. 3.5. Reflection intensities for TE and TM waves as a function of incident angle.

$$\Gamma_{\text{TM}} = -\frac{\sqrt{1 - \sin^2 \theta_1} - (n_1/n_2)\sqrt{1 - (n_1/n_2)^2 \sin^2 \theta_1}}{\sqrt{1 - \sin^2 \theta_1} + (n_1/n_2)\sqrt{1 - (n_1/n_2)^2 \sin^2 \theta_1}} \quad (3.5.40)$$

Figure 3.5 compares the reflection intensities $|\Gamma|^2$ for TE and TM waves and an air-glass boundary as a function of incident angle. The TE reflection intensity increases monotonically, while the TM wave goes through a zero point along the way. The angle at which TM reflection is zero is called Brewster's angle.

The remarkable property of TM waves is that at Brewster's angle all reflection is extinguished. Brewster's angle satisfies the condition

$$\theta_1 + \theta_2 = \pi/2 \quad (3.5.41)$$

That is, the transmitted and reflected waves are perpendicular to one another (see Fig. 3.6(a)). That there is no power in the reflected wave is reasonable based on physical considerations. The radiation pattern of an electric dipole is null along the polar axis. Brewster's condition orients the dipole excitation and the direction of the reflected wave perpendicular to one another. Substituting Brewster's condition (3.5.41) into (3.5.39) gives the requisite incident angle θ_B

$$\theta_B = \tan^{-1}(n_2/n_1) \quad (3.5.42)$$

Brewster's condition cannot be satisfied by TE waves because the electric field of the reflected and transmitted waves are always parallel.

To write the power conservation equation in the form of (3.5.36), the reflection and transmission coefficients are identified to be

$$R = |\Gamma|^2, \quad \text{and} \quad T = \frac{n_1}{n_2} |T|^2$$

As a concluding remark, in addition to incident angle and polarization state, only the refractive-index ratio n_2/n_1 determines the refraction angle, not the absolute refractive indices.

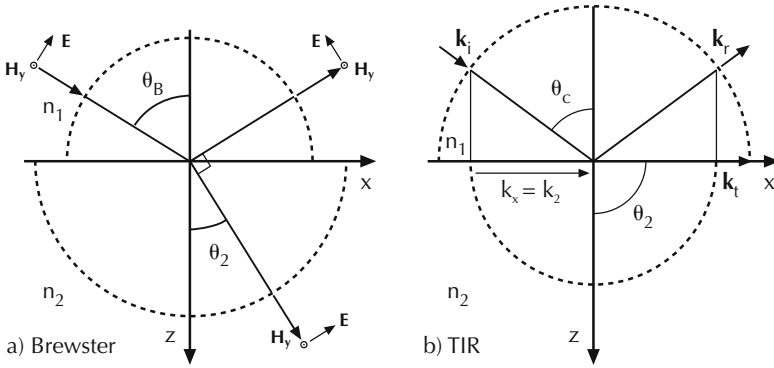


Fig. 3.6. a) The condition at Brewster's angle: the reflected and transmitted waves are perpendicular to one another. Brewster's condition of zero reflectance applies onto to TM waves. b) The critical angle for total internal reflection, $n_2 < n_1$.

3.5.6 Total Internal Reflection

Total internal reflection (TIR) occurs when the phase-matching condition at an interface cannot be satisfied by plane-wave solutions. A plane-wave solution does not exist when phase-matching requires a phase velocity that exceeds the speed of light. In this case all the incident power is reflected.

For TIR to exist, the refractive index in which the incident wave travels must exceed the refractive index beyond the interface: $n_1 > n_2$. This condition is required whether the material is isotropic or anisotropic. In isotropic media the refractive index is independent of polarization state, so TIR between two isotropic media is polarization agnostic.

For TIR to occur at an interface where $n_1 > n_2$, the incident angle must be at or beyond the critical angle. The critical angle for the incident wave is that angle which makes the transmitted wave run parallel to the interface $\theta_2 = \pi/2$ (see Fig. 3.6(b)). At and beyond the critical angle the wavefront period as projected onto the interface from the incident side is shorter than the wavelength of a plane wave on the transmission side. There is no orientation of $\mathbf{k}_t$ that achieves phase matching. Instead, k_t becomes imaginary and the transmitted field decays exponentially into the lower-index media.

Referring to Fig. 3.3(a), the angle where $\lambda_1 = \lambda_2 / \sin \theta_2$ is called the critical angle. Snell's law generates the relation

$$\theta_c = \sin^{-1} n_2/n_1 \quad (3.5.43)$$

Resolving k_2 into normal and transverse components,

$$k_2 = \sqrt{k_x^2 + k_{z2}^2},$$

at and above the critical angle $k_x > k_2$. This condition is only satisfied when k_z is imaginary:

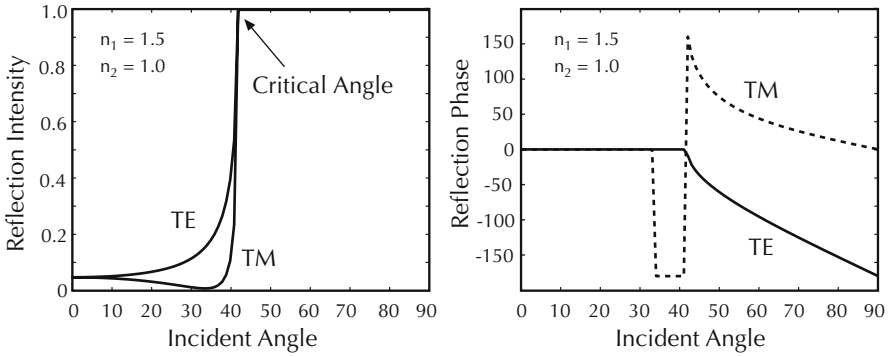


Fig. 3.7. Reflection intensity and phase before and after the onset of total internal reflection. a) Reflection intensity for TE and TM waves. b) Reflection phase shifts for TE and TM. Notice the sign change for TM reflection across the Brewster's angle boundary.

$$k_{z2} = j\alpha_{z2} = j\sqrt{k_x^2 - k_2^2}$$

This shows the exponential decay of the field along the axis normal to the surface. Parallel to the surface the incident and evanescent fields are phase matched.

In comparison to the reflected intensities from air to glass $n : 1.0 \rightarrow 1.5$ plotted in Fig. 3.5, the reflection intensity and phase for the reverse direction $n : 1.5 \rightarrow 1.0$ is plotted in Fig. 3.7. Since both media are isotropic the critical angle is the same for both polarizations. Below the critical angle of $\theta_c \simeq 41.8^\circ$ the reflection coefficients are below unity and there is no phase slip. At and beyond the critical angle, however, the reflection coefficients are unity and a phase develops between the incident and reflected waves. The phase of the TM wave goes through a π phase shift at Brewster's angle.

When the incident angle exceeds the critical angle, the reflection coefficient takes unity magnitude and develops a phase shift. The phase shift is called the Goos-Hänchen shift. Defining the reflection phase as $\angle \Gamma = 2\phi$, the boundary conditions for TE (3.5.32) and TM (3.5.38) are rewritten as

$$1 + e^{2j\phi} = T \quad (3.5.44a)$$

$$q_1(1 - e^{2j\phi}) = q_2 T \quad (3.5.44b)$$

where for TE $q = k_z$ and for TM $q = k_z/\varepsilon$. The Goos-Hänchen phase shifts are

$$\phi_{\text{TE}} = -\tan^{-1}\left(\frac{\alpha_{z2}}{k_{z1}}\right) \quad \text{and} \quad \phi_{\text{TM}} = -\tan^{-1}\left(\frac{\varepsilon_1 \alpha_{z2}}{\varepsilon_2 k_{z1}}\right) \quad (3.5.45)$$

Expanding in terms of the indices and incident angle, the reflection phases are

$$\angle\Gamma_{\text{TE}} = -2 \tan^{-1} \left[\frac{\sqrt{\sin^2 \theta_1 - (n_2/n_1)^2}}{\cos \theta_1} \right] \quad (3.5.46a)$$

$$\angle\Gamma_{\text{TM}} = -2 \tan^{-1} \left[\frac{\sqrt{\sin^2 \theta_1 - (n_2/n_1)^2}}{(n_2/n_1)^2 \cos \theta_1} \right] \quad (3.5.46b)$$

The retardance $\Delta\Gamma$ induced by a TIR reflection, $\Delta\Gamma = (\angle\Gamma_{\text{TE}} - \angle\Gamma_{\text{TM}})$, is

$$\Delta\Gamma = 2 \tan^{-1} \left(\frac{\cos \theta_1 \sqrt{\sin^2 \theta - (n_2/n_1)^2}}{\sin^2 \theta_1} \right) \quad (3.5.47)$$

Since the sign of $\Delta\Gamma$ is positive, the magnitude of phase shift imparted on the TM wave is greater than that for the TE wave.

Total internal reflection can be understood as a partial penetration of the electro-magnetic wave into the material of lower index (cf. Fig. 3.8). The field component normal to the interface is a standing wave where the first null lies within the lower-index material. The phase shift between the interface location and the first field maxima is called the Goos-Hänchen shift. This shift is polarization dependent, as determined above, with the TM wave penetration greater than the TE penetration for the same inclination. Since the Goos-Hänchen phases differ for TE and TM waves, the state of polarization of a reflected field can be transformed with respect to the incident field.

For TE waves, the electric field expressions above the critical angle are

$$\mathbf{E}_y^{(1)} = \hat{y} 2E_o e^{j\phi} \cos(k_z z + \phi) e^{-jk_x x} \quad (3.5.48a)$$

$$\mathbf{E}_y^{(2)} = \hat{y} E_o e^{-j\alpha_z z} e^{-jk_x x} \quad (3.5.48b)$$

The phase factor ϕ in the cosine term of the standing wave indicates that the last null of the standing wave lies below the interface and in region 2, Fig. 3.8(a). If the lower dielectric material were removed and a perfect metal conductor were placed at the location of the last null, the same standing wave pattern would persist.

The existence of the Goos-Hänchen phase shift alters the reflection diagrams of Fig. 3.4: the reflected light ray is no longer coincident with the incident ray at the interface. Rather, the incident ray penetrates into the lower material and reemerges forward-shifted with respect to its point of entry (Fig. 3.8(b)). The forward shift $2x_s$ is also rather confusingly called the Goos-Hänchen shift.

To construct the expression for the forward shift, consider two plane waves incident on the interface at slightly different angles $\theta_i \pm \Delta\theta$. The incident and reflected waves at $z = 0$ are

$$E_{y\pm}^{(i)} = E_o e^{-j(k_x \pm \Delta k_x)x}$$

$$E_{y\pm}^{(r)} = E_o e^{-j(k_x \pm \Delta k_x)x} e^{2j(\phi \pm \Delta\phi)}$$

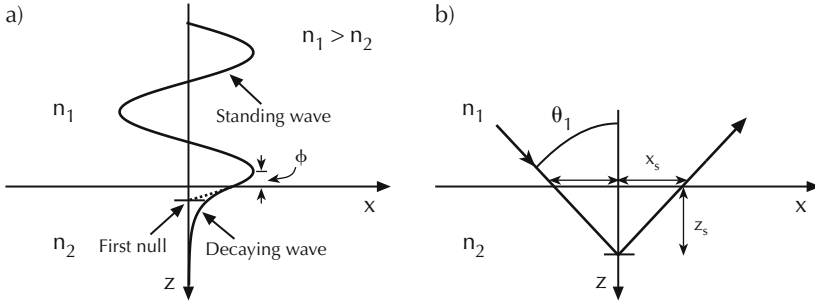


Fig. 3.8. Views of the Goos-Hänchen shift. a) A standing wave oscillates along the normal to the interface; the first null penetrates into the lower-index material. b) The inclined wave penetrates into the lower-index material and reemerges forward-shifted with respect to the incident wave. The penetration depth depends on the polarization state.

The total incident and reflected fields, summed over the two slightly different incident angles, are

$$E_y^{(i)} = 2E_o \cos(\Delta k_x x) e^{-jk_x x} \quad (3.5.49a)$$

$$E_y^{(r)} = 2E_o \cos(\Delta k_x x - 2\Delta\phi) e^{-jk_x x} \quad (3.5.49b)$$

Clearly the reflected field is shifted forward with respect to the incident field. Expanding the Goos-Hänchen shift about k_x gives

$$\phi(k + \Delta k_x) = \phi(k) + \frac{\partial\phi}{\partial k_x} \Delta k_x$$

Substitution into (3.5.49)(b) gives the expression for the reflected wave

$$E_y^{(r)} = 2E_o \cos(\Delta k_x (x - 2x_s)) e^{-jk_x x} \quad (3.5.50)$$

where the lateral Goos-Hänchen shift x_s is defined as

$$x_s = \frac{\partial\phi}{\partial k_x} \quad (3.5.51)$$

The Goos-Hänchen shifts for TE and TM reflections are therefore

$$x_s^{(\text{TE})} = \frac{\tan \theta_1}{\sqrt{k_x^2 - k_2^2}} \quad (3.5.52a)$$

$$x_s^{(\text{TM})} = \frac{x_s^{(\text{TE})}}{\left(\frac{k_x}{k_2}\right)^2 + \left(\frac{k_x}{k_1}\right)^2 - 1} \quad (3.5.52b)$$

The penetration depth for TE incident light is

$$z_s = \frac{1}{\alpha_{z2}} \quad (3.5.53)$$

and the TM penetration is greater. The penetration depth of (3.5.53) makes sense since a point and slope fit to decaying field in (3.5.48b) gives the null location at α_{z2}^{-1} .

3.6 Birefringent Materials

Most glasses when unstrained are isotropic due to their amorphous nature and lack of long-range periodicity. Crystals, however, can exhibit anisotropy due to natural symmetries of the lattice structure. There are seven distinct crystalline classes, one class being cubic, where the atomic lengths of the three unit cell dimensions are equal; three classes having one unique and two identical unit cell dimensions; and three classes having different unit cell dimensions in all three directions. Cubic crystals, being centrosymmetric, are the only isotropic crystalline class. As the polarizability of a material is intimately related to the binding energy of electrons within the lattice, differences in unit cell dimensions of the remaining six classes cause differences in the material susceptibility. Accordingly, the dielectric constant of the material depends on how the bound electrons oscillate.

In the transparent regime, a high binding energy corresponds to a tight spring constant in the electron-oscillator model, which in turn corresponds to a high resonance frequency. A low binding energy corresponds to a low resonance frequency. At an excitation frequency ω below either resonance frequencies, the electron oscillation along the axis (axes) having lower resonance frequency will induce a higher refractive index (Fig. 3.9). A uniaxial crystal has two different refractive indices and a biaxial crystal has three different refractive indices. Both uniaxial and biaxial materials are called birefringent materials. Table 3.3 summarizes the correspondence of crystal class symmetries with susceptibilities.

When a linearly polarized light ray propagates through a crystal such that its electric field is aligned to a crystalline lattice axis, denoted by index i , the equation of motion for the bound electrons is

$$m \frac{\partial^2 r_i}{\partial t^2} = -K_i r_i - eE_i - m\gamma_i \frac{\partial r_i}{\partial t} \quad (3.6.1)$$

where K_i and γ_i are the spring and damping constants for the i^{th} direction. In general, however, the electric fields are not aligned to a crystalline axis and multiple vibrational modes are excited. Substitution of electric susceptibility tensor χ_e for the scalar χ_e provides the mathematical framework to describe this more complex oscillation. The tensor relation between the induced polarization density and incident electric field is

$$\mathbf{P} = \varepsilon_0 \chi_e(\omega) \mathbf{E} \quad (3.6.2)$$

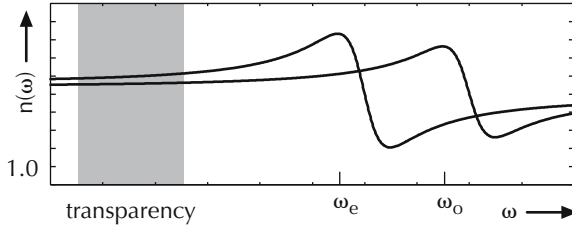


Fig. 3.9. Resonance model of birefringent vibrations and corresponding refractive index.

(cf. (3.5.5)). In the natural coordinate system of the crystal, the tensor χ_e is written as

$$\mathbf{P} = \varepsilon_o \begin{pmatrix} \chi_a & & \\ & \chi_b & \\ & & \chi_c \end{pmatrix} \mathbf{E} \quad (3.6.3)$$

where each diagonal susceptibility may be different and all off-diagonal components are zero. The lack of off-diagonal components in the lattice coordinate system indicates that under pure excitation along a crystalline axis there is no coupling to the other axes.

That the susceptibility is a tensor recasts the dielectric constant ε as a tensor: $\boldsymbol{\varepsilon} = \varepsilon_o(1 + \chi_e)$. The relation of $\mathbf{D}$ to $\mathbf{E}$, while still linear, now depends on the orientation of the electric field. For example, in the kDB system, the electric constitutive relation is

$$\mathbf{D}_k = T \cdot \boldsymbol{\varepsilon} \cdot T^{-1} \mathbf{E}_k \quad (3.6.4)$$

The tensor $T \cdot \boldsymbol{\varepsilon} \cdot T^{-1}$ in general has off-diagonal components, mixing the electron oscillation modes. The following studies analyze the propagation and polarization states of light rays within birefringent media.

3.6.1 Propagation in Uniaxial Materials

Uniaxial birefringence exists in trigonal, tetragonal, and hexagonal crystals. The axis that has the unique lattice constant is called the extraordinary axis; the remaining two axes are called the ordinary axes. A uniaxial crystal can be positive uniaxial or negative uniaxial, depending on the relative refractive indices. The convention is

$$\begin{aligned} \text{positive (+) uniaxial} & \quad n_e > n_o \\ \text{negative (-) uniaxial} & \quad n_e < n_o \end{aligned}$$

Whether a crystal is positive or negative uniaxial depends on the particular constituents and group symmetries.

Table 3.3. Crystal Classes and Birefringence

Class	Unit cell	Susceptibility
<i>Isotropic</i>		
Cubic	$a = c = b$	$\chi_e = \begin{pmatrix} \chi_a & & \\ & \chi_a & \\ & & \chi_a \end{pmatrix}$
<i>Uniaxial</i>		
Trigonal	$a = b \neq c$	$\chi_e = \begin{pmatrix} \chi_o & & \\ & \chi_o & \\ & & \chi_e \end{pmatrix}$
Tetragonal		
Hexagonal		
<i>Biaxial</i>		
Triclinic	$a \neq b \neq c$	$\chi_e = \begin{pmatrix} \chi_a & & \\ & \chi_b & \\ & & \chi_c \end{pmatrix}$
Monoclinic		
Orthorhombic		

After Fowles [5].

In an uniaxial material, tensors ξ^i and ζ^i in C_{DB} are zero, ν is a scalar, and κ is a tensor. The constitutive relations are

$$\mathbf{E} = \kappa \cdot \mathbf{D} \quad (3.6.5a)$$

$$\mathbf{H} = \nu \mathbf{B} \quad (3.6.5b)$$

The impermeittivity tensor κ is diagonal when $\mathbf{D}$ and $\mathbf{E}$ are aligned to its eigenvectors:

$$\kappa = \begin{pmatrix} \kappa_o & & \\ & \kappa_o & \\ & & \kappa_e \end{pmatrix} \quad (3.6.6)$$

where the subscripts o and e refer to ordinary and extraordinary axes, respectively. In diagonal form, $\kappa_i = 1/\varepsilon_i$. Moreover, the extraordinary axis is aligned to the $\hat{z}$ axis of the lattice.

A plane wave propagating within the crystal has a k -vector direction. In the kDB system the k -vector is always aligned to the $\hat{e}_3$ axis, while vectors D_1 and D_2 are aligned to $\hat{e}_1$ and $\hat{e}_2$, respectively. The fields and fluxes, after rotation into kDB coordinates from lattice coordinates, are governed by the coordinate-transformed constitutive relations

$$\mathbf{E}_k = \kappa_k \cdot \mathbf{D}_k \quad (3.6.7a)$$

$$\mathbf{H}_k = \nu \mathbf{B}_k \quad (3.6.7b)$$

where $\kappa_k = T \cdot \kappa \cdot T^{-1}$, or

$$\kappa_k = \begin{pmatrix} \kappa_o & & \\ & \kappa_o \cos^2 \theta + \kappa_e \sin^2 \theta & (\kappa_o - \kappa_e) \sin \theta \cos \theta \\ & (\kappa_o - \kappa_e) \sin \theta \cos \theta & \kappa_o \sin^2 \theta + \kappa_e \cos^2 \theta \end{pmatrix} \quad (3.6.8)$$

The impermeittivity tensor κ_k is independent of ϕ : a uniaxial crystal is isotropic in the plane normal to $\hat{z}$, so the orientation of the k -vector-projection within that plane makes no difference. Substitution of (3.6.7) into the coupled kDB equations (3.3.5) gives

$$\begin{pmatrix} \kappa_{11} & \\ & \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}$$

where

$$\kappa_{11} = \kappa_o$$

$$\kappa_{22} = \kappa_o \cos^2 \theta + \kappa_e \sin^2 \theta$$

Elimination of $\mathbf{B}$ generates the governing equation

$$\begin{pmatrix} u^2 - \nu \kappa_{11} & 0 \\ 0 & u^2 - \nu \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.6.9)$$

In general, there are three non-trivial solutions to (3.6.9). They are

$$\begin{aligned} (1) \quad D_1 \neq 0, D_2 = 0 \quad u &= \pm \sqrt{\nu \kappa_{11}} \\ (2) \quad D_1 = 0, D_2 \neq 0 \quad u &= \pm \sqrt{\nu \kappa_{22}} \\ (3) \quad D_1 \neq 0, D_2 \neq 0 \quad u &= \pm \sqrt{\nu \kappa_{11}} \quad \text{and} \quad u = \pm \sqrt{\nu \kappa_{22}} \end{aligned}$$

For solutions (1) and (2) the k -vector can point in any direction. The dispersion relation for solution (1) is

$$k = \frac{\omega}{c} n_o \quad (3.6.10)$$

Physically, the D_1 flux vector lies along $\hat{e}_1$, which in turn always lies in the (x, y) lattice plane. The D_1 vector therefore never excites the extraordinary vibrational mode; only the ordinary refractive index is experienced. One the other hand, the dispersion relation for solution (2) is

$$k = \frac{\omega}{c} \left[\frac{\cos^2 \theta}{n_o^2} + \frac{\sin^2 \theta}{n_e^2} \right]^{-1/2} \quad (3.6.11)$$

Physically, the D_2 flux vector can point in any direction in the lattice coordinates. Whether D_2 excites the extraordinary vibrational mode, the ordinary

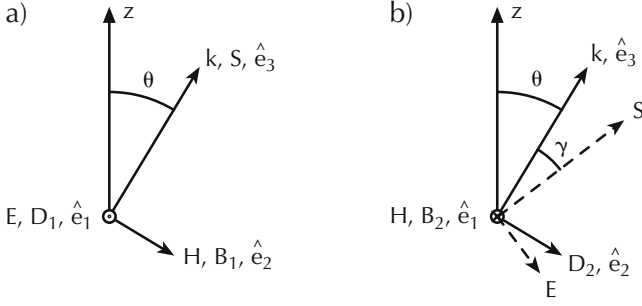


Fig. 3.10. Two solutions to linearly polarized propagation in a (–) uniaxial medium. a) $\mathbf{D} \perp \hat{z}$, therefore $\mathbf{S} \parallel \mathbf{k}$. b) $\mathbf{D} \not\perp \hat{z}$, therefore $\mathbf{S} \not\parallel \mathbf{k}$. For a (+) uniaxial crystal, $\mathbf{S}_e$ lies between $\hat{z}$ and $\mathbf{k}$.

vibrational mode, or a mixture, depends only on θ , the declination angle of the k -vector to the $\hat{z}$ axis. The effective index n_{eff} is defined by identifying terms in (3.6.10, 3.6.11):

$$n_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2 \theta + n_o^2 \sin^2 \theta}} \quad (3.6.12)$$

When the k -vector is aligned with $\hat{z}$, $n_{\text{eff}} = n_o$; and when the k -vector is perpendicular to $\hat{z}$, $n_{\text{eff}} = n_e$. Otherwise, a mixture of vibrational modes is excited and an intermediate refractive index is experienced.

The characteristic direction of the Poynting vectors is another fundamental difference between solutions (1) and (2). As illustrated in Fig 3.10, the k - and Poynting vectors are aligned for solution (1) and are misaligned for solution (2). For solution (1), $D_2 = 0$ and the Poynting vector is

$$\begin{aligned} \mathbf{D}_k &= \hat{e}_1 D_1, \quad \mathbf{E}_k = \hat{e}_1 \kappa_{11} D_1 \\ \mathbf{B}_k &= \hat{e}_2 u/v D_1, \quad \mathbf{H}_k = \hat{e}_2 u D_1 \\ \mathbf{S}_k &= \hat{e}_3 \kappa_{11} u D_1^* D_1 \end{aligned} \quad (3.6.13)$$

In this case $\mathbf{S}_k \parallel \mathbf{k}$. Figure 3.10(a) illustrates that the D_1 vector is perpendicular to $\hat{z}$ and that k - and Poynting vectors are aligned. Solution (2) on the other hand, where $D_1 = 0$, generates a Poynting vector according to

$$\begin{aligned} \mathbf{D}_k &= \hat{e}_2 D_2, \quad \mathbf{E}_k = \hat{e}_2 \kappa_{22} D_2 + \hat{e}_3 \kappa_{32} D_2 \\ \mathbf{B}_k &= -\hat{e}_1 u/v D_2, \quad \mathbf{H}_k = -\hat{e}_1 u D_2 \\ \mathbf{S}_k &= (\hat{e}_3 \kappa_{22} - \hat{e}_2 \kappa_{32}) u D_2^* D_2 \end{aligned} \quad (3.6.14)$$

This is a characteristically different solution. The $\mathbf{E}$ field has a longitudinal component along $\mathbf{k}$ and the Poynting vector is no longer parallel to the k -vector, $\mathbf{S}_k \nparallel \mathbf{k}$. Planes of constant phase lie perpendicular to the k -vector by definition, but those planes are tilted with respect to the energy-flow propagation direction. The walkoff angle γ between the ordinary and extraordinary Poynting vectors is

$$\begin{aligned}\tan \gamma &= -\frac{\kappa_{32}}{\kappa_{22}} \\ &= -\frac{(n_e^2 - n_o^2) \sin \theta \cos \theta}{n_e^2 \cos^2 \theta + n_o^2 \sin^2 \theta}\end{aligned}\quad (3.6.15)$$

The angle γ is a signed value: $\mathbf{S}_e$ is inclined further from the $\hat{z}$ axis than $\mathbf{k}$ in a positive uniaxial material; $\mathbf{S}_e$ is inclined less than $\mathbf{k}$ from $\hat{z}$ in a negative uniaxial material. In either case, $\mathbf{S}_e$ lies in the $(\hat{z}, \hat{k})$ plane. Similar to the reflection and transmission coefficients at a boundary, γ is a function of the n_e/n_o ratio.

There is an important rule for behavior of uniaxial crystals. The rule is

$\mathbf{D} \parallel \hat{z} \text{ or } \mathbf{D} \perp \hat{z} \implies \mathbf{S}_e \parallel \mathbf{k}_e$	When the electric flux vector is either parallel or perpendicular to the $\hat{z}$ axis of the crystal, the extraordinary k - and Poynting vectors coincide.
$\mathbf{D} \nparallel \hat{z} \text{ or } \mathbf{D} \nperp \hat{z} \implies \mathbf{S}_e \nparallel \mathbf{k}_e$	When the electric flux vector is aligned in any other orientation, the extraordinary k - and Poynting vectors are misaligned.

Many designs use the fact that the ordinary and extraordinary rays can be separated by walkoff in a uniaxial crystal.

Returning to (3.6.9), solution (3) allows for the simultaneous existence of D_1 and D_2 , but only when $\mathbf{k}$ is aligned along $\hat{z}$. With $D_1 \neq 0$ and $D_2 \neq 0$, (3.6.9) can only be satisfied for $\theta = 0$. Physically, the k -vector points along $\hat{z}$ and the perpendicular plane contains only the ordinary vibrational mode; only the ordinary refractive index is experienced and the material appears isotropic.

In general, only linearly polarized plane waves can propagate in a uniaxial birefringent crystal. There are two solutions for each orientation of the k -vector, each solution having a distinct wavenumber, effective refractive index, and polarization. An arbitrarily polarized light ray incident on a birefringent crystal is decomposed into two linearly polarized light rays, one associated with each wavenumber. Refraction into and out of a birefringent crystal shows the distinction between these two solutions.

The behavior of light in a birefringent medium, and its refraction into and out of the medium, is described geometrically by the indicatrix. The

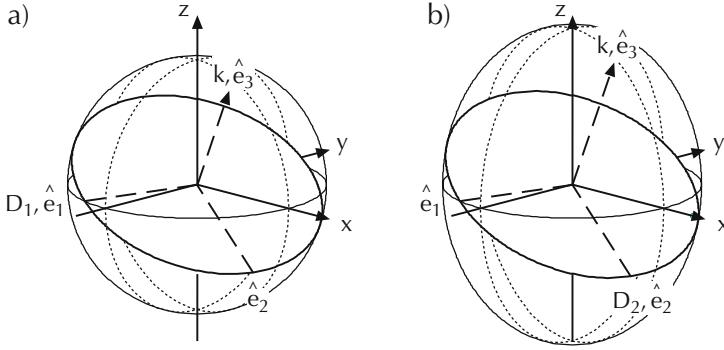


Fig. 3.11. The ordinary and extraordinary indicatrices. a) The ordinary indicatrix is isotropic to k -vector orientation. b) The extraordinary indicatrix is an ellipsoid, with major axis along $\hat{z}$. This is a negative uniaxial indicatrix, the $\hat{z}$ axis being longer than the others.

indicatrix encompasses all possible effective refractive indices of the medium, is associated with the k -vector and polarization vectors, and generates the orientation of the Poynting vector.

For each k -vector there are two indicatrices: one for $D_1 \neq 0$ and the other for $D_2 \neq 0$. In the former case, dispersion relation (3.6.10) is isotropic. Resolving the k -vector into components along the lattice coordinates, the ordinary indicatrix is a sphere (Fig. 3.11(a)), which follows the expression

$$\frac{k_x^2}{n_o^2} + \frac{k_y^2}{n_o^2} + \frac{k_z^2}{n_o^2} = \frac{\omega^2}{c^2} \quad (3.6.16)$$

For the $D_2 \neq 0$ case, the following coordinate associations are made:

$$\begin{aligned} k_z &= k \cos \theta \\ k_y &= k \sin \theta \sin \phi \\ k_x &= k \sin \theta \cos \phi \end{aligned}$$

Substitution into (3.6.11) generates an ellipsoidal indicatrix (Fig. 3.11(b)), which follows the expression

$$\frac{k_x^2}{n_e^2} + \frac{k_y^2}{n_e^2} + \frac{k_z^2}{n_o^2} = \frac{\omega^2}{c^2} \quad (3.6.17)$$

This is the extraordinary indicatrix. The major and minor axes of this indicatrix are aligned to the axes of the lattice, the major axis aligned to $\hat{z}$ axis, and the axis lengths are the wavenumbers $k_{x,y,z}$ along the associated lattice axes. Equivalently, the axis lengths are the refractive indices along the lattice directions.

For either indicatrix the wavenumber is determined by the length of the k -vector at the point of intersection between that vector with the ellipsoid surface. The group velocity is the tangent to the surface at this intersection:

$$v_g = \nabla_k \omega \quad (3.6.18)$$

where $\nabla_k = \hat{k} \partial / \partial k$. Recall that the group velocity is the result of different wavelengths travelling at different speeds. In a birefringent medium a change of k -vector at a fixed frequency is sufficient to alter the propagation speed. A perturbation analysis of Maxwell's equations (3.3.1a,b) shows the geometric interpretation. Expanding these equations to first order in $\delta \mathbf{k}$ and making the indicated dot products gives

$$\begin{aligned} \mathbf{H}^* \cdot (\delta \mathbf{k} \times \mathbf{E} + \mathbf{k} \times \delta \mathbf{E}) &= \omega \delta \mathbf{B} \\ -\mathbf{E} \cdot (\delta \mathbf{k} \times \mathbf{H}^* + \mathbf{k} \times \delta \mathbf{H}^*) &= -\omega \delta \mathbf{D}^* \end{aligned}$$

Taking the difference and applying the vector identity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ yields

$$2\delta \mathbf{k} \cdot (\mathbf{E} \times \mathbf{H}^*) = \omega (\mathbf{H}^* \cdot \delta \mathbf{B} - \delta \mathbf{H}^* \cdot \mathbf{B} + \mathbf{E} \cdot \delta \mathbf{D}^* - \delta \mathbf{E} \cdot \mathbf{D}^*)$$

Substitution of $\mathbf{D}$ and $\mathbf{B}$ with the inverse constitutive relations (3.6.5), recognizing that $\mathbf{E} \cdot \boldsymbol{\epsilon}^* \cdot \mathbf{E}^* = \mathbf{E}^* \cdot \boldsymbol{\epsilon}^\dagger \cdot \mathbf{E}$ (cf. § 3.2), and applying lossless conditions $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^\dagger$ and $\mu = \mu^\dagger$, reduces the expansion to

$$\begin{aligned} 2\delta \mathbf{k} \cdot (\mathbf{E} \times \mathbf{H}^*) &= \omega ((\mathbf{E}^* \cdot \boldsymbol{\epsilon} \cdot \delta \mathbf{E})^* - (\mathbf{E}^* \cdot \boldsymbol{\epsilon} \cdot \delta \mathbf{E}) + \\ &\quad (\mathbf{H}^* \cdot \mu \delta \mathbf{H}) - (\mathbf{H}^* \cdot \mu \delta \mathbf{H})^*) \end{aligned}$$

The right-hand side of (3.6.19) is an entirely imaginary quantity. As the time-averaged Poynting vector is defined as $\langle S \rangle = \Re(\mathbf{E} \times \mathbf{H}^*)/2$, this last expression is further reduced to

$$\delta \mathbf{k} \cdot \langle \mathbf{S} \rangle = 0 \quad (3.6.19)$$

Geometrically, $\langle \mathbf{S} \rangle$ is normal to the indicatrix surface at the intersection point k . As this coincides with the direction of the group velocity, is it clear that the energy flow travels at the group velocity and may not coincide with the direction of the k -vector.

3.6.2 Refraction at an Interface

Refraction and reflection at an interface between isotropic and uniaxial materials are treated nearly the same way as detailed in §3.5, but with the following modifications:

- 1) the o -ray and e -ray are calculated separately;
- 2) the effective index of the e -ray determines the refraction angle and reflection coefficient; and

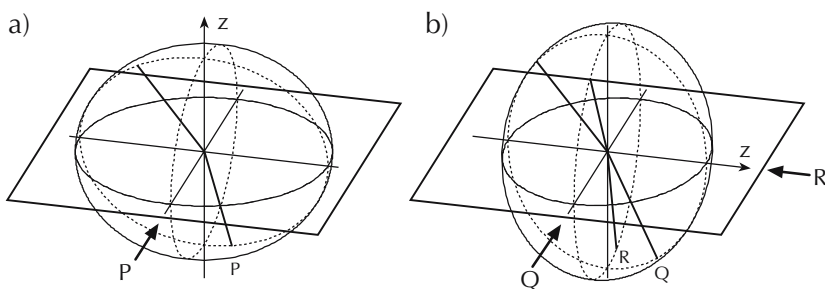


Fig. 3.12. Two orientations of the extraordinary indicatrix at a boundary. a) The $\hat{z}$ axis is normal to the interface, view P is isotropic about $\hat{z}$. b) The $\hat{z}$ axis is in the plane of the interface. Rays Q and R are orthogonal slices through the indicatrix and have different refraction angles.

3) the Poynting vector of the e -ray is not generally coincident with the k -vector and must be calculated separately.

Snell's law remains intact for birefringent materials because it enforces phase matching along the interface.

Work with birefringent crystals requires the distinction between the physical shape of the crystal and the internal orientation of the lattice. Crystal cuts can be limited by the brittleness of the material, but there are nonetheless several customary orientations that can be cut. As illustrated in Fig. 3.12, two orientations that are common are for the $\hat{z}$ normal to or lying in the interface plane. The figures illustrate the extraordinary indicatrix of a positive uniaxial crystal intersected by a plane, and rays P , Q , and R refracting into the interface.

Figure 3.12(a) illustrates the $\hat{z}$ axis cut perpendicular to the interface plane. The vertical plane defined by view P is isotropic about $\hat{z}$; the effective index of e -ray P depends only on its inclination. Figure 3.12(b) illustrates the $\hat{z}$ axis cut within the interface plane. The vertical plane defined by view Q forms an elliptical intersection with the indicatrix while the plane defined by view R forms a circle. The effective index of an e -ray refracting into a uniaxial material with this cut depends both on the inclination angle and azimuth orientation about the interface normal.

Figures 3.13(a–f) show the refractive index manifolds for views P , Q , and R for positive and negative uniaxial crystals. In each figure, the ordinary and extraordinary indicatrices are shown in plan view and bisected by the interface plane. The center of each indicatrix must lie in the plane of the interface and be located where the respective k -vector breaches the interface. As illustrated, the centers of the e - and o -indicatrices coincide.

The drawings illustrate the relative refraction for the e - and o -rays. The horizontal components of the k_e - and k_o -vectors are equal, satisfying phase matching along the interface. The e - and o -Poynting vectors are normal to

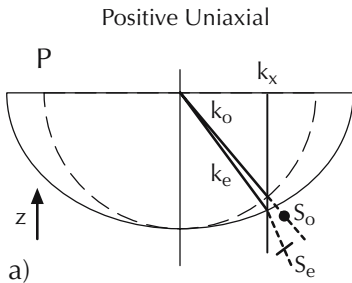


Fig. 3.13a. Refraction manifold for orientation P in a positive uniaxial crystal. Ordinary (dashed) and extraordinary (solid) indicatrices are bisected by the interface plane. Poynting vectors $\mathbf{S}$ are normal to the respective indicatrices. The o -ray is TE.

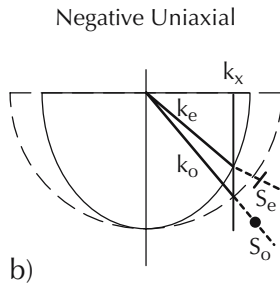


Fig. 3.13b. Refraction manifold for orientation P in a negative uniaxial crystal. Ordinary (dashed) and extraordinary (solid) indicatrices are bisected by the interface plane. Poynting vectors $\mathbf{S}$ are normal to the respective indicatrices. The o -ray is TE.

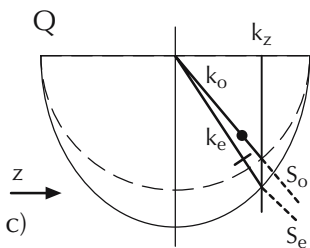


Fig. 3.13c. Refraction manifold for orientation Q . The o -ray is TE.

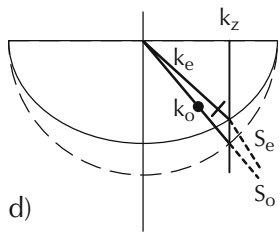


Fig. 3.13d. Refraction manifold for orientation Q . The o -ray is TE.

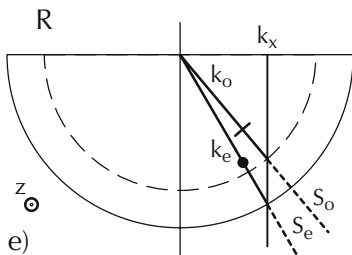


Fig. 3.13e. Refraction manifold for orientation R . The o -ray is TM. This is the only orientation that is isotropic for both e -rays and o -rays.

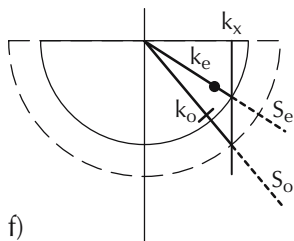


Fig. 3.13f. Refraction manifold for orientation R . The o -ray is TM. This is the only orientation that is isotropic for both e -rays and o -rays.

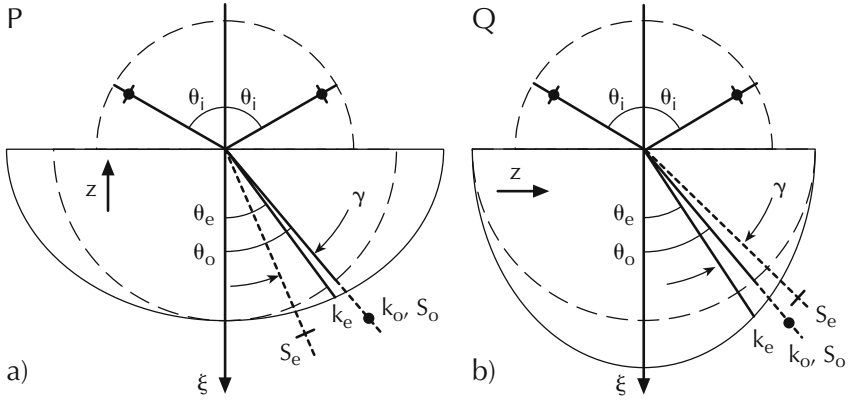


Fig. 3.14. Refraction from an isotropic material into a positive uniaxial material, views P and Q. a) Poynting vector S_e lies between S_o and $\hat{z}$. The e -ray is slow. b) Poynting vector S_e lies outside of S_o . The e -ray is again slow. In both cases the o -ray is TE.

the corresponding indicatrices at the point of intersection with the respective k -vectors. The polarization orientation is also illustrated for each ray. The ordinary ray always has its polarization state perpendicular to the extraordinary axis $\hat{z}$. Depending on the direction of $\hat{z}$ at the interface, the ordinary ray may be either TE or TM. Finally, the “fast” axis is always the axis that has the lower refractive index; like TE and TM, the fast axis can be either the e - or o -ray, depending on orientation of $\hat{z}$ to the interface.

Refraction into positive uniaxial crystals for views P and Q is shown in detail in Figs. 3.14(a,b). For view P (Fig. 3.14(a)), the o -ray sees a refractive index of n_o and is TE. Snell’s law determines the angle of refraction:

$$n_i \sin \theta_i = n_o \sin \theta_o \quad (3.6.20)$$

The Poynting vector S_o coincides with k_o . The e -ray, which is TM, sees an effective refractive index n_{eff} given by

$$n_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2 \theta_e + n_o^2 \sin^2 \theta_e}} \quad (3.6.21)$$

where θ_e is the angle between k_e and the $\hat{z}$ axis. Snell’s law then

$$n_i \sin \theta_i = n_{\text{eff}} \sin \theta_e \quad (3.6.22)$$

However, this version of Snell’s law contains the angle θ_e both in n_{eff} and the sine term. Given n_e , n_o , and θ_i , (3.6.22) can be solved for θ_e :

$$\tan \theta_e = \frac{n_e}{n_o} \frac{n_i \sin \theta_i}{\sqrt{n_e^2 - n_i^2 \sin^2 \theta_i^2}} \quad (3.6.23)$$

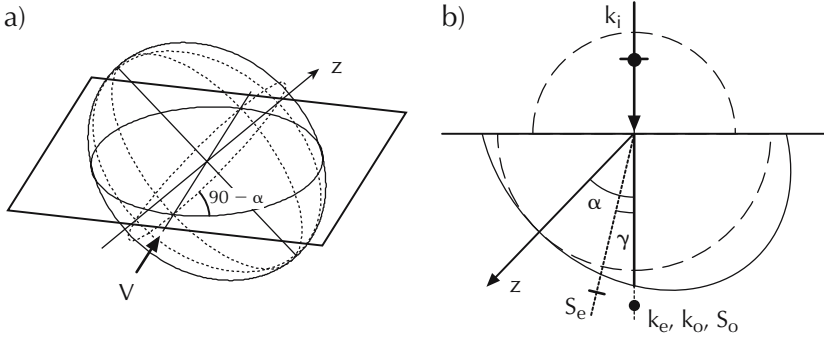


Fig. 3.15. A walkoff cut: a uniaxial crystal cut where the $\hat{z}$ axis is inclined by α from the normal. a) The (positive) extraordinary indicatrix as intersected by the interface plane. View V is indicated. b) Plan view of refraction at interface. Ordinary (dashed) and extraordinary (solid) indicatrices are shown. $\mathbf{S}_e$ is inclined by γ from the normal.

One can verify that when $n_e \rightarrow n_o$, (3.6.23) reduces to (3.6.20). The normal to the extraordinary indicatrix at the point of intersection with $\mathbf{k}_e$ determines the direction of the $\mathbf{S}_e$ vector. The angle γ between the $\mathbf{S}_o$ and $\mathbf{S}_e$ is calculated from (3.6.15) using $\theta = \theta_e$. For view P with a positive uniaxial crystal, γ is negative and $\mathbf{S}_e$ lies between $\mathbf{S}_o$ and $\hat{z}$.

View Q is for a cut such that $\hat{z}$ lies in the interface plane; this is a waveplate cut. When $\theta_i = 0$, the ordinary and extraordinary k - and Poynting vectors coincide but the refractive index of the e -ray differs from that of the o -ray. As the two rays transit the crystal, the e - and o -rays slip relative to one another, which in turn transforms the polarization state from input to output.

Figure 3.15 illustrates a third important crystal cut, the walkoff cut. A walkoff crystal is used to polarize the input light and spatially separate the resultant ordinary and extraordinary components. The crystal is cut such that its $\hat{z}$ -axis is inclined from the plane of the interface. Since no walkoff occurs at 0° and 90° inclination, there is an intermediate inclination angle that maximizes the angular separation of $\mathbf{S}_e$ and $\mathbf{S}_o$.

Figure 3.15(b) illustrates the refraction manifold for an inclined-cut positive-uniaxial crystal for view V (Fig. 3.15(a)). The $\hat{z}$ -axis is inclined by α with respect to the normal. When the incident light is normal to the surface, the ordinary component is not refracted and continues in the same direction. The ordinary Poynting vector runs parallel to the k_o vector. The extraordinary component, however, behaves differently. Since the input angle (as illustrated) is normal, the extraordinary k -vector also continues into the crystal without refraction and runs parallel to the k_o vector. However, the point of intersection of k_e with the extraordinary indicatrix leads to a deflected Poynting vector. The deflection for a positive-uniaxial crystal is in the direction of the $\hat{z}$ axis.

The deflection angle γ , also known as the walkoff angle, is governed by (3.6.15) with α in place of θ (note that n_e is not replaced by n_{eff} : the

deflection equation is written in terms of the cardinal indices). In order to maximize the walkoff angle γ , (3.6.15) is differentiated with respect to α and set to zero. The cut angle $\tilde{\alpha}$ that maximizes the walkoff angle is

$$\tan(\tilde{\alpha}) = \frac{n_e}{n_o} \quad (3.6.24)$$

and the corresponding deflection angle is

$$\tan(\gamma_{\max}) = -\frac{1}{2} \left(\frac{n_e}{n_o} - \frac{n_o}{n_e} \right) \quad (3.6.25)$$

The governing quantity is the ratio n_e/n_o , rather than the difference. A positive-uniaxial crystal having 10% birefringence, that is, $n_e/n_o = 1.1$, gives a maximum walkoff angle $\gamma_{\max} \simeq 5.45^\circ$ for a cut angle of $\tilde{\alpha} \simeq 47.7^\circ$.

A walkoff crystal makes two parallel but laterally displaced, orthogonally polarized output rays when the input ray is perpendicular to the input face, and the input and output faces are cut parallel to one another. The index of the ordinary ray is n_o , while that of the extraordinary ray is (compare (3.6.21))

$$n_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2(\theta_e + \alpha) + n_o^2 \sin^2(\theta_e + \alpha)}} \quad (3.6.26)$$

For a crystal cut at $\tilde{\alpha}$ the effective index is

$$n_{\text{eff}} = \sqrt{\frac{1}{2} (n_e^2 + n_o^2)} \quad (3.6.27)$$

This is, in fact, just the average of the permittivities along the cardinal directions.

The effective path length L_{eff} of the extraordinary ray is the length traversed by the k_e vector times the effective index. That is, the path length of $\mathbf{S}_e$ does not enter into the effective path. The path-length difference for a walkoff crystal of length L is $L_{\text{eff}} - L_o = (n_{\text{eff}} - n_o) L$.

3.6.3 Total Internal Reflection

Total internal reflection (TIR) for birefringent materials is more interesting than for isotropic materials because its onset differs for e - and o -rays. Otherwise, the physics of TIR remains the same as detailed in §3.5.6. This section analyzes a birefringent polarizer that uses selective TIR, and asymmetric TIR where the input and output angles are unequal.

Figure 3.16 illustrates a polarizing birefringent wedge. The $\hat{z}$ axis of the wedge is cut normal to the plane of the page and the output face is tilted by angle θ_p from the input face. The incident light is normal to the input face. Since the polarization of the o - and e -rays is perpendicular to the $\hat{z}$ axis, the Poynting vectors coincide with their respective k -vectors. No refraction

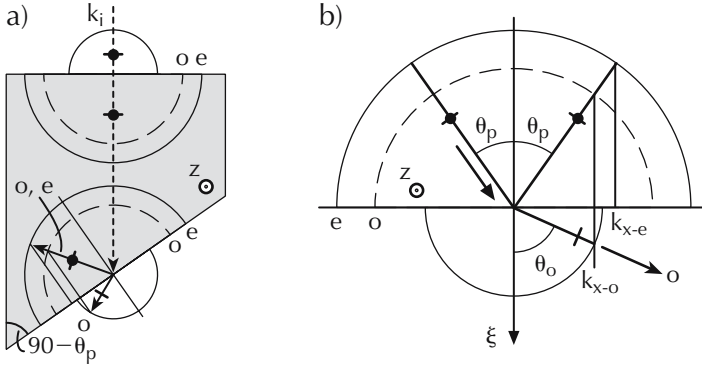


Fig. 3.16. Polarizing wedge. a) Cross-section of a polarizing birefringent wedge cut with $\hat{z}$ pointing out of the page. Only the o -ray survives refraction. b) Plan view of output refraction manifold.

is experienced at the input. However, the inclination of the output face cuts off transmission for the e -ray but not the o -ray. This is because the refractive indices of the two rays differ. The critical angles for the two polarizations are

$$\theta_{c,e} = \sin^{-1}(n_i/n_e)$$

for the e -ray and

$$\theta_{c,o} = \sin^{-1}(n_i/n_o)$$

for the o -ray, where n_i is the isotropic index, possibly air. As long as θ_p falls within $\theta_{c,e} \leq \theta_p \leq \theta_{c,o}$ then only the o -ray emerges from the output face. The e -ray experiences TIR. The angle at which the o -ray emerges with respect to the axis of the input ray is

$$\theta_o = \sin^{-1}\left(\frac{n_o}{n_i} \sin \theta_p\right)$$

The o -ray is TM at the second interface and will suffer reflection, reducing its transmission. The reflection coefficient is determined by (3.5.40). The Shirasaki prism solves this problem by setting θ_p to Brewster's angle (see §4.7.3).

Another example of the effects of birefringent total internal reflection is the asymmetric reflection created in a walkoff-cut crystal, Fig. (3.17). The left side of the figure illustrates the configuration, where the $\hat{z}$ axis is cut at angle α and the incident light is normal to the input. The k -vector does not refract into the crystal, but the extraordinary Poynting vector tilts upward by angle γ . The Poynting vector propagates until it hits the roof of the crystal, whereupon it experiences TIR as long as the outside refractive index is suitably low. After total internal reflection the k - and Poynting vectors point downward. The k - and Poynting vectors refract at the output surface and subsequently run parallel to one another.

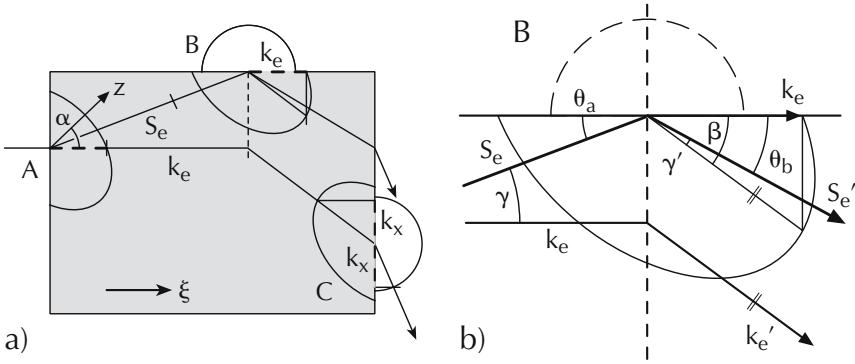


Fig. 3.17. TIR from roof of walkoff block. a) Path of e -ray k - and Poynting vectors. Position A: normal incidence, walkoff of extraordinary ray. Position B: TIR from roof, asymmetric reflection and direction change of k -vector. Position C: refraction at output. b) Detail of Position B.

Figure (3.17b) illustrates the detail of the TIR at the roof of the walkoff crystal. The extraordinary-index refraction manifold is shown tilted by angle α with respect to the roof. The incident vector $\mathbf{k}_e$ is parallel to the roof. The reflected wave must maintain phase matching along the roof, and, unlike an isotropic material, the ellipsoidal shape of the extraordinary indicatrix allows for two solutions: $\mathbf{k}_e$ that runs parallel with the roof, and $\mathbf{k}'_e$ that tilts downward at angle β . To calculate the angle β , the effective index n'_{eff} associated with $\mathbf{k}'_e$ (when projected along the horizontal axis) must match the effective index n_{eff} associated with $\mathbf{k}_e$. Thus,

$$n'_{\text{eff}} \cos \beta = n_{\text{eff}}$$

where

$$n'_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2(\alpha + \beta) + n_o^2 \sin^2(\alpha + \beta)}}$$

Solving for $\tan \beta$ gives

$$\tan \beta = \frac{2(n_e^2 - n_o^2) \sin \alpha \cos \alpha}{n_e^2 \sin^2 \alpha + n_o^2 \cos^2 \alpha} \quad (3.6.28)$$

The vector $\mathbf{k}'_e$ is tilted downward by angle β . The deviation angle γ' between Poynting vector $\mathbf{S}'_e$ and $\mathbf{k}'_e$ is calculated from (3.6.15), replacing θ with $\alpha + \beta$. The reflection angle θ_b with respect to the roof is $\theta_b = \beta - \gamma'$. Generally speaking, $\theta_a \neq \theta_b$, although the values can be close.

When the tilt angle is optimized for maximum walkoff, $\tan \tilde{\alpha} = n_e/n_o$. Substituting this angle into (3.6.28) gives

$$\tan \beta(\tilde{\alpha}) = \left(\frac{n_e}{n_o} - \frac{n_o}{n_e} \right) \quad (3.6.29)$$

For $n_e/n_o = 1.1$, $\beta \simeq 10.8^\circ$. Another relevant identity is

$$\tan(\tilde{\alpha} + \beta(\tilde{\alpha})) = \left(\frac{n_e}{n_o}\right)^3 \quad (3.6.30)$$

Combined, the inclination of $\mathbf{S}'_e$ to $\mathbf{k}'_e$ is given by

$$\tan \gamma'(\tilde{\alpha}) = \frac{1 - (n_e/n_o)^2}{1 + (n_e/n_o)^4} \frac{n_e}{n_o} \quad (3.6.31)$$

Substitution yields $\gamma' \simeq 5.36^\circ$. Therefore, the reflected angle is $\theta_b \simeq 5.44^\circ$. This is compared with $\theta_a \simeq 5.45^\circ$. Here the reflected and incident TIR angles are nearly the same. This is because $\tilde{\alpha}$ maximizes γ ; at the maximum the system is stationary. Asymmetry occurs for lower walkoff angles.

3.6.4 Polarization Transformation

A waveplate is a birefringent crystal cut so that its extraordinary axis lies in the plane of the input, and the input and output crystal faces are parallel. A light ray normally incident on the input face is resolved into ordinary and extraordinary components, according to the relative orientation of the input polarization and the e -axis of the waveplate. The k - and Poynting vectors within the crystal propagate collinearly. The wavenumbers for the two components are

$$k_e = \frac{\omega}{c} n_e, \quad k_o = \frac{\omega}{c} n_o \quad (3.6.32)$$

Therefore the phase velocities differ. The difference in phase velocity results in the fast component walking through the slow component, slipping one full optical wave every birefringent beat length.

Strictly speaking, the ordinary and extraordinary waves within the waveplate propagate independently from one another, and it is accurate to say only linearly polarized fields exist. Once the fields emerge from the waveplate they continue to run collinearly but have a phase slip between them. This phase slip makes the output polarization (generally) different from the input polarization. Within a waveplate, the phase slip at any position z_o is represented by a polarization rotation *as if* the waveplate ended at z_o .

Figure 3.18 illustrates the polarization evolution along a long waveplate in laboratory coordinates and Stokes coordinates. In the laboratory frame, an input polarization state $|s\rangle$ is resolved into ordinary and extraordinary components at the input face of the waveplate. These two components have different wavelengths and travel at different phase velocities. The phase slip φ between the components is called the birefringent phase and as a function of travel distance z is defined by

$$\varphi(\omega, z) = \frac{\omega}{c} \Delta n z \quad (3.6.33)$$

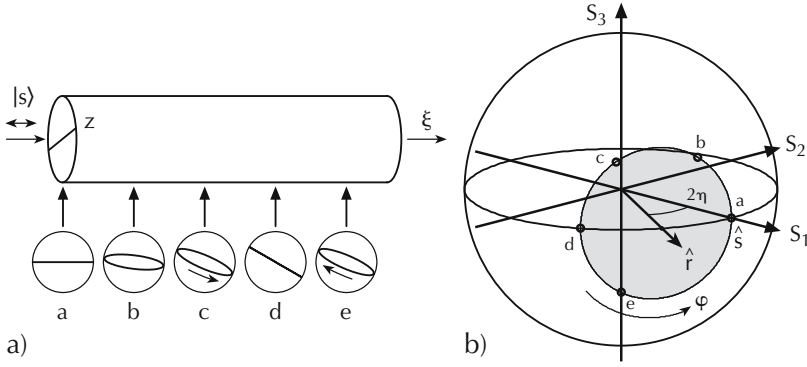


Fig. 3.18. Polarization transformation along a waveplate. a) The input SOP $|s\rangle$ is projected onto $\hat{z}$, which is tilted by angle η within the polarization plane. The projection generates two collinear waves having different wavelengths. The phase slip along the waveplate transforms the state of polarization from state a through e . b) Stokes picture of polarization change. Precession of $\hat{s}$ about $\hat{r}$.

where Δn is the birefringence $\Delta n = n_e - n_o$. The travel distance over which the components slip by one full wave is called the birefringent beat length. This distance is defined by

$$\Lambda = \frac{\lambda_0}{\Delta n} \quad (3.6.34)$$

where λ_0 is the free-space wavelength. The positive uniaxial crystal YVO_4 has a birefringence of $\Delta n \simeq 0.2$; the corresponding beat length is only $7.5 \mu\text{m}$ at $\lambda_0 \simeq 1.5 \mu\text{m}$.

The phase slip through the waveplate changes the polarization state. Figure 3.18(b) illustrates the corresponding Stokes transformation. As only the phase changes but not the power split between e - and o -components, the precession circle is traced. The birefringent axis of the waveplate lies on the equator of the Poincaré sphere. The polarization evolution is periodic with travel distance.

The birefringent phase φ also depends on frequency. The free-spectral range is the analogue to the birefringent beat length but for frequency change, not length change. However, there is an important subtlety: the ordinary and extraordinary refractive indices are generally frequency dependent. To first order the refractive index n is replaced by the group index n_g . Accounting for the group index, the free-spectral range (FSR) of a waveplate of length L is defined by $\omega_2 - \omega_1 = 2\pi \times \text{FSR}$, or

$$\text{FSR} = \frac{c}{\Delta n_g L} \quad (3.6.35)$$

The free-spectral range is given in cyclic frequency. Subsection entitled, “Free-Spectral Range and Group Index” starting on page 158 details the relation between phase and group indices.

The accrual of phase slip due to frequency change is tantamount to a relative delay between the two components. This delay is called the differential group delay (DGD). As delay is defined as $\tau = \partial\varphi/\partial\omega$, the DGD for a waveplate is

$$\tau = \frac{\Delta n_g L}{c} = \frac{1}{\text{FSR}} \quad (3.6.36)$$

With this definition of differential group delay, the birefringent phase as a function of frequency is written $\varphi = \omega\tau$. An output polarization state precesses about the birefringent axis (of a single homogeneous section) in Stokes space just the same as if the length changed, but in the case of length change the precession angle is $\varphi = 2\pi z/\Lambda$.

3.7 Gyrotropic Materials

Gyrotropic materials are those whose optical properties are effected by either an applied magnetic field or their own internal magnetization. In either case the eigen-axes are elliptical or circular rather than linear. Moreover, unique to gyrotropic materials is their nonreciprocal polarization rotation. This non-reciprocal rotation is due to a reorientation of the eigen-axes between forward and backward transit. The term “non-reciprocal” is of unfortunate historical origin because one typically thinks of a non-linear system being non-reciprocal. Non-reciprocity of a non-linear system means one cannot reverse time and end up with what was at the beginning. Gyrotropic processes, instead, are fully reciprocal under time reversal. Time reversal not only flips the propagation direction but also reverses the electron spin which is the source of magnetization. Non-reciprocity for gyrotropic media refers to polarization transformation for forward and backward propagation.

There are a vast array of materials and conditions that exhibit gyrotropic behavior. Gyrotropic effects are known in gases, liquids, and solids. Gyrotropic effects are also known to exist from radio-frequency through the ultraviolet. A full exposition of gyrotropy requires a quantum mechanical formalism, which is not the thrust of this text. Nonetheless, there are two steps involved when studying these materials. First, a microscopic-level analysis determines the origin and character of the gyrotropy, culminating in a constitutive relation. Second, given the constitutive relation, a *kDB* analysis predicts the interaction of light with the material. Common to most all gyrotropic materials is the constitutive relation of the permittivity tensor:

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon & -j\varepsilon_g & \\ j\varepsilon_g & \varepsilon & \\ & & \varepsilon_z \end{pmatrix} \quad (3.7.1)$$

The diagonal entries look like a uniaxial birefringent material. The off-diagonal entries are characteristic of gyrotropy. These entries are purely imaginary and keep the tensor $\boldsymbol{\varepsilon}$ Hermitian. For low, non-optical frequencies, the permeability tensor $\boldsymbol{\mu}$ takes the form of (3.7.1) and $\boldsymbol{\varepsilon}$ is a scalar.

3.7.1 Magnetic Material Classes

There are five classes of magnetic materials that exhibit magneto-optical effects: diamagnetic, paramagnetic, ferromagnetic, antiferromagnetic, and ferrimagnetic. Diamagnetic and paramagnetic materials have no intrinsic magnetization; these materials exhibit weak magneto-optical effects. Ferro-, antiferro- and ferrimagnetic materials do have intrinsic magnetization and their magneto-optical effects are dominated by their internal magnetization rather than an external field (unless of high coercivity). These materials can exhibit strong magneto-optical effects.

A diamagnetic material has no magnetic dipole moments at a microscopic level. Only through coercion from an external field does this material class exhibit gyrotropic behavior. The external field splits the excited-state electron levels of the constituent atoms, thereby inducing a refractive index splitting. A paramagnetic material does possess intrinsic magnetic dipole moments among some if not all of the constituent atoms. These magnetic moments exist in the ground state and tend toward cancellation on a nano-scale. An externally applied magnetic field splits the ground state, rather than the excited state, of the atoms, inducing a refractive index splitting. A finite temperature populates the upper level of the split ground state, again leading to general cancellation of the magnetization. Magneto-optical effects can be stronger in paramagnetic materials than in diamagnetic materials, but at the cost of a fall higher temperature sensitivity.

Ferro-, antiferro-, and ferrimagnetic materials are more complex. These materials all exhibit long-range spin ordering. A homogeneously ordered region is called a domain. Domains exist even in the absence of an applied magnetic field and the magnetization within a domain can be orders of magnitude stronger than for dia- or paramagnetic materials. In ferromagnetic materials, the spins of unpaired valence electrons align, generating a large magnetization within a domain. However, on a microscopic scale neighboring domains are coerced to align anti-parallel and buck a large spontaneous magnetization. An external field, however, coerces the domains to align preferentially in the direction of the applied field, which in turn induces a strong magnetic reaction to the field. In antiferromagnetic materials the spins of adjacent unpaired valence electrons anti-align and cancel. While the process of spin-orbit coupling places antiferromagnetic materials in the same category as ferromagnetic materials, magnetization of this type is weak. Finally, ferrimagnetic materials exhibit a coupling of two interstitial magnetic sublattices that couple antiferromagnetically. Generally, one sublattice has a stronger magnetization than the other, offering some degree of overall magnetization.

Ferro-, antiferro-, and ferrimagnetic materials have such strong intrinsic magnetization that their magneto-optical properties are dominated by the internal fields rather than applied fields (of a coercion order one would expect in a telecommunications environment). Often, such complex material structures require measurements of their magneto-optical properties since a

first-principles derivation is too complex. Moreover, these materials can have significant temperature dependence since thermal vibration disrupts the zero-Kelvin alignment. In fact, at zero Kelvin, the magnetization of a ferromagnet is largest, the magnetization decreases monotonically with increased temperature, and finally vanishes at the Curie temperature T_c .

In this text, a short derivation using classical electron motion is presented to motivate the construction of the constitutive relation with linear entries. Indeed, the Lorentz force is the only tool presented in this text to address the present analysis. Only diamagnetic materials yield to the Lorentz-force analysis. The reader is cautioned about adopting this derivation to other materials. Rather, references [3, 6] and the references found within provide a solid foundation on which to pursue more detailed studies of the microscopic behavior.

3.7.2 Permittivity of Diamagnetic Materials

Magnetization of diamagnetic materials is generated by an externally applied magnetic field. As discussed in §3.4, the magnetic field of an electro-magnetic wave is not strong enough to generate an appreciable force on the bound electron since $|\omega \mathbf{r} \mathbf{B}| \ll |e \mathbf{E}|$. However, an external biasing field can easily influence the motion. Consider once again the bound-electron equation of motion (3.4.2) rewritten in time-harmonic form and with an external DC magnetic field aligned along $+\hat{z}$:

$$(-\omega^2 + \omega_o^2) \mathbf{r} + j\omega(e/m) (\mathbf{r} \times B_z) = -(e/m)\mathbf{E} \quad (3.7.2)$$

The spring-constant resonance is $\omega_o = K/m$. Before the susceptibilities are calculated, intuition can be developed by solving (3.7.2) under simplifying assumptions. Consider a ray propagating along the $+\hat{z}$ direction; in this case $\mathbf{r} \cdot \hat{z} = 0$. Separating (3.7.2) into its components and solving for $\mathbf{r}$ gives

$$\begin{pmatrix} r_x \\ r_y \end{pmatrix} = \frac{a_2}{1 - a_1^2} \begin{pmatrix} 1 & -ja_1 \\ ja_1 & 1 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}$$

where $a_1 = \omega(e/m)B_z/(\omega_o^2 - \omega^2)$ and $a_2 = -(e/m)/(\omega_o^2 - \omega^2)$. The eigenvalue and eigenvector solutions are

$$\lambda_+ = 1 + a_1, \quad \lambda_- = 1 - a_1$$

$$\mathbf{E}_+ = \begin{pmatrix} 1 \\ j \end{pmatrix}, \quad \mathbf{E}_- = \begin{pmatrix} 1 \\ -j \end{pmatrix}$$

Counter-clockwise and clockwise polarization states of $\mathbf{E}$ are the eigenstates of the system. These two states in turn induce a circular motion of the bound electrons about the B_z field, the counter-clockwise orbit having the larger radius of the two. The circular electron trajectory is called cyclotron motion.

As the susceptibility is intimately related to the electron motion, one should expect that the eigen-polarizations of the system, when simplified in this manner, are circular and that the refractive indices are different for ccw and cw polarizations, the larger index corresponding to the larger electron radius.

Returning to the detailed analysis of susceptibilities based on the Lorentz force, recall that the polarization density is defined as $\mathbf{P} = -N\mathbf{e}\mathbf{r}$. The equation of motion (3.7.2) is recast in terms of $\mathbf{P}$ and $\mathbf{E}$:

$$(\omega_o^2 - \omega^2) \mathbf{P} + j\omega(e/m) (\mathbf{P} \times B_z) = Ne^2/m\mathbf{E} \quad (3.7.3)$$

Expanding the curl, the cartesian components in matrix form are

$$\begin{pmatrix} (\omega_o^2 - \omega^2) & +j\omega\omega_c & 0 \\ -j\omega\omega_c & (\omega_o^2 - \omega^2) & 0 \\ 0 & 0 & (\omega_o^2 - \omega^2) \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \frac{Ne^2}{m} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}$$

where the cyclotron resonance frequency ω_c is defined as

$$\omega_c = \frac{eB_z}{m} \quad (3.7.4)$$

The dielectric susceptibility tensor χ_e relates the components of the electric field to the polarization density: $\mathbf{P} = \varepsilon_o\chi_e \mathbf{E}$. Inverting (3.7.4) gives the components of the electric susceptibility tensor

$$\chi_e = \begin{pmatrix} \chi_{11} & -j\chi_{12} & \\ j\chi_{12} & \chi_{11} & \\ & & \chi_{33} \end{pmatrix} \quad (3.7.5)$$

where

$$\chi_{11} = \frac{Ne^2}{m\varepsilon_o} \frac{(\omega_o^2 - \omega^2)}{(\omega_o^2 - \omega^2)^2 - \omega^2\omega_c^2} \quad (3.7.6a)$$

$$\chi_{12} = \frac{Ne^2}{m\varepsilon_o} \frac{\omega\omega_c}{(\omega_o^2 - \omega^2)^2 - \omega^2\omega_c^2} \quad (3.7.6b)$$

$$\chi_{33} = \frac{Ne^2}{m\varepsilon_o} \frac{1}{(\omega_o^2 - \omega^2)} \quad (3.7.6c)$$

The susceptibility χ_{12} is directly related to the strength and direction of the fixed magnetic field B_z . The susceptibility χ_{11} experiences a second-order detuning due to the field, but is generally comparable in magnitude to χ_{33} when far away from absorption resonance. The permittivity tensor ε relates to the susceptibility in the usual way: $\varepsilon = \varepsilon_o(1 + \chi_e)$. The entries in ε correspond to those in χ_e by

$$\varepsilon = \varepsilon_o(1 + \chi_{11}) \quad (3.7.7a)$$

$$\varepsilon_z = \varepsilon_o(1 + \chi_{33}) \quad (3.7.7b)$$

$$\varepsilon_g = \varepsilon_o\chi_{12} \quad (3.7.7c)$$

where the subscripts z and g denote the dielectric constant along the $\hat{z}$ axis and the gyrotropic dielectric value, respectively. The impermeability tensor κ required for the kDB calculations is

$$\begin{pmatrix} \varepsilon & -j\varepsilon_g \\ j\varepsilon_g & \varepsilon \\ & & \varepsilon_z \end{pmatrix}^{-1} = \begin{pmatrix} \kappa & j\kappa_g \\ -j\kappa_g & \kappa \\ & & \kappa_z \end{pmatrix} \quad (3.7.8)$$

where

$$\kappa = \frac{\varepsilon}{\varepsilon^2 - \varepsilon_g^2}, \quad \kappa_g = \frac{\varepsilon_g}{\varepsilon^2 - \varepsilon_g^2}, \quad \kappa_z = \frac{1}{\varepsilon_z} \quad (3.7.9)$$

Note from (3.7.8) that indeed $\varepsilon = \varepsilon^\dagger$, so the system remains ideally lossless: no work is done on the electrons due to the fixed magnetic field, all energy coupled into the cyclotron motion is recovered.

3.7.3 Propagation in Gyrotropic Materials

The constitutive relations for gyrotropic materials are

$$\mathbf{E} = \kappa \cdot \mathbf{D}$$

$$\mathbf{H} = \mu \mathbf{B}$$

where κ is given by (3.7.8). Rotation into the kDB frame gives

$$\mathbf{E}_k = \kappa_k \cdot \mathbf{D}_k \quad (3.7.10a)$$

$$\mathbf{H}_k = \mu \mathbf{B}_k \quad (3.7.10b)$$

where the transformed impermeability tensor relation is

$$\kappa_k = \begin{pmatrix} \kappa & j\kappa_g \cos \theta & j\kappa_g \sin \theta \\ -j\kappa_g \cos \theta & \kappa \cos^2 \theta + \kappa_z \sin^2 \theta & (\kappa - \kappa_z) \sin \theta \cos \theta \\ -j\kappa_g \sin \theta & (\kappa - \kappa_z) \sin \theta \cos \theta & \kappa \sin^2 \theta + \kappa_z \cos^2 \theta \end{pmatrix} \quad (3.7.11)$$

There is no ϕ dependence in κ . The eigenvectors of (3.7.8) are circular in the plane perpendicular to $\hat{z}$, so κ is invariant to any rotation ϕ . Substitution of (3.7.11) into the coupled kDB equations (3.3.5) gives

$$\begin{pmatrix} \kappa_{11} & j\kappa_{12} \\ -j\kappa_{12} & \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \quad (3.7.12a)$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \quad (3.7.12b)$$

where

$$\kappa_{11} = \kappa, \quad \kappa_{12} = \kappa_g \cos \theta, \quad \text{and} \quad \kappa_{22} = \kappa \cos^2 \theta + \kappa_z \sin^2 \theta$$

Elimination of $\mathbf{B}$ generates the governing equation

$$\begin{pmatrix} (u^2/\nu - \kappa) & -j\kappa_g \cos \theta \\ j\kappa_g \cos \theta & (u^2/\nu - \kappa) + (\kappa - \kappa_z) \sin^2 \theta \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.7.13)$$

Notice that the sign of the off-diagonal elements changes with $\theta + \pi$: reversal of propagation direction generates a transposition of the eigenvectors. More on this to follow.

In order to arrive at a compact expression for the eigenvectors and eigenvalues, (3.7.13) is first rearranged to the form

$$\begin{pmatrix} p & -jq \\ jq & p + 2 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.7.14)$$

where

$$p = \frac{2(u^2/\nu - \kappa)}{(\kappa - \kappa_z) \sin^2 \theta}, \quad q = \frac{2\kappa_g \cos \theta}{(\kappa - \kappa_z) \sin^2 \theta}$$

The eigenvectors and corresponding eigenvalues of (3.7.14) are

$$|r_{\pm}\rangle = \frac{1}{\sqrt{2(1 + q^2 \pm \sqrt{1 + q^2})}} \begin{pmatrix} q \\ j(1 \pm \sqrt{1 + q^2}) \end{pmatrix} \quad (3.7.15)$$

and

$$p_{\pm} + 1 \pm \sqrt{1 + q^2} = 0 \quad (3.7.16)$$

The gyrotropic angle ψ is defined as

$$\tan 2\psi = \frac{2\kappa_g \cos \theta}{(\kappa - \kappa_z) \sin^2 \theta} \quad (3.7.17)$$

Identifying the gyrotropic angle with (3.7.15), the eigenvectors are expressed in the more revealing form:

$$|r_+\rangle = \begin{pmatrix} \sin \psi \\ j \cos \psi \end{pmatrix}, \quad |r_-\rangle = \begin{pmatrix} \cos \psi \\ -j \sin \psi \end{pmatrix}$$

In Stokes space the eigenvectors are

$$\hat{r}_+ = \begin{pmatrix} -\cos 2\psi \\ 0 \\ \sin 2\psi \end{pmatrix}, \quad \hat{r}_- = \begin{pmatrix} \cos 2\psi \\ 0 \\ -\sin 2\psi \end{pmatrix}$$

The eigenvectors of (3.7.13) correspond to polarization states that can propagate through the gyrotropic medium without change. The eigenstates of polarization are elliptical and the ellipsoid axes in the laboratory frame are aligned

to the horizontal and vertical. The eccentricity depends on the gyrotropic angle ψ , which in turn depends on the propagation direction, wavelength, and strength of the applied magnetic field. In contrast to birefringent materials where the eigen-polarizations are linear, gyrotropic materials can have any eigen-polarization along a line of longitude through S_1 . An input polarization state is resolved onto the two eigen-polarization states and in general the two states propagate with different phase velocities and energy-flow directions.

The phase-velocity eigenvalues of (3.7.13) are

$$u_{\pm}^2 = \nu \kappa_o \kappa_{\pm} \quad (3.7.18)$$

where $\varepsilon_o = 1/\kappa_o$ and

$$\kappa_{\pm} = \frac{\kappa}{\kappa_o} - \frac{(\kappa - \kappa_z) \sin^2 \theta}{2\kappa_o} \left[1 \pm \sqrt{1 + \left(\frac{2\kappa_g \cos \theta}{(\kappa - \kappa_z) \sin^2 \theta} \right)^2} \right] \quad (3.7.19)$$

The dispersion relations are therefore

$$k_{\pm} = \frac{\omega}{c} \frac{1}{\sqrt{\kappa_{\pm}}} \quad (3.7.20)$$

It follows that the effective refractive indices are $n_{\pm} = 1/\sqrt{\kappa_{\pm}}$. Heuristically, the refractive index has the form $n_{\pm} = n_o - \delta n \pm \delta n_g$, where n_o is the intrinsic refractive index, δn is a field-induced diminution of n_o , and δn_g is the splitting factor due to the cyclotron motion of the bound electrons.

Like birefringent media, the Poynting vectors in gyrotropic media do not necessarily align with the k -vectors of the eigenstates. Inserting (3.7.11) into (3.7.10) and remembering that $D_3 = 0$, the fields and flux densities are

$$\begin{aligned} \mathbf{D}_k &= \hat{e}_1 D_1 + \hat{e}_2 D_2, \\ \mathbf{E}_k &= \hat{e}_1 (\kappa_{11} D_1 + \kappa_{12} D_2) + \hat{e}_2 (\kappa_{21} D_1 + \kappa_{22} D_2) + \\ &\quad \hat{e}_3 (\kappa_{31} D_1 + \kappa_{32} D_2), \\ \mathbf{B}_k &= u/\nu (-\hat{e}_1 D_2 + \hat{e}_2 D_1), \\ \mathbf{H}_k &= u (-\hat{e}_1 D_2 + \hat{e}_2 D_1), \end{aligned}$$

The Poynting vector is therefore

$$\mathbf{S}_k = -\hat{e}_1 E_3 H_2^* + \hat{e}_2 E_3 H_1^* + \hat{e}_3 (E_1 H_2^* - E_2 H_1^*) \quad (3.7.21)$$

The Poynting vector is generally not aligned to the k -vector. In contrast to birefringent materials, the gyrotropic Poynting vector is tilted away from the k -vector along both the $\hat{e}_1$ and $\hat{e}_2$ axes, and out of the plane of $\mathbf{k}$ and $\hat{z}$. Only when $E_3 = 0$ does the Poynting vector align with the k -vector. The E_3 field component is

$$E_3 = j\kappa_g \sin \theta D_1 + (\kappa - \kappa_z) \sin \theta \cos \theta D_2$$

which shows that $E_3 = 0$ only when either $B_z = 0$ and/or $\sin \theta = 0$. Since $B_z \neq 0$ by design, the Poynting vector aligns with the k -vector only when the propagation direction is aligned with or counter to the magnetic field direction.

3.7.4 Faraday Rotation

Applications of gyrotropic materials in telecommunication components are generally limited to applications where the Poynting vector aligns with the k -vector. In the presence of a magnetic field, this is only possible when $\sin \theta = 0$. The polarization-transforming effect that is induced when a ray travels along the $\pm \hat{z}$ direction is called Faraday rotation. As detailed below, Faraday rotation is a nonreciprocal polarization rotation. The absence of reciprocity enables key components such as isolators and circulators to be realized.

The following analysis details propagation in the $+\hat{z}$ and $-\hat{z}$ directions. Since the signs will quickly become difficult to track, the analysis is divided into a first section that details $\theta = 0$ propagation and a second section that details $\theta = \pi$ propagation.

When $\theta = 0$, the ray travels along the $+\hat{z}$ direction. The governing equation (3.7.13) simplifies to

$$\begin{pmatrix} u^2 - \nu\kappa & -j\nu\kappa_g \\ j\nu\kappa_g & u^2 - \nu\kappa \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.7.22)$$

The gyrotropic angle ψ , as determined by (3.7.17), is

$$\psi = \pi/4$$

The eigenvectors of forward propagation are therefore

$$|r_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ j \end{pmatrix}, \quad |r_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -j \end{pmatrix} \quad (3.7.23)$$

The eigenvectors of the system are circular polarization states. The precession axis $\hat{r}$ in Stokes space for the forward propagation direction is $\hat{r} = +S_3$. The circular eigenstates of polarization are reminiscent of the circular electron orbital motion that was calculated at the beginning of §3.7.2. It should be no surprise that the cyclotron electron motion produces circular eigenpolarization states.

The phase-velocity eigenvalues of the system are

$$u_{\pm} = \sqrt{\nu(\kappa \mp \kappa_g)} \quad (3.7.24)$$

The wavenumbers and effective indices are found by substituting $u = \omega/k$, $\nu = 1/\mu_o$, and (3.7.9) for the permittivities. The wavenumbers are

$$k_{\pm} = \frac{\omega}{c} \sqrt{\frac{\varepsilon_r^2 - \varepsilon_{gr}^2}{\varepsilon_r \mp \varepsilon_{gr}}} \quad (3.7.25)$$

where $\varepsilon_r = \varepsilon/\varepsilon_o$ and $\varepsilon_{gr} = \varepsilon_g/\varepsilon_o$. In terms of susceptibility, the wavenumbers are

$$k_{\pm} = \frac{\omega}{c} \sqrt{1 + \chi_{11} \pm \chi_{12}} \quad (3.7.26)$$

The gyrotropic effect splits the intrinsic susceptibility by $\pm\chi_{12}$.

Since the $+\hat{z}$ propagation of the ray keeps the Poynting vectors aligned with the k -vectors, and further assuming normal incidence onto the gyrotropic medium to ensure alignment of the k -vectors, the polarization state of an incident ray is resolved into right- and left-circular components which co-propagate through the medium with different phase velocities. The circular basis of the eigenmodes makes the polarization state transform along a line of latitude on the Poincaré sphere. Indeed, Table 2.1 on page 67 shows that for an eigenvector that points to $+S_3$, the transformation matrix U is

$$U = \begin{pmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix} \quad (3.7.27)$$

where the birefringent phase is defined as usual:

$$\varphi = (k_+ - k_-)z \quad (3.7.28)$$

U rotates an input state of polarization about the $+S_3$ axis by angle φ and the rotation direction is right-handed. This rotation is illustrated in Fig. 3.19(a).

The contour traced by $-\pi \leq \varphi \leq \pi$ is a line of constant latitude. Recall from §1.4 that a family of polarization states on a line of latitude has constant ellipticity and handedness. It is the tilt of the major axis that rotates with longitude. In the laboratory frame, the effect of Faraday rotation is to rotate the major axis of the input polarization ellipse by angle $(\varphi/2)$.

Next, consider when $\theta = \pi$; the ray travels along the $-\hat{z}$ direction while the magnetic field still points in the $+\hat{z}$ direction. The governing equation (3.7.13) simplifies to

$$\begin{pmatrix} u^2 - \nu\kappa & j\nu\kappa_g \\ -j\nu\kappa_g & u^2 - \nu\kappa \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.7.29)$$

The gyrotropic angle ψ , as determined by (3.7.17), is

$$\psi = -\pi/4$$

The eigenvectors of backward propagation are therefore

$$|r_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -j \end{pmatrix}, \quad |r_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ j \end{pmatrix} \quad (3.7.30)$$

with corresponding eigenvalues

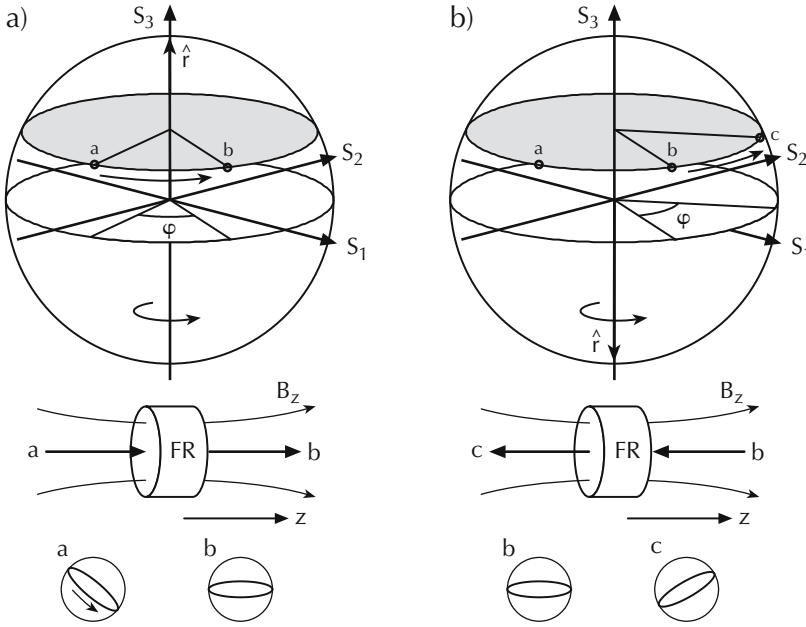


Fig. 3.19. Nonreciprocal polarization transformation via Faraday rotation. a) Forward propagation along the direction of the fixed magnetic field. Input polarization (a) right-hand precesses about $+S_3$ by φ to state (b). b) Backward propagation against the direction of the magnetic field. Polarization state (b) left-hand precesses about $-S_3$ by $|\varphi|$ to state (c).

$$u_{\pm} = \sqrt{\nu(\kappa \mp \kappa_g)} \quad (3.7.31)$$

In comparison with (3.7.23-3.7.24), the eigenvectors for backward propagation have flipped sign, while the eigenvalues remain unchanged. Rather than precessing about the $+S_3$ axis, the polarization state of a ray travelling backward precesses about the $-S_3$ axis. The transformation matrix U for backward propagation is

$$U = \begin{pmatrix} \cos(\varphi/2) & \sin(\varphi/2) \\ -\sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix} \quad (3.7.32)$$

where φ is defined by (3.7.28).

The off-diagonal sign of U has changed in comparison with (3.7.27) because of precession about $-S_3$. However, the precession expression remains the same. Since precession increases with forward travel and decreases with backward travel, the net result of the backward polarization rotation is a left-hand rotation about $-S_3$ by $|\varphi|$. Figure 3.19(b) illustrates this.

The analyses for $\pm \hat{z}$ propagation show that polarization states are transformed via right-hand-rule rotation about $+S_3$ for forward and backward

propagation. Regardless of travel direction, the polarization state precesses in the same direction in Stokes space. This is completely different behavior in comparison with a optical activity (covered in the following section), where a polarization transformation due to forward travel is undone by backward travel.

The nonreciprocal character of Faraday rotation is not tantamount to violation of time reversibility. Time reversal would change the spin direction of the electrons that generate the fixed magnetic field, in addition to mapping ccw to cw circular polarizations and vica-versa. All polarization transformations induced by Faraday rotation would be undone under time reversal.

At this point it is almost trivial to point out that when the input polarization state is linear, the output state is also linear but with its major axis rotated by the medium. This behavior is the basis for all telecommunications-grade Faraday rotators. As a rule, when the input polarization is a linear state, the output state is the orthogonal linear state after $\varphi = \pi$ rotation. There is some confusion in the literature that Faraday rotation magically drives *any* input state to its orthogonal state. This analysis has shown the invalidity of such a view.

For linear input polarizations, the Faraday angle θ_F is defined as

$$\theta_F = \varphi/2 \quad (3.7.33)$$

The rotation angle θ_F is the physical, not Stokes, rotation of the linear state. The physical Faraday angle θ_F is an important quantity when considering isolator and circulator designs because this angle determines the relative orientation between two polarizers.

3.7.5 The Verdet Constant

Considering that an external magnetic field is not necessarily uniform throughout the diamagnetic medium, the overall retardation from one end to the other is

$$\varphi = \int_0^L (k_+ - k_-) dz$$

In the transparency regime, the wavenumbers are related to the refractive indices in the usual way: $k_{\pm} = \omega n_{\pm}/c$. The splitting of the refractive index for small χ_{12} in light of (3.7.26) is

$$n_+ - n_- = \frac{\chi_{12}}{\sqrt{1 + \chi_{11}}} \quad (3.7.34)$$

The denominator is approximately the intrinsic refractive index to the same degree of approximation. Referring to (3.7.7a), the average refractive index is $n \simeq \sqrt{1 + \chi_{11}}$, or

$$n^2 = 1 + \frac{Ne^2}{m\varepsilon_o} \frac{1}{\omega_o^2 - \omega^2}$$

A small susceptibility χ_{12} occurs when the cyclotron resonance frequency follows $\omega\omega_c \ll (\omega_o^2 - \omega^2)$. Inserting this condition into (3.7.7b) eliminates the susceptibility terms from (3.7.34):

$$n_+ - n_- \simeq \frac{Ne^3 B_z}{nm^2 \varepsilon_o} \frac{\omega}{(\omega_o^2 - \omega^2)^2} \quad (3.7.35)$$

The refractive index splitting is also related to the dispersion of the intrinsic index $dn/d\lambda$ by

$$n_+ - n_- \simeq -\frac{\lambda^2 e}{2\pi cm} \left(\frac{dn}{d\lambda} \right) B_z$$

Putting all of this together, the functional form of the Faraday rotation angle is

$$\theta_F = V \int_0^L B_z dz \quad (3.7.36)$$

where the Verdet constant is identified as

$$V = -\frac{e\lambda}{2mc} \left(\frac{dn}{d\lambda} \right) \quad (3.7.37)$$

This expression is known as the Becquerel formula of the Verdet constant [3]. The negative sign is cancelled out by $dn/d\lambda$ for usual dispersions. The Verdet constant measures the rotary power of a material and can be used as a point of comparison between different materials. In SI units, the Verdet constant is measured in (rad/(m T)).

Fundamentally, the Verdet constant is a function of frequency and varies only to second order with the magnetic field strength. Generally it is well documented that the Verdet constant has a λ^{-2} wavelength dependence, as indicated in (3.7.35) [9]. Akin with the study of refractive index and material susceptibility, the Verdet constant of (3.7.37) is based here on a single oscillator model. More complicated materials may have multiple contributions to the Verdet constant, and the Verdet constant can certainly go negative.

3.7.6 Faraday Rotation in Ferrous Materials

The preceding analysis of diamagnetic materials resulted in an equation that linearly relates the Faraday rotation angle θ_F to the applied magnetic field B_z (cf. (3.7.36)). However, the Verdet constant for diamagnetic materials is too weak to find useful applications in telecommunications. In order to achieve a compact size the magnetization must be strong. Rare-earth iron garnets are ferrimagnetic materials that have orders of magnitude stronger rotary power than the best diamagnetic materials for the same applied magnetic field.

As outlined in §3.7.1, ferro-, antiferro-, and ferrimagnetic materials exhibit domain structure on a microscopic scale, where the magnetization within a domain is uniform and the domains spontaneously orient to prevent a large

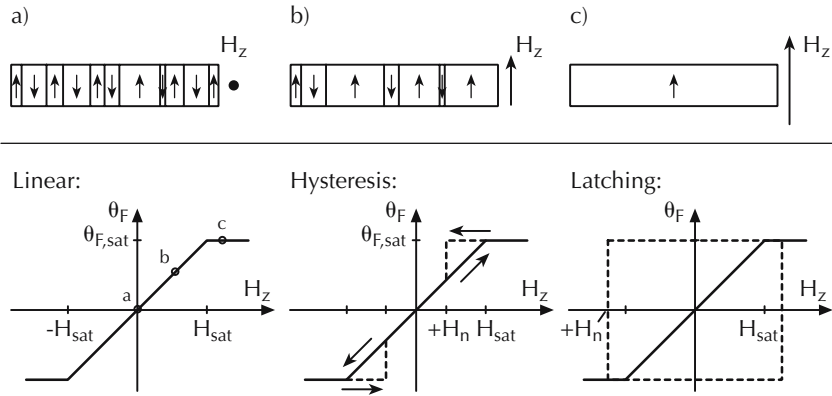


Fig. 3.20. Illustration of domain structure and alignment with applied external field. a) A demagnetized ferrimagnet: equal number of spin-up and spin-down domains. b) Partially magnetized material; some spin-down domains remain. c) Saturated material; a single spin-up domain spans the material. Below, saturation curves of θ_F vs. H_z : linear, hysteretic, and latching [14, 15].

constructive magnetic field (at least at high enough temperatures or in a demagnetized sample). In the presence of an external magnetic field the domains reorient preferentially along the field lines. However, unlike diamagnetic materials where the magnetization linearly tracks the applied field strength, all materials with domain structures saturate once a sufficiently strong field is applied. Saturation occurs when the domain boundaries are pushed to the edges of the material and only a single, unidirectional domain is left. No further magnetization occurs with increased applied field. Moreover, all magneto-optical effects are dominated by the internal magnetization rather than the applied field.

Figure 3.20 illustrates domain alignment and various magnetization curves. Figures 3.20(a–c) illustrate a demagnetized ferrimagnet having equal number of spin-up and spin-down domains (a), a partially magnetized ferrimagnet where spin-up domains expand at the expense of spin-down domains (b), and a saturated ferrimagnet where a single aligned domain spans the materials. Beyond the saturation field H_{sat} there is no further magnetization of the material.

The magnetization curves in the lower half of Fig. 3.20 illustrate three types of material responses. A linear magnetization curve has a one-to-one correspondence between H_z and θ_F ; above H_{sat} angle θ_F is no longer responsive and remains fixed at $\theta_{F,\text{sat}}$. A hysteretic magnetization curve is one where the magnetization, or equivalently the Faraday rotation angle θ_F , traces one path for increasing H_z and a separate path for decreasing H_z . In particular, once the applied field exceeds the saturation field, θ_F remains fixed at $\theta_{F,\text{sat}}$ until the applied field goes below the nucleation field H_{nuc} . The nucle-

ation field is the field strength where internal fields coerce a fracturing of the single domain to reduce the potential energy of the system. A latching magnetization curve has particular interest for component applications because once the saturation field is applied, the Faraday rotation remains at $\theta_{F,\text{sat}}$ even for zero external field.

When the Faraday rotation can saturate, the Verdet constant is no longer a reasonable measure since, by definition, the Verdet constant is the constant of proportionality between field and rotation. Instead, the specific rotation θ'_F is defined as

$$\theta'_F = \frac{\theta_{F,\text{sat}}}{L} \quad (3.7.38)$$

or the saturation rotation $\theta_{F,\text{sat}}$ per unit length. The specific rotation has units of (rad/m) [3].

In order for a ferrimagnet such as iron garnet to work well as a Faraday rotator it must be saturated. A demagnetized or partially magnetized element scatters light to an unacceptable degree. The light is scattered because the domains locally impart a polarization rotation, but the domains have no coherence or alignment. A radiation field with numerous k -vectors and polarization components must be constructed to match the boundary condition of the scattered light on the output face of the element.

That the ferrimagnet must operate in saturation is certainly a benefit because sensitivity to the applied field strength is eliminated. Without this built-in nonlinearity, much care would have to go into designing an external magnetic field that is highly uniform throughout the volume. A calculation of the magnetic field profile generated by a toroidal magnetic is given in Appendix B. Moreover, the saturation nonlinearity aids with the stringent aging requirements of all telecommunication components since the fixed magnet will degrade over time and still, with proper design, exceed the saturation field. A key design goal for an iron garnet is to have a low saturation field so that the requisite magnet can be small.

The remaining factors that must be accounted for when using iron garnets are the temperature sensitivity and wavelength variation of the specific rotation $\theta'_F(T, \lambda)$, and the wavelength dependence of the element [13–15].

3.8 Optically Active Materials

Materials that are chiral exhibit optical activity. A chiral material is one where the crystalline unit cell, or the molecular structure, differs from its mirror image; that is, the molecules have twist. A chiral molecule and its mirror image are called isomers. Most organic molecules are chiral, including sugar and DNA. For example, dextrose is right-handed sugar and fructose is left-handed sugar.

Chiral materials have a different radiation response to an optical field because nearest-neighbor dipole polarizations add constructively because of the

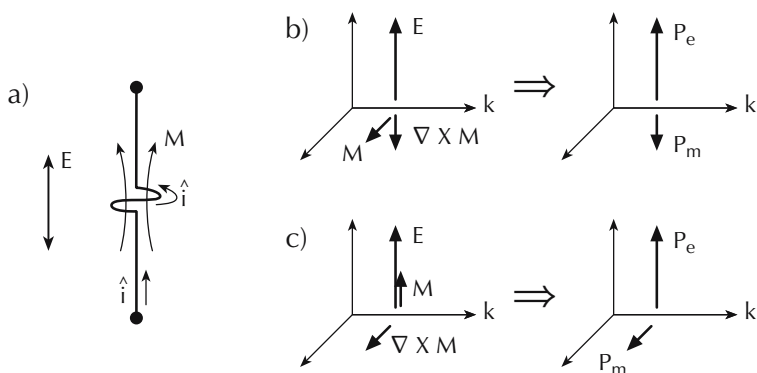


Fig. 3.21. Simple model of chiral molecular and induced polarization components [7]. a) Perfectly conducting wire of length ℓ with right-handed single turn. The applied electric field $\mathbf{E}$ induces current i , which generates both an electric polarization component $\mathbf{P}$ and a parallel magnetization component $\mathbf{M}$. b) An achiral material. $\mathbf{M}$ is perpendicular to $\mathbf{E}$ ($\mathbf{M} \parallel \mathbf{H}$), so magnetically induced polarization component $\mathbf{P}_m$ is parallel to $\mathbf{P}_e$ and makes an insignificant contribution. c) A chiral material. By construction, $\mathbf{M}$ is parallel to $\mathbf{E}$, generating magnetization contribution $\mathbf{P}_m$ perpendicular to $\mathbf{P}_e$. This one-sided persistent bias rotates the polarization state of the propagating wave.

molecular twist. The symmetry of isotropic and anisotropic materials cancels all nearest-neighbor dipole contributions, which is why they were ignored in the preceding sections. The twist of chiral materials induces both electric and magnetic dipole responses to an external electric field; these coupled responses cause optical activity.

Optical activity causes the right-hand or left-hand rotation of a linear polarization state as it propagates through the medium. A material that induces right-hand polarization rotation is called dextrorotatory and a material that induces left-hand rotation is called levorotatory. There is no *a priori* relation between the handedness of an isomer and dextro- or levo-rotatory behavior. However, if a right-handed isomer is dextrorotary, then its left-handed counterpart will be levorotatory.

The microscopic origins of optical activity can vary widely depending on the nature of the molecular or atomic structure. Optically active materials can be either reciprocal or non-reciprocal, bi-isotropic or bianisotropic, lossless or lossy. The constitutive relations differentiate these possible conditions. The common theme underlying optical activity is a coupling of the magnetic field to the electric polarizability, and the electric field to the magnetic polarizability. As an example, consider a perfectly conducting wire of length ℓ having a single spiral turn half-way along its length (Fig. 3.21(a)). When an external electric field generates a current that oscillates up and down along the wire, the current through the turn generates a magnetic flux in the direction of

the electric field. The external electric field thus elicits a collinear electric and magnetic response.

The handedness of the spiral turn determines whether the magnetic response is parallel or antiparallel with the electric response. Moreover, flipping the wire upside-down does not change the wire's handedness; handedness is an inherent property of the wire structure.

To pursue this example further, Hagan [7] considers Maxwell's equations in the absence of a current source:

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} (\mu_o \mathbf{H} + \mu_o \mathbf{M}) \quad (3.8.1a)$$

$$\nabla \times \mathbf{H} = \frac{\partial}{\partial t} (\varepsilon_o \mathbf{E} + \mathbf{P}) \quad (3.8.1b)$$

In a charge-free region, the divergence of the electric field is zero. Rewriting in time-harmonic form and using the vector identity $\nabla \times \nabla \times = \nabla(\nabla \cdot) - \nabla^2$ gives

$$\nabla^2 \mathbf{E} = -\omega^2 \mu_o \mathbf{P}_{\text{eff}} \quad (3.8.2)$$

where an effective polarization density $\mathbf{P}_{\text{eff}}$ is identified as

$$\mathbf{P}_{\text{eff}} = \left(\varepsilon \mathbf{E} - \frac{j}{\omega} \nabla \times \mathbf{M} \right) \quad (3.8.3a)$$

$$= \mathbf{P}_e + \mathbf{P}_m \quad (3.8.3b)$$

The polarization of the medium has two contributors: $\mathbf{P}_e$ associated with the linear dielectric and $\mathbf{P}_m$ associated with the curl of the induced magnetic flux. In achiral materials $\mathbf{E} \cdot \mathbf{H} = 0$, so the curl of the magnetic flux ($\nabla \times \mathbf{M}$) is parallel or antiparallel with the polarization $\mathbf{P}_e$. Since generally $|\varepsilon \mathbf{E}| \gg |\mathbf{M}/c|$, the $\mathbf{P}_m$ contribution is negligible. However, in chiral materials, the $\mathbf{P}_m$ component lies perpendicular to $\mathbf{P}_e$. That is, since a component of $\mathbf{M}$ is generated parallel with $\mathbf{E}$, $\nabla \times \mathbf{M}$ lies perpendicular. It is the $\mathbf{P}_m$ component that distinguishes optical activity from other processes.

In order to analyze this further, the magnetization must be related to the electric field. The canonical constitutive relation between $\mathbf{M}$ and $\mathbf{H}$ is

$$\mathbf{M} = \chi_m \mathbf{H} \quad (3.8.4)$$

Now, given the particular geometry of the wire with embedded loop, the magnetic flux generated by current flow through the loop is parallel to the applied field. Moreover, the current lags the voltage due to the loop inductance. With these considerations, (3.8.4) can be rewritten as

$$\mathbf{M} = \pm j \chi_m |\mathbf{H}| \hat{\mathbf{E}} \quad (3.8.5)$$

where $|\mathbf{H}|$ is the magnitude of the magnetic field and $\hat{\mathbf{E}}$ is the direction of the electric field. The $\pm$ sign accounts for the handedness of the loop, $(-)$ for

a right-hand loop and (+) for a left-hand loop. From (3.8.1b), the magnetic field magnitude is

$$|\mathbf{H}| = \frac{\omega}{k} \left| \varepsilon \mathbf{E} - \frac{j}{\omega} \nabla \times \mathbf{M} \right| \quad (3.8.6)$$

Since $\varepsilon \mathbf{E}$ dominates the right-hand side magnitude, the curl term is neglected. The curl of $\mathbf{M}$ in (3.8.5) is then just

$$\nabla \times \mathbf{M} \simeq \pm j \chi_m \frac{\omega \varepsilon}{k} \nabla \times \mathbf{E} \quad (3.8.7)$$

The effective polarization density is therefore

$$\mathbf{P}_{\text{eff}} = \varepsilon \mathbf{E} + \varepsilon \beta \nabla \times \mathbf{E} \quad (3.8.8)$$

where the chirality parameter is defined as $\beta = \pm \chi_m / k$. The sign of the chirality parameter β designates the handedness, and the units of β are length. The constitutive relation between $\mathbf{D}$ and $\mathbf{E}$ are therefore

$$\mathbf{D} = \varepsilon_{\text{DBF}} (\mathbf{E} + \beta \nabla \times \mathbf{E}) \quad (3.8.9)$$

This is the Drude-Born-Fedorov constitutive relation for reciprocal chiral material. Detailed investigation of conservation of energy [4] requires a complementary constitutive relation for the magnetic flux vector,

$$\mathbf{B} = \mu_{\text{DBF}} (\mathbf{H} + \beta \nabla \times \mathbf{H}) \quad (3.8.10)$$

The notation of ε_{DBF} and μ_{DBF} follows from [12] and is used to distinguish these values when compared to more general forms of the constitutive relations. Of immediate importance is the existence of $\nabla \times$ terms in the constitutive relations. The ∇ part is a spatial derivative, which physically means that neighboring fields contribute to the polarization. The cross in $\nabla \times$ indicates that the neighboring field contributions are perpendicular to the applied field. The presence of the $\nabla \times$ terms in (3.8.9-3.8.10) generates a persistent bias perpendicular to the propagation direction which rotates the fields in a circular motion. Circular states of polarization are in fact the eigenstates of optical activity.

The above derivation is heuristic but not particularly rigorous. More sophisticated calculations start with dipole-dipole interactions within chiral molecules and proceed to generate coupled equations of bound-electron motion. From these the spatially averaged polarization and magnetization vectors are derived [2]. While telecommunications applications rarely require more detailed knowledge, bio-optics is replete with applications of molecular mechanics and optical activity. The interested reader is referred to [1, 2, 12].

3.8.1 Propagation in Bi-Isotropic Media

The most general constitutive relations for bi-isotropic optically active media are [12]

$$\mathbf{D} = \varepsilon \mathbf{E} + (\chi_T - j\kappa_P)\sqrt{\mu_o\varepsilon_o} \mathbf{H} \quad (3.8.11a)$$

$$\mathbf{B} = \mu \mathbf{H} + (\chi_T + j\kappa_P)\sqrt{\mu_o\varepsilon_o} \mathbf{E} \quad (3.8.11b)$$

where χ_T is the Tellegen magnetoelectric parameter and κ_P is the Pasteur chirality parameter. When $\chi_T \neq 0$ the medium is non-reciprocal. When $\kappa_P \neq 0$ the medium is chiral. Either or both terms can be non-zero. The constitutive relations (3.8.11) follow from (3.8.9-3.8.10) for a proper association of constants $(\varepsilon_{DBF}, \mu_{DBF}, \beta) \rightarrow (\varepsilon, \mu, \kappa_P)$ [12] and for $\chi_T = 0$. That is, the phenomenological derivation of the preceding section was based on a purely reciprocal effect. A Tellegen material is well beyond the scope of this text, but suffice it to say that it is a complex material having bound-electric and magnetic permanent dipoles. A Pasteur material is an isotropic chiral material, where the axial direction of the helices are uniformly randomly oriented. An alignment of the helical axes makes the material anisotropic, which has the implication that κ_P becomes a tensor.

In the following analysis, a lossless, reciprocal, bi-isotropic medium is considered. The constitutive relations (3.8.11) form the scalar transfer relation

$$\mathbf{C}_{EH} = \begin{pmatrix} \varepsilon & -j\kappa_P \\ j\kappa_P & \mu \end{pmatrix} \quad (3.8.12)$$

The *kDB* formalism requires the inverse constitutive relations $\mathbf{C}_{DB} = \mathbf{C}_{EH}^{-1}$. The bi-isotropic *kDB* constitutive relations are written as

$$\mathbf{E} = \kappa \mathbf{D} + j\chi \mathbf{B} \quad (3.8.13a)$$

$$\mathbf{H} = -j\chi \mathbf{D} + \nu \mathbf{B} \quad (3.8.13b)$$

There is no transformation of the constitutive parameters into *kDB*, as they are all scalars. Substitution into the coupled *kDB* equations (3.3.5) yields

$$\kappa \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = - \begin{pmatrix} j\chi & -u \\ u & j\chi \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \quad (3.8.14a)$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = - \begin{pmatrix} -j\chi & u \\ -u & -j\chi \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \quad (3.8.14b)$$

Elimination of $\mathbf{B}$ generates the governing equation

$$\begin{pmatrix} u^2 + \chi^2 - \kappa\nu & 2ju\chi \\ -2ju\chi & u^2 + \chi^2 - \kappa\nu \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \quad (3.8.15)$$

The eigenvectors of (3.8.15) are circular states:

$$|r_+\rangle = \begin{pmatrix} 1 \\ j \end{pmatrix}, \quad |r_-\rangle = \begin{pmatrix} 1 \\ -j \end{pmatrix} \quad (3.8.16)$$

with corresponding phase-velocity eigenvalues

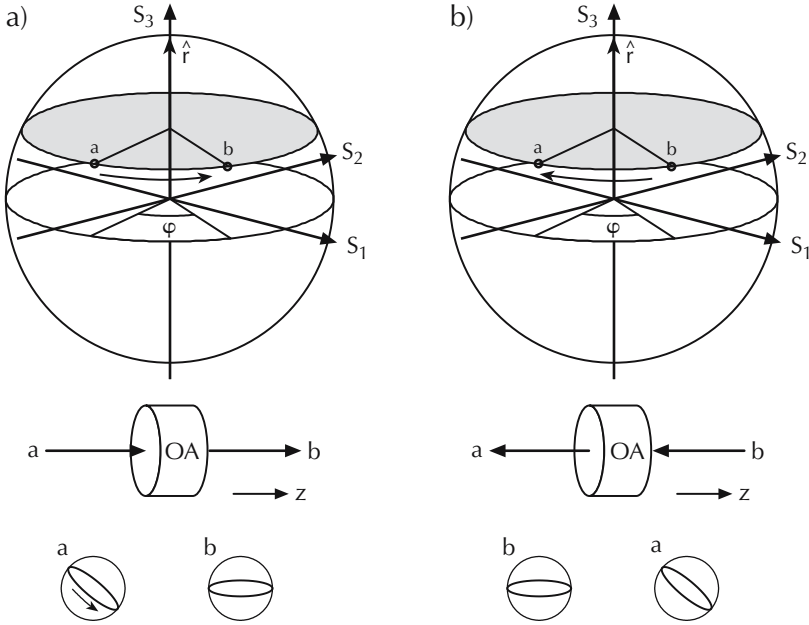


Fig. 3.22. An optically active medium is reciprocal. a) Forward travel. The precession axis is $+S_3$; the eigenvectors are circular polarization states. Transit through the medium transforms an input polarization state from (a) to (b). b) Backward travel. The precession axis remains $+S_3$. Transit through the medium transforms input polarization state (b) to (a).

$$u_{\pm} = \sqrt{\kappa\nu} \pm \chi \quad (3.8.17)$$

The corresponding wavenumbers are

$$k_{\pm} = \frac{\omega}{c} \frac{n}{1 \pm n\chi/c} \quad (3.8.18)$$

where n is an average index. For positive χ values, right-hand circularly polarized light travels slower than left-hand polarization.

As the constitutive relations in the present model are scalars, the Poynting vector coincides with the k -vector in the medium. Polarization transformation therefore occurs through transit. In Stokes space, the precession axis $\hat{r}$ points along $+S_3$. The precession angle is determined by the phase difference (3.7.28). The transformation matrix U is therefore

$$U = \begin{pmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix} \quad (3.8.19)$$

For positive χ values the precession in Stokes space for $+z$ travel follows $|\varphi|$.

Polarization transformation due to optical activity is similar to that of Faraday rotation: the precession axes for both mechanisms are $\pm S_3$. This is in stark contrast to birefringent transformation where the precession axis lies in the (S_1, S_2) plane. However, unlike Faraday rotation, when $\chi_T = 0$ optically active media is reciprocal (see Fig. 3.22). Consider the upper-right-hand matrix component from Faraday and OA governing equations (3.7.13, 3.8.15):

Faraday	$-j\kappa_g\nu \cos \theta$
Optical Activity	$2ju\chi$

For Faraday rotation, the sign of the off-diagonal term changes when the propagation direction is reversed. This is not the case with optical activity, where the sign is unaffected by direction. Therefore the eigenvectors for OA do not change when the propagation direction is reversed.

Like Faraday rotation, a chiral medium will rotate a linear polarization state from one angle to another. The rotary power ρ of a chiral material is this polarization rotation per unit length. From the above analysis, the rotary power is $\rho = \varphi/2z$. Biot's law (circa 1812) gives a phenomenological although rather accurate wavelength dependence of the rotary power:

$$\rho = a + \frac{b}{\lambda^2} \quad (3.8.20)$$

Drude in the nineteenth century proposed an extension of Biot's to account for multiple material resonances, akin to Sellmeier's equations. Drude's equation is

$$\rho = \sum_i \frac{b_i}{\lambda^2 - \lambda_o^2} \quad (3.8.21)$$

These models are complex to derive. For more information on the relevant expressions and the tools required for derivation, see [10].

References

1. H. Ammari, K. Hamdache, and J. Nédélec, "Chirality in the Maxwell equations by the dipole approximation," *SIAM Journal of Applied Math.*, vol. 59, pp. 2045–2059, 1999.
2. D. J. Caldwell and H. Eyring, *The Theory of Optical Activity*. New York: Wiley-Interscience, 1971.
3. M. N. Deeter, G. W. Day, and A. H. Rose, *CRC Handbook of Laser Science and Technology, Supplement 2: Optical Materials*. Boca Raton, Florida: CRC Press, 1995, ch. Magneto-optic Materials, pp. 367–402.
4. F. Fedorov, "On the theory of optical activity of crystals. I. Energy conservation law and optical activity tensors, optics and spectroscopy," *Optics and Spectroscopy*, vol. 6, pp. 85–93, 1959.
5. G. R. Fowles, *Introduction to Modern Optics*. New York: Dover Publications, 1989.
6. V. J. Fratello and R. Wolfe, *Handbook of Thin Film Devices, Vol. 4: Magnetic Thin Film Devices*. San Diego: Academic Press, 2001, ch. Epitaxial Garnet Films for Nonreciprocal Magneto-Optic Devices, pp. 93–141.
7. D. J. Hagan, private communication, 2002, from lecture notes, School of Optics, University of Central Florida. [Online]. Available: <http://www.creol.ucf.edu/>
8. H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*. Englewood Cliffs, New Jersey: Prentice-Hall, 1989.
9. A. Jain, J. Kumar, F. Zhou, L. Li, and S. Tripathy, "A simple experiment for determining verdet constants using alternating current magnetic fields," *Am. J. Phys.*, vol. 67, pp. 714–717, 1999.
10. W. Kaminsky, "Experimental and phenomenological aspects of circular birefringence and related properties in transparent crystals," *Rep. Prog. Phys.*, vol. 63, pp. 1575–1640, 2000.
11. J. A. Kong, *Electromagnetic Wave Theory*. New York: John Wiley & Sons, 1989.
12. I. Lindell, A. Sihvola, S. Tretyakov, and A. Viitanen, *Electromagnetic Waves in Chiral and Bi-Isotropic Media*. Boston, Massachusetts: Actech House, 1994.
13. K. B. Rochford, A. H. Rose, and G. Day, "Magneto-optic sensors based on iron garnets," *IEEE Transactions on Magnetics*, vol. 32, no. 5, pp. 4113–4117, 1996.
14. K. Shirai, K. Ishikura, and N. Takeda, "Low saturated magnetic field bismuth-substituted rare earth iron garnet single crystal and its use," U.S. Patent 5,512,193, Aug. 30, 1996.
15. K. Shirai and N. Takeda, "Faraday rotator," U.S. Patent 5,535,046, July 9, 1996.

Elements and Basic Combinations

4.1 Wavelength-Division Multiplexed Frequency Grid

Specification of an internationally recognized standard for the channel frequency locations of a wavelength-division multiplexed optical communication system is essential for component development and system interoperability. Optical components such as signal lasers, multiplexers and demultiplexers, wavelength add-drop filters, and interleavers require a channel location specification to design to.

The current standard for dense wavelength-division multiplexed (DWDM) transmission is specified by the document ITU G694.1 [30]. An anchor frequency of 193.1 THz is defined off of which all other channel locations can be derived. Channel spacings are defined for 12.5 GHz, 25 GHz, 50 GHz, and 100 GHz separations. Figure 4.1 illustrates a partial frequency grid at 100 GHz channel separation. The channel frequency locations f_n are defined as

$$f_n = 193.1 + n \times C \quad (\text{THz}) \quad (4.1.1)$$

where n is a positive or negative integer including zero, and where $C = 0.0125$, $C = 0.025$, $C = 0.050$, or $C = 0.100$. To convert between frequency and wavelength, the value of the speed of light found on page 4 is used by the Standard. Other than the anchor frequency and channel separations there is no adopted standard for how many channels can run on a DWDM system or where those channels are located. The standard is agnostic to implementation.

Nonetheless, there is the practical issue that an optically amplified multi-span link requires periodic amplification of the optical signal. An optical amplifier amplifies all channels at once [18]. There are, however, limitations on the bandwidth of the gain due to the physical nature of the rare-earth ions such as erbium that are doped into the optical fiber. There is no single bandwidth that can be attributed to an optical amplifier because the useable bandwidth is governed by both the particular architecture of the amplifier as well as the system application. What is true is that communications carriers who pur-

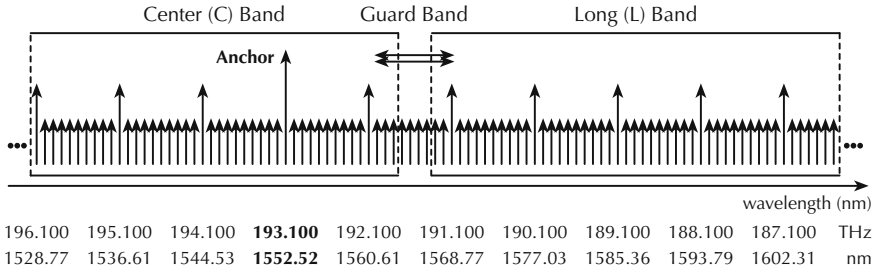


Fig. 4.1. Overview of ITU-T G.694.1 spectral grid for DWDM applications. The anchor frequency is located at 193.100 THz. As illustrated, channel centers are spaced by 100 GHz above and below the anchor frequency. While the Standard is open to higher and lower frequencies than indicated, rough demarcation of center (C) and long (L) bands, with possible intermediate guard band, is shown.

chase the optical transport systems want the lowest overall system cost for the largest aggregate transmission.

There has evolved a banding of the spectrum based on the optical amplifier architectures that are economically manufactured. The C-band, or center band, is the original band where an erbium-doped optical amplifier provides high gain efficiency. This band is often delineated by the range 192.1 to 196.1 THz. The L-band, or long band, is recently available using erbium-doped fiber. That band is often delineated by the range 186.1 to 191.1 THz. These ranges vary from vendor to vendor. Because the C- and L-bands have different gain efficiencies for pumps wavelengths of 1480 or 980 nm, separate amplifiers are built for these each band. A Raman amplifier can pump the entire spectrum seamlessly, however. For diode-pumped systems, a band-separation filter splits the two bands prior to amplification and then combines them prior to transmission. Typically there is a guard band between the C- and L-bands to accommodate the band-separation filter. However, recently demonstrated filter improvements can eliminate the need for a guard band.

There is no standard for the deviation of laser or filter center frequency from the center frequency of the grid [29]. The end-of-life specification depends on many factors, including the maximum foreseeable channel density. For the purposes of analysis in this text, the allowable beginning-of-life frequency deviation will be taken as $\Delta f = \pm 2.5$ GHz.

There are several reasons the definition of the spectral grid is important for component designers. These reasons include

- The tolerance on the FSR of a periodic component must allow for channel alignment across a band.
- The frequency centering of a periodic component with the correct FSR must align to the channel locations.
- The bandwidth of polarization transforming elements such as waveplates must cover a band.

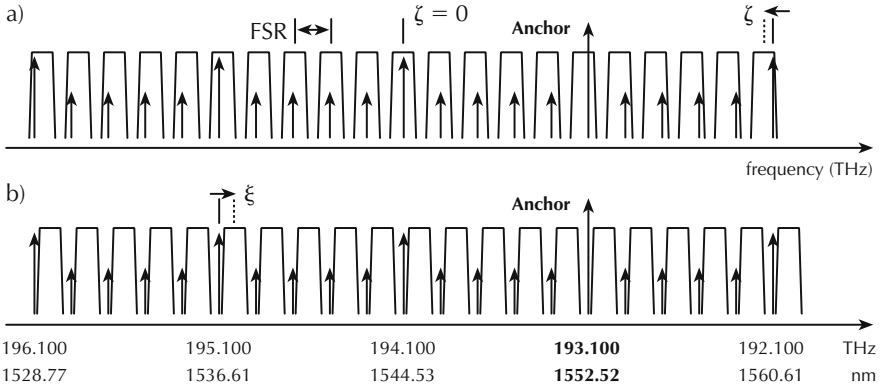


Fig. 4.2. Two types of filter placement errors in relation to the DWDM spectral grid. a) FSR error leads to walkoff between spectral grid and filter centers. While at one frequency the grid and filter may align ($\zeta = 0$) at the band edges the filter is misaligned. b) Frequency location error ξ , often called phase error. Even if the FSR is within tolerance, the filter center frequencies may suffer a common misalignment to the spectral grid.

Figure 4.2 illustrates the first two error types. In Fig. 4.2(a) the FSR of a periodic element such as an interleaver filter is too small, leading to a walkoff over the band. Denote the frequency separation between the grid and the filter at either band end as ζ . The tolerable FSR error of a component is then

$$\frac{\delta \text{FSR}}{C} = \frac{|C - \text{FSR}|}{C} = \frac{|\zeta|}{NC} \quad (4.1.2)$$

where C is the designed channel separation and N is the number of channels between band center and band edge. For example, in a C-band designed with 40 channels on 100 GHz centers, $N = 20$. With the aforementioned filter location tolerance of $\Delta f_n = \pm 2.5$ GHz, the FSR tolerance of the filter is

$$\frac{\delta \text{FSR}}{C} = \frac{|2.5|}{20 \times 100} = 0.125\% \quad (4.1.3)$$

For a resonant element such as a Fabry-Perot, a 0.125% tolerance for a 1 mm long cavity is about one micron. The broader the band coverage the tighter the cavity-length tolerance.

In Fig. 4.2(b) $\delta \text{FSR} = 0$, but there is a common frequency offset error ξ across all the channels. The frequency tolerance is generally more than the FSR tolerance because the error does not accumulate across multiple channels. As such, the frequency tolerance δf is

$$\frac{\delta f}{\text{FSR}} \simeq \frac{|\Delta f|}{C} \quad (4.1.4)$$

As with the preceding example, the frequency tolerance is $\delta f = 2.5\%$ FSR.

In relation to polarization elements and material dispersion, the spectral coverage of the system should be tolerated by the component. For a C-band that covers 192.1 to 196.1 THz, the spectral coverage $\mathcal{S}_C$ of this band is

$$\mathcal{S}_C = \frac{196.1 - 192.1}{194.1} \simeq 2.061\% \quad (4.1.5)$$

Likewise, an L-band that covers 186.1 to 191.1 THz has a spectral coverage $\mathcal{S}_L$ of

$$\mathcal{S}_L = \frac{191.1 - 186.1}{188.6} \simeq 2.651\% \quad (4.1.6)$$

The combined spectral coverage is $\mathcal{S}_{C+L} = 5.23\%$. These are broad coverages for waveplates, Faraday rotators, and anti-reflection coatings to cover. In many cases compromises or increased complexities are required to satisfy these bandwidth demands.

4.2 Properties of Select Materials

A central theme of this text is to provide information necessary to realize the components whose architectures are detailed in subsequent chapters. To this end it is essential to have on hand the physical and optical properties of the relevant materials. This section tabulates the properties of select isotropic optical materials, select birefringent materials, and select metal and alloy materials used for packaging. Faraday rotators made from iron-garnet derivatives are covered at the end. These are not intended to be an exhaustive tabulation but rather a short reference for the commonly used materials.

4.2.1 Isotropic Glass Materials

Isotropic glass materials are used both as optical elements and as packaging and assembly parts. There are a large number of well-characterized glasses available from major suppliers [28, 38, 45]. Principal factors used to select a low-loss glass for optical transmission use include its refractive index at the wavelength of use; the refractive index dispersion; its thermal-optic coefficient, or change of refractive index with temperature; and its thermal-expansion coefficient [44]. The refractive index and its dispersion will govern such attributes such as the angle of a prism made from a particular glass, while the thermal-optic coefficient is necessary to tolerance the component over the required temperature range. The thermal expansion coefficient is important because a package assembled from a variety of materials must maintain its integrity over its lifetime. If one part expands significantly more than the others then adhesion, for example, can be compromised.

Glass parts are sometimes used for packaging and assembly parts as well. Two key factors when choosing such as glass are its thermal expansion coefficient and its ultraviolet (UV) transmissivity. A glass package part is often

Table 4.1. Select Isotropic Glass Materials

Property	Fused silica ^(a)	N-BK7 ^(b)	Units
Density	2.2	2.51	g/cm ³
Hardness	5-6		Mohs
Thermal conductivity	1.31	1.114	W/(m·K)
Thermal expansion	0.5 ^(c)	7.1 ^(d)	×10 ⁻⁶ /K
Refractive index $n_{1529.6}$	1.44424	1.50091	
Thermal optic ^(e)			
1060 nm line		2.4	×10 ⁻⁶ /K
e-line	9.9	3.0	
Sellmeier coefficients ^(f)			λ in μm
B1	0.66942	1.03961	
B2	0.43458	0.23179	
B3	0.87169	1.01047	
C1	0.00448	0.00600	
C2	0.01328	0.02002	
C3	95.3414	103.561	

^(a) Reported by Schott [45].

^(b) Reported by Schott [46].

^(c) 25°C-100°C, ^(d) -30°C-+70°C, ^(e) 20°C-40°C, $\Delta n_{\text{rel}}/\Delta T$.

^(f) $n^2(\lambda) - 1 = \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3}$, [45].

used because its thermal expansion matches with other glassy parts that are in the transmission path. Another reason is that the transmission parts need to be visible during assembly, for alignment and/or for UV tacking with epoxy. Finally, glass windows are sometimes brazed into metal packages to make a clear path for collimators while maintaining hermetic integrity.

While a complete glass catalog should be referred to in order to choose an optical glass for a particular application, Table 4.1 provides select material properties of two commonly used transmission and packaging glasses: fused silica and BK7.

4.2.2 Birefringent Crystals

Birefringent crystals are the basic building blocks for the birefringent components detailed in the following chapters. The crystal materials may be divided into two application regimes: applications requiring high birefringence and those requiring low birefringence. Rutile and yttrium orthovanadate (YVO₄) are examples of very high birefringent material, both having about 10% birefringence at 1.55 μm. Crystalline quartz is a readily available low birefringent material, having a birefringence of 0.0084 at 1.55 μm.

Materials of intermediate birefringence are nonetheless required for practical reasons. For example, the birefringent phase of a birefringent crystal is tem-

Table 4.2. Select Birefringent Material Properties

Property	YVO ₄			LiNbO ₃			α-BBO			CaCO ₃ (Calcite)			Units
	a	c		a	c		a	c		a	c		
Crystal type	Tetragonal			Hexagonal			Hexagonal			Hexagonal			
Birefringent type	+ uniaxial			− uniaxial			− uniaxial			− uniaxial			
Space group	D4h			R3c			R3			R3c			
Density	4.22 ^(a)			4.65 ^(b)			3.84 ^(s)			2.711 ^(s)			g/cm ³
Hardness ^(a)	5			5			4.5			3			Mohs
Hydroscopic susceptibility ^(a)	none			none			low			low			
Lattice constants	7.12 ^(c)	6.29	5.151 ^(b)	13.866	12.547 ^(s)	12.736	4.990 ^(s)	17.060					Å
Thermal conductivity	5.10 ^(a)	5.23	4.2 ^(b)		0.08 ^(s)	0.80	5.1 ^(s)	6.2					W/(m·K)
Thermal expansion	4.43 ^(a)	11.37	15 ^(b)	7.5	0.5 ^(s)	33.3	−3.7 ^(s)	25.1					×10 ^{−6} /K
Refractive index	1.9447 ^(e)	2.1486	2.2112 ^(f)	2.1381	1.6749 ^(g)	1.5555	1.6629 ^(h)	1.4885					
Birefringence	+0.2039			−0.0732			−0.1194			−0.1744			
Thermal optic	8.5 ^(c)	3.0	4.1 ^(f)	35.0	−9.3 ^(a)	−6.6	2.1 ^(s)	11.9					dn/dT ×10 ^{−6} /K
Group index ^(d)	1.9787	2.1914	2.2643	2.1856	1.6587	1.5295	1.6586	1.4820					
Group birefringence ^(d)	+0.2127			−0.0787			−0.1292			−0.1766			
Thermal optic ^(d)	7.99	4.59	15.14	33.08	1.41	5.21	−4.84	1.29					$\frac{1}{n_g L} \frac{dn_g L}{dT} \times 10^{-6}/K$
Sellmeier coefficients	A1	3.77834 ^(a)	4.59905	4.9048 ^(f)	4.582		2.69705 ^(a)	2.18438					λ in μm
	A2	0.069736	0.1105	0.1176	0.09921		0.01921	0.00873					
	A3	0.04724	0.04813	0.04753	0.04448		0.0182	0.01018					
	A4	0.010813	0.01227	0.02715	0.02194		0.01516	0.00244					

^(a) Reported by Casix [32]; ^(e) reported for 1550 nm, ^(g) reported for 532 nm, ^(h) reported for 1300 nm.
^(b) Reported by Crystal Technology [14], ^(f) from Sellmeier at 1.55 μm and 24.5°C.
^(s) Reported by Handbook of Optics [3].
^(d) Reported in this text by Damask, 25°C–100°C, 1515–1575 nm. Sellmeier expression: $n^2(\lambda) = A_1 + \frac{A_2}{\lambda^2 - A_3} - A_4\lambda^2$

Table 4.3. Select Birefringent Material Properties

Property	Crystal Quartz		Rutile		Lead Molybdate		Tellurium Dioxide		Magnesium Fluoride		Units
	SiO ₂ ^(s)		TiO ₂ ^(s)		PbMoO ₄ ^(c)		TeO ₂ ^(s)		MgF ₂ ^(s)		
	a	c	a	c	a	c	a	c	a	c	
Crystal type	Hexagonal		Tetragonal		Tetragonal		Tetragonal		Tetragonal		
Birefringent type	+ uniaxial		+ uniaxial		− uniaxial		+ uniaxial		+ uniaxial		
Space group	P3 ₂ 21		P4 ₂ /mmm		I4 ₁ /a		P4 ₁ 2 ₁ 2		P4 ₂ /mmm		
Density	2.648		4.25		6.95		6.019		3.171		g/cm ³
Hardness	7		7		3		4		6		Mohs
Hydroscopic susceptibility	none		none				none		none		
Lattice constants	4.9136	5.4051	4.594	2.962	5.4312	12.1065	4.810	7.613	4.623	3.053	Å
Thermal conductivity	7.5	12.7	8.3	11.8					30	21	W/(m·K)
Thermal expansion	12.38	6.88	6.86	8.97	12.4	26.7	15.0	4.9	9.4	13.6	×10 ^{−6} /K
Thermal optic	−6.2 ^(st)	−7.0	4 ^(st)	−9		190 ^(ct)	9 ^(st)	8	0.88 ^(st)	0.32	dn/dT ×10 ^{−6} /K
Refractive index	1.5352	1.5440	2.432	2.683	2.260	2.170	2.18	2.32	1.3734	1.3851	
Birefringence	+0.0088		+0.251		−0.090		+0.14		+0.0117		deg/mm
Optical rotary power							11.5 ^(d)				

^(s) Reported by Handbook of Optics [3]: ^(st) reported at 546 nm, ^(st) reported at 405 nm, ^(st) reported at 644 nm, ^(st) reported at 1.15 μm.

^(c) Reported by Isomet Corporation [31], ^(ct) Δn_g = −0.104, d(Δn_g)/dT is reported.

^(d) At λ = 1.55 μm. Follows Biot's law: ρ = ρ_o/λ².

perature dependent. But two crystals having complementary thermal-optic coefficients can be cascaded to mitigate the temperature dependence passively. Useful complementary crystals for YVO_4 are LiNbO_3 and lead molybdate, both of which have an intermediate birefringence. Alternatively, intermediate birefringent crystals such as LiNbO_3 have strong electro-optic coefficients, making them useful for switching and polarization-control applications.

Tables 4.2 and 4.3 provide a compilation of data on nine widely used birefringent materials. All of these materials are synthetic except for calcite, which is found in nature and mined. Because the crystal quality can vary and the material is not sufficiently hydrophobic, calcite is not a favored component-grade material. In comparison with rutile, YVO_4 is favored because it is more easily grown to large boule sizes and is easier to handle at the grinding and polishing stages. To compensate for the temperature dependence of YVO_4 , LiNbO_3 is commonly used, although lead molybdate has been recently proposed. Tellurium dioxide is a reasonably high birefringent crystal with the added feature that it can be grown in high-purity dextrorotatory and laevorotatory chiralities, enabling the crystal to be used for optical activity. α -BBO is a negative uniaxial crystal that is not commonly used due to its intermediate birefringence, but one interleaver design pairs its negative uniaxial property with the positive uniaxial YVO_4 to make a compound crystal that imparts no Poynting vector walkoff.

Finally, crystalline quartz is extensively used for waveplates because of its high material quality and its low birefringence. The birefringent beat length of quartz is about $184\ \mu\text{m}$ at $1.55\ \mu\text{m}$ wavelength, so a true zero-order half-wave waveplate at this wavelength is $92\ \mu\text{m}$ thick. In contrast, the birefringent beat length of LiNbO_3 is about $19\ \mu\text{m}$; an equivalent half-wave waveplate is $9.5\ \mu\text{m}$ thick, a nearly impossible thickness to reproducibly attain.

4.2.3 Iron Garnets

Faraday rotators (FR) provide the requisite nonreciprocal polarization rotation necessary for isolators and circulators. The optical properties of Faraday rotation were detailed in §3.7.6. Here the material properties and design goals of iron garnet FRs are outlined.

An iron garnet is a ferrimagnetic material that has several orders of magnitude greater rotary power than diamagnetic materials. Practical iron garnets are most commonly grown by the liquid-phase epitaxy. A substrate such as gadolinium gallium garnet (GGG) is used to nucleate the crystalline growth and support the film. After the film is grown, usually to $250\text{--}500\ \mu\text{m}$ thick, the substrate is ground away and both sides of the remaining film are polished. Anti-reflection coatings are then applied and the film is diced into square parts, typically $2 \times 2\ \text{mm}^2$.

To use an iron garnet FR, the finished part is placed at the center of an annular permanent magnet such as samarium-cobalt (Sm-Co). The major face of the garnet part is placed in the clear aperture of the permanent magnet and,

Table 4.4. Non-Latching Iron Garnet Design Goals

High specific Faraday rotation	$\theta'_F > 0.1^\circ/\mu\text{m}$
Low thickness for $\theta'_F = 45^\circ$	$t < 500 \mu\text{m}$
Low absorption	$\text{IL} < 0.1 \text{ dB}$
Low saturation magnetization	$H_{\text{sat}} < 400 \text{ Oe}$
Good substrate lattice matching	match to GGG or like
Low pitting density	$< 10/\text{cm}^2$
Temperature range	$-20 \text{ to } +70^\circ\text{C}$
Low temperature coefficient	$d \theta'_F /dT < 0.07^\circ/\text{C}$
Low wavelength dependence	$d \theta'_F /d\lambda < 0.08^\circ/\text{nm}$
High Curie temperature	$T_c > 150^\circ\text{C}$

accordingly, the magnetization vector within the film is normal to the film surface. The strength of the magnet must be sufficient to saturate fully the domain structure of the garnet over the specified temperature range. Multi-magnet schemes have been proposed to enhance the magnetic field in the region surrounding the magnet [23, 25], although these concepts are not currently used in telecom-grade components. The direction of magnetization sets the direction of polarization rotation, whether clockwise or counterclockwise. Light transits the garnet part either parallel or anti-parallel to the magnetization direction.

To attain component-quality performance, the FR must have a low saturation magnetization H_{sat} so the permanent magnetic can be small, a low absorption, a low temperature-dependent specific rotation (defined by (3.7.38) on page 135), and a low wavelength-dependent specific rotation. Moreover, the film must closely lattice match to the substrate, over the range of room and growth temperatures, to enable crystal growth. The requisite qualities of a telecommunications-grade iron garnet used in isolators and circulators, as opposed to magneto-optic sensor applications, are listed in Table 4.4.

Yttrium iron garnet (YIG) is an early garnet material that was a suitable replacement for diamagnetic FRs of the time. With a chemical formula of $\text{Y}_3\text{Fe}_5\text{O}_{12}$, the iron content is the sole contributor to the magnetization as Y has no net magnetic moment. In relation to the component requirements, however, a YIG film must be $\sim 2.7 \text{ mm}$ to achieve a rotation of $\theta_F = 45^\circ$. Also, YIG has a large saturation magnetization ($\sim 1800 \text{ G}$), requiring a large permanent magnet.

To reduce the requisite film thickness, bismuth exchanged rare-earth iron garnets (Bi:RIG) were introduced. With the chemical formula $(\text{BiRE})_3\text{Fe}_5\text{O}_{12}$, the combined bismuth and rare-earth (RE) ions greatly enhance the specific rotation. There are, however, tradeoffs due to the exchange of (BiRE) for yt-

trium. The bismuth ion increases the lattice constant of the film; it increases the temperature dependence of the specific rotation; it increases the thermal expansion of the film; and it increases the possibility of pitting in the film. The lattice mismatch limits bismuth incorporation to about one atom in three. These deleterious effects may be somewhat compensated by the selection and concentration of the rare-earths. Addition of terbium (Tb), for example, will decrease the temperature dependence. Which rare-earth atoms are suitable depends in large part on their absorption spectra and the operating wavelength of the garnet. At 1.55 μm , Tb, gadolinium (Gd), holmium (Ho), and europium (Eu) are all suitable to varying degrees. However, at 980 nm their absorption is too high and other means, such as heavy bismuth loading and a 25 μm film [21], is all that can be expected.

To reduce the saturation magnetization, gallium and aluminum can be substituted for iron as in $(\text{BiRE})_3(\text{FeGaAl})_5\text{O}_{12}$. Introduction of Ga and/or Al, however, concurrently reduces the Curie temperature (the temperature at which the magnetization is zero) and increases the temperature dependence of the specific rotation.

Very interesting studies have been conducted to tailor the overall material properties through introduction and balancing of rare-earth ions as well as gallium and aluminum. With the optimal balance, there is a window of material compositions in which all of the iron garnet design goals listed in Table 4.4 can be achieved. A single source that presents the various tradeoffs is [21]. The combined patent work of [1, 27, 47–51, 53] provides many practical details about compositions and materials processing.

As a specific example, $(\text{Tb}_{1.69}\text{Bi}_{1.31})(\text{Fe}_{4.38}\text{Ga}_{0.42}\text{Al}_{0.20})\text{O}_{12}$ [27] exhibits $H_{\text{sat}} = 340$ Oe at $+60^\circ\text{C}$, $\theta'_F = 0.099^\circ/\mu\text{m}$ and a temperature dependence of $0.062^\circ/\text{C}$. It should be noted that the wavelength dependence of iron garnets is generally small and does not follow Biot's law.

All of the above described iron garnets follow the linear saturation curve of Fig. 3.20. These garnets are non-latching and require the presence of a permanent magnet to maintain alignment of the magnetic domains. Latching garnets, in contrast, require only a one-time poling by a strong magnet and then retain their magnetization indefinitely under normal conditions. A latching garnet is perfectly hysteretic, as illustrated by the latching curve in Fig. 3.20. As a practical matter, once the garnet is poled, proper orientation in a component is critical: reversing the part will create high transmission rather than high isolation. To make the direction of magnetization easy to identify, reference [52] reports the idea of making the AR coating on the two sides of the film different in color. A light purple can be made from a three-layer coating while a bluish purple can be made from a single-layer coating.

A linear hysteresis loop results from the film's natural tendency to fracture into domains rather than remain in a single domain. The magneto-static energy of the film is proportional to the square of the saturation magnetization: a high H_{sat} provides sufficient energy for domain break up. However, doping with Ga and/or Al decreases the saturation magnetization, increasing

Table 4.5. Select Alloy Packaging Materials

Property	Kovar ^(a)	Invar-36 ^(b)	Units
Element			
Carbon	0.020%	0.020%	
Silicon	0.20%	0.20%	
Cobalt	17%	-	
Manganese	0.30%	0.35%	
Nickel	29.0%	36.0%	
Iron	balance	balance	
Density	8.39	8.09	g/cm ³
Thermal Conductivity	17.3	10.5	W/(m·K)
Thermal Expansion ^(c)	5.85	1.30	cm/cm/K
Electrical Resistivity ^(d)	487	820	$\mu\Omega$ -mm
Melting Point	1470	1445	°C
^(a) Reported by Carpenter [12].			
^(b) Reported by Carpenter [11].			
^(c) 25°C-100°C, ^(d) 70.0°F.			

in turn the hysteresis of the material. By lowering H_{sat} to below 100 G, latching can occur. As an example, $(\text{Bi}_{0.75}\text{Eu}_{1.5}\text{Ho}_{0.75})(\text{Fe}_{4.1}\text{Ga}_{0.9})\text{O}_{12}$ has a thickness of 86 μm for 45° rotation and a saturation magnetization of 14 Oe [7]. However, the temperature dependence necessarily increases due to the high Ga concentration. The temperature dependence of the preceding film is reported as $-0.093^\circ/\text{C}$ [21].

The latching garnet is perfectly suited for an isolator placed at the output of a diode laser within the hermetic housing [22]. A semiconductor laser diode, used for signal and pump lasers, requires a miniature housing where elimination of the permanent magnet is a significant advantage. To maintain the lasing wavelength, laser diodes sub-mounts are attached to Peltier thermal-electric coolers that maintain the temperature. The latching garnet is also placed on the cooler. Under these conditions the latching garnet performs well. In passive component applications, however, where the garnet must remain stable over a wide temperature range, low temperature-dependent garnets are a better choice.

4.2.4 Packaging Alloys

Packaging materials play an integral role in component design and assembly because they house the optical core and must resist environmental interference. Industry standards such as [58] place stringent conditions on the heat and humidity a component must tolerate over its 25-year lifetime. Generally, telecommunications-grade components require hermetic sealing to protect the optical surfaces and attachment epoxies from heat and humidities. Specialty

components such as air-gap interleavers do a final frequency tuning by adjusting an inert-gas overpressure prior to sealing the package, after which that overpressure is expected to be maintained.

Table 4.5 provides information on two commonly used alloys, Kovar and Invar. These alloys are suitable for their machinability and low thermal expansion coefficient. Typically these alloys are plated with one or two mils thickness nickel and gold, in that order. The nickel promotes the gold adhesion, and the gold prevents corrosion while providing a quality surface on which to solder and weld.

4.3 Fabry-Perot and Gires-Tournois Interferometers

Two elementary components are the Fabry-Perot (FP) and Gires-Tournois (GT) interferometers. Both interferometers are resonators that store energy two partially reflecting mirrors when on resonance between. For the Fabry-Perot, leakage of the stored energy through the partial reflectors interferes constructively (destructively) with the input light to alter the transmission (reflection) properties. By comparison, on anti-resonance there is no energy stored and the transmission is minimum. The Gires-Tournois is also a resonator but with the second mirror fully reflecting. The reflection coefficient is unity but there is a frequency-dependent delay associated with resonance. The response of these two resonators is typically derived using an infinite series construct. Here, scattering and transmission matrices are first derived and then a matrix concatenation is used to reach the solution. In either formalism the final equations are the same, but the transformation matrix approach is applicable to a broader range of calculations.

The first step is to derive a scattering matrix S across an index-step boundary such as a partially reflecting mirror. The matrix relates the outputs across the boundary to the inputs (Fig. 4.3(S)):

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad (4.3.1)$$

Since the system is lossless, the scattering matrix must be unitary:

$$S^\dagger S = I \quad (4.3.2)$$

The scattering matrix elements are found from (3.5.31) on page 97. A phase reference plane is established on either side of the boundary. On the side of the reflection, the phase reference is chosen so that $\Gamma/|\Gamma| = +1$, or $2k_{z1}z = 2n\pi$. On the side of transmission, the phase reference is chosen so that $T/|T| = -j$, or $k_{z2}z = (2n - 1/2)\pi$. One can say that for a given wavelength the phase reference plane for reflection coincides with the boundary surface and the phase reference plane for transmission is set back by a quarter-wave on the far

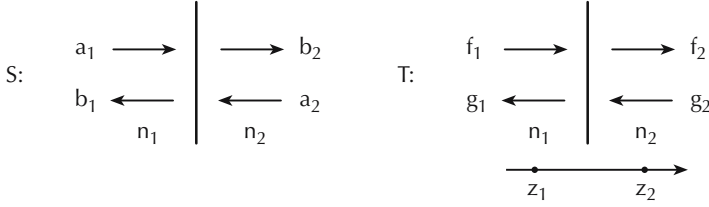


Fig. 4.3. Scattering and transmission coefficients across an index-step boundary. S) Scattering coefficients, inputs denoted by a , outputs denotes by b . T) Transmission coefficients, forward waves denoted by f , backward waves denoted by g .

side of the boundary. With these definitions and enforcing the unitary property of the scattering matrix, the scattering matrix of a partially transmissive mirror are

$$S = \begin{pmatrix} r & -jt \\ -jt & r \end{pmatrix} \quad (4.3.3)$$

where r and t are the reflection and transmission field amplitudes, respectively. The reflection and transmission coefficient related by

$$r^2 + t^2 = 1 \quad (4.3.4)$$

ensures a lossless interface. The physical interpretation of the $-j$ multiplier to the transmission coefficient is that the passage of a wave across the boundary is delayed by $\pi/2$ with respect to a wave reflected.

Next the scattering matrix is converted to a transformation matrix. The latter matrix relates the forward and backward waves at one location z_1 to the forward and backward waves at another location z_2 (Fig. 4.3(T)). A transformation matrix can be concatenated with other transmission matrices to find the response of a more complicated system. This is not the case for the scattering matrix, which is the reason for the conversion. In regard to the figure, the field amplitudes of the scattering and transformation frameworks are related in the following way:

$$\begin{aligned} f_1 &= a_1, & f_2 &= b_2, \\ g_1 &= b_1, & g_2 &= a_2. \end{aligned}$$

The field amplitudes are related to the transformation matrix as

$$\begin{pmatrix} f_1 \\ g_1 \end{pmatrix} = \begin{pmatrix} t_1 & t_2 \\ t_3 & t_4 \end{pmatrix} \begin{pmatrix} f_2 \\ g_2 \end{pmatrix} \quad (4.3.5)$$

The transformation matrix for the partially transmissive mirror is therefore

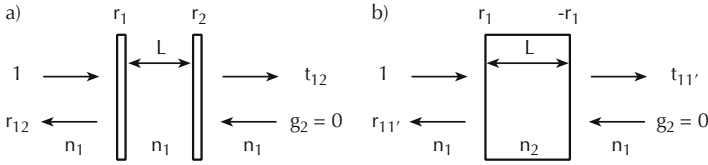


Fig. 4.4. Fabry-Perot interferometers with planar partially reflective mirrors. a) Two mirrors with reflectivities r_1 and r_2 separated by distance L . b) Model of a solid of index n with equal magnitude Fresnel reflections on either face. One reflection is the negative of the other.

$$T(z_1, z_2) = \frac{j}{t} \begin{pmatrix} 1 & -r \\ r & -1 \end{pmatrix} \quad (4.3.6)$$

Note that the determinant of (4.3.6) is $\det(T) = -jt$; reflection appears as loss to the forward-going waves.

A Fabry-Perot is modelled with two partially reflecting mirrors separated by gap L . Free propagation to position z' from z , where $z' > z$, is simply

$$T(z', z) = \begin{pmatrix} e^{-jkL} & 0 \\ 0 & e^{jkL} \end{pmatrix} \quad (4.3.7)$$

where the wavenumber k is given by the dispersion relationship $k = \omega n L / c$. Two Fabry-Perots are illustrated in Fig. 4.4, one with an air gap and the other made of a solid block.

An important simplification for finding the transmission and reflection coefficients of the FP is to null one input, in this case $g_2 = 0$ as illustrated in Fig. 4.4. The transmission and reflection amplitudes t_{12} and r_{12} , respectively, are then determined from $t_{12} = f_2/f_1$ and $r_{12} = g_1/f_1$. The left- and right-hand sides are related by the FP transformation matrix

$$T_{\text{fp}}(z_1, z_2) = -\frac{1}{t_1 t_2} \begin{pmatrix} 1 & -r_1 \\ r_1 & -1 \end{pmatrix} \begin{pmatrix} e^{-jkL} & 0 \\ 0 & e^{jkL} \end{pmatrix} \begin{pmatrix} 1 & -r_2 \\ r_2 & -1 \end{pmatrix} \quad (4.3.8)$$

Expansion of (4.3.8) and identification of terms leads to

$$t_{12} = -\frac{t_1 t_2}{e^{-jkL} - r_1 r_2 e^{jkL}} \quad (4.3.9a)$$

$$r_{12} = +\frac{r_1 e^{-jkL} - r_2 e^{jkL}}{e^{-jkL} - r_1 r_2 e^{jkL}} \quad (4.3.9b)$$

The transmitted and reflected powers are the squared-magnitudes of the preceding expressions.

4.3.1 Fabry-Perot Response

The general transmission and reflection coefficients for a Fabry-Perot are given by (4.3.9). Two special cases are considered here.

For the first special case, $r_1 = r_2 = r$. The power coefficients are

$$|t_{11}|^2 = \frac{(1 - R)^2}{1 + R^2 - 2R \cos(2kL)} \quad (4.3.10a)$$

$$|r_{11}|^2 = \frac{2R(1 - \cos(2kL))}{1 + R^2 - 2R \cos(2kL)} \quad (4.3.10b)$$

where $R = r^2$, $T = t^2$, $R + T = 1$, and coefficient T and transformation matrix T are distinguished by context.

For the second special case, illustrated in Fig. 4.4(b), the reflection coefficients on the two boundaries have opposite sign: $r_2 = -r_1$. The power transmitted is

$$|t_{11'}|^2 = \frac{(1 - R)^2}{1 + R^2 + 2R \cos(2kL)} \quad (4.3.11)$$

where the effect of the negative second reflection coefficient is to shift the transmission (and reflection) spectrum by a half period as compared with (4.3.10). An exemplar transmission spectrum is illustrated in Fig. 4.5, where the FP interferometer is a solid block.

The transmission spectrum of a Fabry-Perot is periodic and depends on the phase accumulation in the resonator. The transmission is related to the resonator phase as

$$|t_{11}|^2 = \begin{cases} 1 & \text{resonance: } 2kL = 2n\pi, \\ \frac{(1 - R)^2}{(1 + R)^2} & \text{anti-resonance: } 2kL = (2n + 1)\pi \end{cases} \quad (4.3.12)$$

Of course, if there is loss in the system then the transmission maximum is less than unity. Also, the conditions for resonance and anti-resonance for the solid-block FP are switched from those in (4.3.12).

In general, a modulation depth MD is defined as

$$\text{MD} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (4.3.13)$$

where I is intensity. Substitution of (4.3.12) into (4.3.13) gives the modulation depth for a Fabry-Perot interferometer:

$$\text{MD}_{\text{FP}} = \frac{2R}{1 + R^2} \quad (4.3.14)$$

The range of modulation depth is from zero to unity and is governed solely by the mirror reflectivity R .

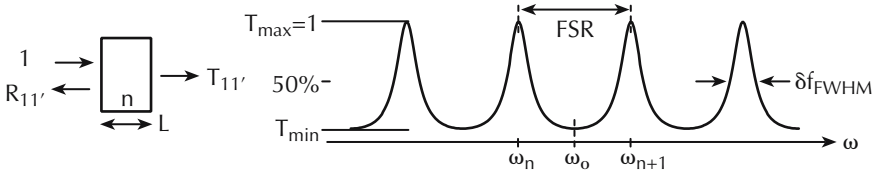


Fig. 4.5. Solid Fabry-Perot interferometer; the second reflection coefficient is the opposite sign of the first. For a unit input there is frequency-dependent transmission and reflection. Left: exemplar transmission spectrum for solid FP. A comb of transmission peaks exists in frequency, spaced by the free-spectral range (FSR) of the cavity. The modulation depth is governed solely by the boundary reflectivity.

Free-Spectral Range and Group Index

As illustrated in Fig. 4.5 the transmission (and reflection) spectrum is periodic. The periodic structure is dictated by the resonator phase $\phi = 2kL$. The free-spectral range (FSR) of the interferometer is defined as the frequency shift required to offset the spectrum by one full period. That is, the change in phase must be $\Delta\phi = 2\pi$. Expanding the phase to show radial frequency ω ,

$$\phi = 2\frac{\omega}{c} n(\omega)L \quad (4.3.15)$$

the phase difference is related to the frequency difference by

$$\begin{aligned} \Delta\phi &= \frac{2L}{c} (\omega_2 n(\omega_2) - \omega_1 n(\omega_1)) \\ &= \frac{2L}{c} \Delta(\omega n(\omega)) \end{aligned} \quad (4.3.16)$$

The difference on the right-hand right may be written

$$\begin{aligned} \Delta(\omega n(\omega)) &= \frac{d(\omega n(\omega))}{d\omega} \Delta\omega \\ &= \left(n + \omega \frac{dn}{d\omega} \right) \Delta\omega \end{aligned} \quad (4.3.17)$$

After the more customary conversion to wavelength from frequency, the group index is defined as

$$n_g = n - \lambda \frac{dn}{d\lambda} \quad (4.3.18)$$

Substitution of (4.3.17-4.3.18) into (4.3.16) and conversion to cycle frequency from radial frequency yields

$$\Delta\phi = 2\pi \frac{2n_g L}{c} \Delta f \quad (4.3.19)$$

The free-spectral range $\text{FSR} = \Delta f$ is defined for $\Delta\phi = 2\pi$, or

$$\text{FSR} = \frac{c}{2n_g L} \quad (4.3.20)$$

That the FSR is defined by the group index is a statement that it is the round-trip time of the optical energy, and not phase, that matters. The difference between phase and group index is critically important when building precision instruments such as optical clocks, where the phase and group velocities in a resonant cavity must be locked [33, 37]. Optical fibers also exhibit a difference between phase and group indices, a difference that varies from unspun to spun fibers, across vintages, and across fiber types [36].

Resonant Bandwidth

Also illustrated in Fig. 4.5 is the full-width half-maximum δf_{FWHM} of the transmission spectrum. For a symmetric lossless Fabry-Perot, as long as the mirror reflectivity exceeds

$$R \geq \frac{\sqrt{2} - 1}{\sqrt{2} + 1} \quad (4.3.21)$$

then the transmission will fall below 50%. Setting (4.3.10) to 50% and solving for the resonant bandwidth δf_{FWHM} yields

$$\delta f_{\text{FWHM}} = \frac{2}{\pi} \text{FSR} \sin^{-1} \frac{1 - R}{2\sqrt{R}} \quad (4.3.22)$$

In the highly resonant case, where $\delta\phi_{\text{FWHM}} \ll \pi$, then

$$\delta f_{\text{FWHM}} \simeq \frac{1 - R}{\pi\sqrt{R}} \text{FSR} \quad (4.3.23)$$

Fabry-Perot interferometers are classically used as narrowband filters because the transmission on resonance is unity and a high mirror reflectivity makes the resonant bandwidth narrow. For practical applications, FP filters are not good wavelength-division multiplexing filters unless they are cascaded to make a multi-pole filter with higher roll off and flatter transmission profiles. The FP resonator is sometimes used, however, as a frequency standard on which to lock a transmission laser frequency.

Tolerancing the Resonance Frequency

The following tolerance analysis illustrates the necessary specifications for a telecommunications-grade Fabry-Perot interferometer. Consider a FP etalon made of BK7 glass that nominally aligns its transmission spectrum to a DWDM grid of $C = 100$ GHz in the C-band. The frequency-dependent argument ($2kL$) of the transmission expression (4.3.11) nominally aligns to the DWDM comb of frequencies f_n :

$$2\pi f_n \frac{2n_g L}{c} = (2n + 1)\pi \quad (4.3.24)$$

Nominally $\text{FSR} = C$, so

$$f_n = \left(n + \frac{1}{2}\right) C \quad (4.3.25)$$

Due to errors, however, the free-spectral range may have an error. That error requires an offset from f_n to reach the nominal frequency comb f_n . That is,

$$2\pi \frac{f_n \pm \delta f}{\text{FSR} \pm \delta \text{FSR}} = 2\pi \frac{f_n}{C} \quad (4.3.26)$$

Substitution of the nominal grid (4.3.25) into (4.3.26) gives the error proportionality

$$\frac{\delta \text{FSR}}{C} = \frac{\delta f}{f_n} \quad (4.3.27)$$

Using the specification in §4.1 of $\Delta f_n = \pm 2.5$ GHz and a center frequency of $f_n = 194.1$ THz, the allowable error on the free-spectral range is

$$\left| \frac{\delta \text{FSR}}{C} \right| \leq 13 \text{ ppm} \quad (4.3.28)$$

This error tolerance is almost impossibly small.

To relax the tolerance one can separate the frequency-location error from the FSR error. That is, rather than insisting the FSR exactly match the grid exactly, the FSR is allowed to walkoff the grid over some bandwidth. This is illustrated in Fig. 4.2(a). The calculation made in (4.1.3) gives a 1250 ppm error tolerance to the FSR.

Separate from the FSR accuracy, the resonant peaks must still align to the grid. At center frequency the number of modes in the cavity is $n \simeq 1941$. The length of BK7 to realize the cavity, using round numbers, is $L = 1.000$ mm. As there are 1941 modes in the cavity, the mode length is $l = 0.515 \mu\text{m}$. A change in cavity length by $0.515 \mu\text{m}$ times an integral number adds or removes modes to the cavity, but a non-integral length change shifts the frequency location of the resonant peaks. Using the frequency-error tolerance $\Delta f_n = \pm 2.5$ GHz, any change in cavity index-length product must be within

$$\left| \frac{\delta(n_g L)}{n} \right| \leq \frac{1.5}{1941} \times 2.5\% \quad (4.3.29)$$

or

$$|\delta(n_g l)| \leq 0.020 \mu\text{m} \quad (4.3.30)$$

Expansion of the index-length product to account for temperature dependence gives

$$\delta(n_g l) = n_g \Delta T \frac{dl}{dT} + l \Delta T \frac{dn_g}{dT} \quad (4.3.31)$$

Using the thermal expansion and thermal optic coefficients for BK7 found in Table 4.1 and accounting for a $\pm 50^\circ\text{C}$ temperature swing, the error budget due to temperature change is

$$n_g \Delta T \frac{dl}{dT} + l \Delta T \frac{dn_g}{dT} \simeq 0.683 \text{ } \mu\text{m} \quad (4.3.32)$$

In comparison with (4.3.30) one can clearly see that a bulk Fabry-Perot cavity does not meet the tolerance requirements for a telecommunications component. These cavities require active temperature control to meet the necessary specifications. In practice, the cavities are made with an air gap sealed hermetically into a package, and the construction materials are selected to minimize temperature-dependent expansion.

4.3.2 Gires-Tournois Response

A Gires-Tournois interferometer replaces the second mirror with a perfect reflector. In this case $t_2 = 0$ and $r_2 = 1$. Therefore $t_{12} = 0$ and the reflection coefficient (4.3.9) simplifies to

$$r_{\text{GT}} = \frac{r e^{-jkL} - e^{jkL}}{e^{-jkL} - r e^{jkL}} \quad (4.3.33)$$

The magnitude is unity for all frequencies: $|r_{\text{GT}}| = 1$. The GT interferometer is an all-reflection filter, but the phase is frequency dependent. Making the denominator read-valued, the reflection coefficient is

$$r_{\text{GT}} = \frac{(r e^{jkL} - e^{-jkL})^2}{1 + r^2 - 2r \cos(2kL)} \quad (4.3.34)$$

Denoting the phase of the reflection as $\Phi_{\text{GT}} = \angle r_{\text{GT}}$, the GT phase is

$$\Phi_{\text{GT}} = -2 \tan^{-1} \left(\frac{1+r}{1-r} \tan(kL) \right) \quad (4.3.35)$$

The phase reference plane from which Φ_{GT} is measured is located on the interface of the leading mirror.

In the limiting cases, when $r = 0$ the phase propagates normally over length $2L$ while when $r = 1$ the reflection is negative one. Between these two limits there is variation of the phase. The effect of the resonant cavity is more clearly shown through the group delay τ_g , where

$$\tau_g = \frac{d\Phi_{\text{GT}}}{d\omega} \quad (4.3.36)$$

or, by expansion

$$\tau_g = -\frac{2n_g L}{c} \frac{2h}{(1+h^2) + (1-h^2) \cos(2kL)} \quad (4.3.37)$$

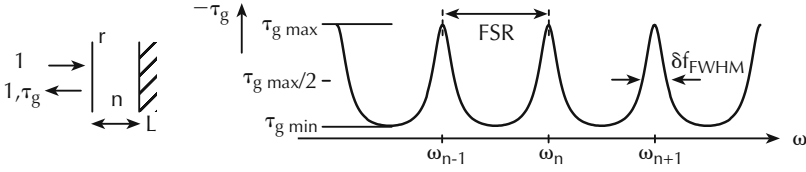


Fig. 4.6. Solid Gires-Tournois interferometer (GTI) with gap length L and refractive index n . The GTI has 100% reflection at all frequencies, but there is frequency-dependent delay of the reflection. The delay spectrum is similar to the transmission spectrum of a Fabry-Perot interferometer in that it is periodic. The FSR of the group-delay comb is dictated by the gap length and refractive index, and the maximum delay is dictated by the FSR and the reflectivity r of the partial reflector.

where

$$h = \left(\frac{1+r}{1-r} \right) \quad (4.3.38)$$

(compare (4.3.10) on page 157). Figure 4.6 illustrates an exemplar group-delay spectrum. On resonance, $2kL = 2n\pi$ and the group delay is maximum because the resonance wavelength is an integral multiple of the cavity length, allowing the storage of energy. On anti-resonance the group delay is at a minimum as there is no energy stored. Indeed the maximum and minimum group delays are

$$\tau_{g,\max} = -\frac{1}{\text{FSR}} \left(\frac{1+r}{1-r} \right) \quad 2kL = 2n\pi, \quad (4.3.39a)$$

$$\tau_{g,\min} = -\frac{1}{\text{FSR}} \left(\frac{1-r}{1+r} \right) \quad 2kL = (2n+1)\pi \quad (4.3.39b)$$

One can see that the group delay is related to the inverse of the free-spectral range and the leading mirror partial reflectivity. The inverse free-spectral range is a unit of delay for the cavity. As with the Fabry-Perot interferometer, the free-spectral range is related to cavity parameters by (4.3.20).

Conservation of Delay – Bandwidth Product

The Gires-Tournois interferometer group-delay response obeys a conservation rule in that the product of peak delay and resonance bandwidth is a constant for fixed r . Following the Fabry-Perot example, the resonance bandwidth δf_{FWHM} of the group delay can be calculated at one-half the maximum level. Setting (4.3.37) equal to $\tau_{g,\max}/2$, one finds that

$$\delta f_{\text{FWHM}} = \frac{\text{FSR}}{\pi\sqrt{h^2 - 1}} \quad (4.3.40)$$

The peak delay/bandwidth product is therefore

$$\begin{aligned}
\tau_{g,\max} \times \delta f_{\text{FWHM}} &= -\frac{h}{\pi\sqrt{h^2 - 1}} \\
&= -\frac{1 + r}{2\pi\sqrt{r}}
\end{aligned} \tag{4.3.41}$$

The partial reflectivity of the front mirror alone governs the peak delay/bandwidth product. The FSR of the cavity plays no role but rather is scaled out. Higher peak delays are achieved with higher reflectivities (cf. (4.3.39)), but the bandwidth (4.3.40) narrows by a commensurate amount. The partial reflectivity of the leading mirror is the only degree of freedom for this simple GT, so independent control of bandwidth and peak delay is not possible.

4.4 Temperature Dependence of Select Birefringent Crystals

There is little data available on the temperature and wavelength dependence of the refractive and group indices for components-grade birefringent crystals in the 1520–1610 nm wavelength region. Yet birefringent components such as interleavers require a detailed specification of these properties. Moreover, for components such as the off-axis compound crystals, detailed in §4.5, the design requires separate determination of the ordinary and extraordinary indices, not just their difference.

Measurement of the refractive and group indices requires different techniques. The method presented below is suitable to ascertain the temperature- and course wavelength-dependence of the group indices along the two birefringent axes. The group index accounts for first-order wavelength dependence and determines the group velocity (cf. (4.3.18)). The refractive index is defined at a specific wavelength and determines the refraction properties. The measurement technique here involves the response to a variation in the input wavelength, so it is the group index that is measured.

4.4.1 Experimental Setup

Crystal samples of YVO_4 , LiNbO_3 , $\alpha\text{-BBO}$, and calcite were fabricated into Fabry-Perot resonators. The samples were cut as waveplates, with the extraordinary axis lying in the polished face of the crystal, perpendicular to the light. The sample thicknesses were nominally 1.000 mm long, and each was coated with a single thin-film layer on either face to bring the Fresnel reflectivity to approximately 50%. The parallelism specification called for less than 5 arcsec of wedge.

Figure 4.7(a) illustrates the experimental setup. The Fabry-Perot etalon response of each sample was measured with a super-luminescent diode (SLD) having 100 nm bandwidth and less than 0.25 dB ripple, and an optical spectrum analyzer with 0.05 nm resolution. The measurement band was 1515–1575 nm for all samples. The optical spectrum analyzer (OSA) was set to

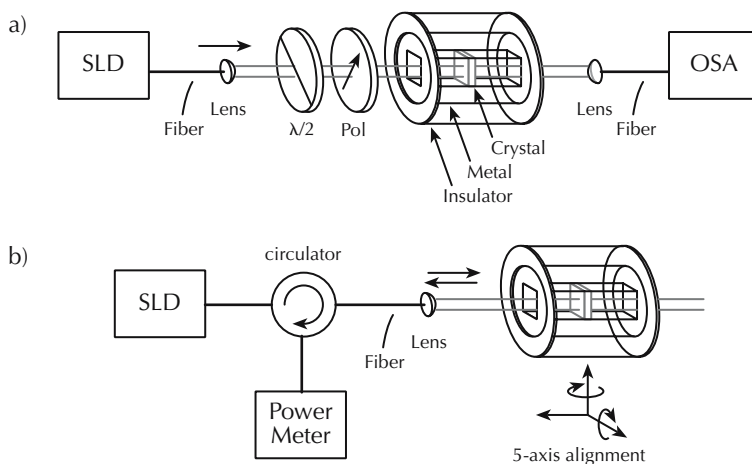


Fig. 4.7. Illustration of experimental setup to measure temperature dependence of group index. a) A superluminescent diode (SLD) provides broadband, ripple-free, polarized light in the 1515–1575 nm band. The polarizer is aligned to the extraordinary or ordinary axis of the sample crystal etalon, the latter being loaded in a heating cell. The light is collected by an optical spectrum analyzer. b) To ensure that the crystal sample is perpendicular to the light, back-reflection is collected through a circulator and the 5-axis alignment stage is adjusted to maximize this reflection.

maximum sensitivity and eight averages per scan. It is noted that the spectra were stable and averaging created little change.

To couple light through a sample, a single-mode fiber was routed from the SLD to a collimator, the collimator expanded the beam to approximately 0.75 mm in diameter, and the collimated beam transited the sample. The beam was refocused through a second lens to a single-mode fiber which was routed to the OSA. Since the samples are birefringent, an in-line Polarcor polarizer was placed before the sample. Rotation of the polarizer selected either the ordinary or extraordinary axis of the sample, or a mixture of the two. For the measurements, the polarizer was always rotated for maximum extinction of one axis or another. No measurement was made as to the extinction ratio, but visual inspection of the spectrum on the OSA indicated that the extinction was better than 10 dB. Since the SLD generates a linearly polarized white light, a half-wave waveplate was located before the polarizer and independently rotated to maximize transmission through the polarizer.

For each experiment, the sample crystal was loaded into a small brass fixture that supplied resistive heating. The aperture was rectangular and the position of one wall of the aperture was adjustable by a screw. In order to allow the crystal to expand physically the screw was lightly tightened. The brass fixture was insulated with delrin and teflon. A closed-loop temperature controller controlled the heating of the fixture to within $\pm 0.1^\circ\text{C}$. The samples

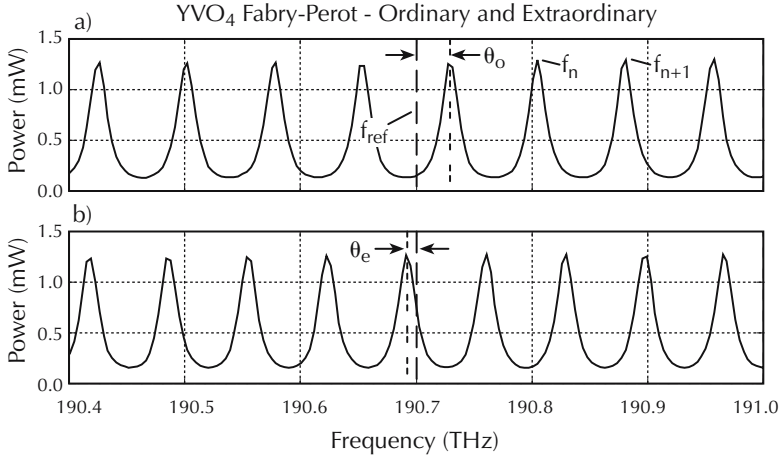


Fig. 4.8. Measured Fabry-Perot etalon response of YVO_4 crystal sample. a) Transmission response along the ordinary axis. b) Transmission response along the extraordinary axis. Note the free-spectral range along the two axes is different.

were taken from 25°C to 100°C in 5°C increments. Five minutes were allowed for thermal stabilization at each temperature.

A critical attribute of this experiment is the optical path through the sample. When the sample is canted the path length increases, but the increase is not an easily measurable quantity. To guarantee that the crystal was positioned perpendicular to the beam, a preliminary alignment was done. Figure 4.7(b) illustrates the setup with the waveplate or polarizer removed, and an optical circulator inserted between the SLD and sample. The staging that holds the sample was adjusted to maximize the back-reflection. Once maximized the sample was considered aligned. This measurement was performed before and after was temperature cycled to assess the degree of position change. It is believed that position shifts did not effect the data to within the present level of accuracy.

Figure 4.8 shows the measured transmission response of a YVO_4 etalon along the ordinary and extraordinary axes over a narrow bandwidth. In both spectra there is a comb of resonant peaks and the period of the peaks differs for the two axes. As YVO_4 is a positive uniaxial crystal, the free-spectral range of the extraordinary axis is narrower than that of the ordinary axes. Each peak, for either spectra, corresponds to a resonant mode. As the frequency is increased, more modes are added to the cavity; this is indicated at frequencies f_n and f_{n+1} in Fig. 4.8(a). Additionally, a reference frequency f_{ref} is defined to measure spectral shift with temperature. The choice of f_{ref} is arbitrary but remains constant throughout. A spectral phase θ is defined between resonant frequency f_n and the reference frequency as

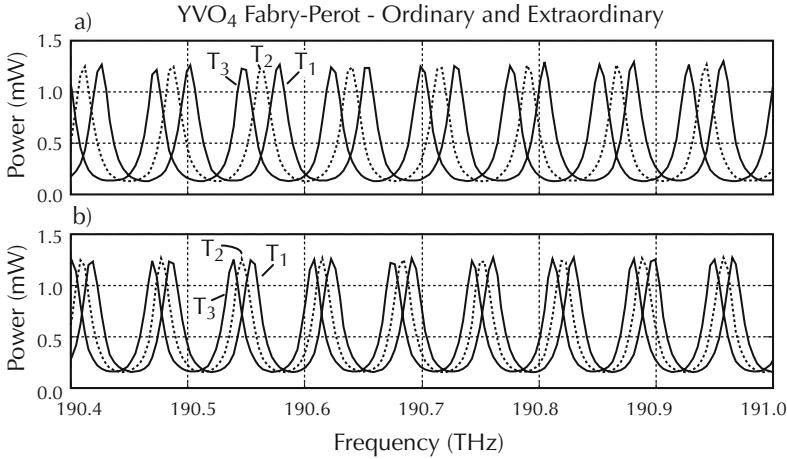


Fig. 4.9. Measured Fabry-Perot etalon response of YVO₄ crystal sample over increasing temperature, 10°C increments. The comb of resonant frequencies shifts to lower frequency, indicating an increase in the index-length product of the sample. a) Ordinary axis. b) Extraordinary axis. Note the temperature-dependent shift differs between the two axes.

$$\theta_n = 2\pi \frac{f_n - f_{\text{ref}}}{\text{FSR}} \quad (4.4.1)$$

This spectral phase will be used to extract the thermal-optic coefficient in the following.

Figure 4.9 shows the measured transmission response of the YVO₄ crystal as the temperature is increased from T_1 to T_3 . For both axes the comb of resonant peaks shifts to lower frequencies. When the cavity expands or when the refractive index increases with increased temperature, a lower frequency is required to maintain the same mode order n that corresponds to resonant frequency f_n . The figure shows that for both the ordinary and extraordinary axes the product of the index and length increases with temperature, detuning the resonance comb to lower frequencies.

What is also evident in Fig. 4.9 is that the index-length product changes by different degrees for the ordinary and extraordinary axes. This is typical of all birefringent crystals. One effect of this birefringent temperature dependency is that the birefringent phase changes with temperature. This is a distinct disadvantage for birefringent filters that need to remain locked to the DWDM grid.

4.4.2 Quadratic Temperature-Dependence Model

The transmission expression (4.3.11) on page 157 is a function of phase $\phi = 2kL$. Expansion of the phase with temperature to second order gives

$$\phi(\Delta T) \simeq \phi_o + \phi' \Delta T + \frac{1}{2} \phi'' (\Delta T)^2 \quad (4.4.2)$$

Similarly, the equivalent expression $2kL$ is expanded to second order as

$$\frac{2\omega}{c} n_g L \simeq \frac{2\omega}{c} \left((n_g L) + \frac{d(n_g L)}{dT} \Delta T + \frac{1}{2} \frac{d^2(n_g L)}{dT^2} (\Delta T)^2 \right) \quad (4.4.3)$$

In light of these expansions, the temperature-dependent phase $\phi(\Delta T)$ can be approximated to second order as

$$\phi(\Delta T) \simeq \frac{2\pi f}{\text{FSR}} \left(1 + K^{(1)} \Delta T + \frac{1}{2} K^{(2)} (\Delta T)^2 \right) \quad (4.4.4)$$

where the linear and quadratic temperature coefficients are defined as

$$K^{(1)} = \frac{1}{n_g L} \frac{d(n_g L)}{dT} \quad (4.4.5a)$$

$$K^{(2)} = \frac{1}{n_g L} \frac{d^2(n_g L)}{dT^2} \quad (4.4.5b)$$

Extension of the temperature dependence to second order is necessary because the first-order term can be identically cancelled by proper selection of complementary crystals, as detailed in §4.4.5. However, the quadratic term cannot be so cancelled, leaving a residual error that may be significant given the requirements for telecommunications-grade components.

4.4.3 Association of Resonant Peak Shift With Temperature Coefficients

The frequency locations of the etalon resonances shift with temperature due to changes in the $n_g L$ product. At the n^{th} resonance the optical phase is $\phi = 2n\pi$. When the temperature changes the peak frequencies shift to maintain this resonant condition. One can write the resonance frequency on the n^{th} mode for two different temperatures as

$$f_n(T_1) = \frac{nc}{2n_g L} \quad (4.4.6a)$$

$$f_n(T_2) = \frac{nc}{2(n_g L + \delta(n_g L))} \quad (4.4.6b)$$

The change $\delta(n_g L)$ can then be expressed as a function of resonant peak frequencies:

$$\frac{\delta(n_g L)}{n_g L} = \frac{f_n(T_1)}{f_n(T_2)} - 1 \quad (4.4.7)$$

Using the quadratic expansion (4.4.3) and temperature coefficient definitions (4.4.5), the ratio of resonant frequencies is related to the temperature dependence as

$$\frac{f_n(T)}{f_n(T + \Delta T)} - 1 = K^{(1)}\Delta T + \frac{1}{2} K^{(2)}(\Delta T)^2 \quad (4.4.8)$$

If the frequency peaks can unambiguously be determined from the data, then estimates of $K^{(1)}$ and $K^{(2)}$ can be determined from (4.4.8).

A problem with the determination of $f_n(T)$ is that the resonant peak locations of the etalon do not coincide with the frequencies at which the OSA measures the transmitted power. This can be observed by close inspection of Fig. 4.8. The periodic nature of the spectrum, however, can be used to advantage by applying a Fourier analysis. Using a Fourier transform to extract the spectral phase of a mode provides a certainty enhancement of the phase value. The phase difference between two temperatures determines $f_n(T + \Delta T)$ via

$$f_n(T + \Delta T) = \frac{\text{FSR}}{2\pi} (\theta_n(T + \Delta T) - \theta_n(T)) + f_n(T) \quad (4.4.9)$$

as long as $\text{FSR}(T + \Delta T) \simeq \text{FSR}(T)$.

There is a certain tradeoff for the Fourier analysis. On one hand the spectral phase accuracy is improved by taking the Fourier transform over a large number of periods. On the other hand, association of the resultant spectral phase to a particular mode is increasingly less certain as the transform window includes more modes. The presence of $f_n(T)$ on the right-hand side of (4.4.9) requires a certainty of the mode to which the spectral phase is associated. The wavelength range for these measurements is 1515–1575 nm, or about 3.9% spectral coverage. If the Fourier transform were taken over, the entire data set there would be a $\pm 2\%$ error in certainty of f_n .

The tradeoff used here is to partition the data set into subsets having 128 points, or about 3.8 nm, each. The certainty of f_n is then $\pm 0.12\%$. The spectral phase of the fundamental tone for each partition was extracted from its Fourier transform, and the sequence of phases for a temperature ramp was inserted into (4.4.9). This was done for each data partition and the results were compared. Overall, there was no discernible trend from partition to partition, although small differences were evident.

For each partition and at each temperature the free-spectral range was estimated via

$$\text{FSR}^{(\text{est})}(T) = \frac{f_k(T) - f_j(T)}{N_p - 1} \quad (4.4.10)$$

where the approximate peak frequencies $f_{j,k}$, were taken at either end of the partition, and N_p is the number of peaks in a partition. The FSR estimates over all temperatures were compared. There was no trend of the FSR estimates, which is reasonable since only one or two modes were added over the total temperature range, depending on the material.

4.4.4 Group Index and Thermal-Optic Coefficients

Table 4.6 tabulates the results of the data analysis, and Figs. 4.10-4.11 show the determination of $\delta(n_g L)$ over temperature for the four crystals. The co-

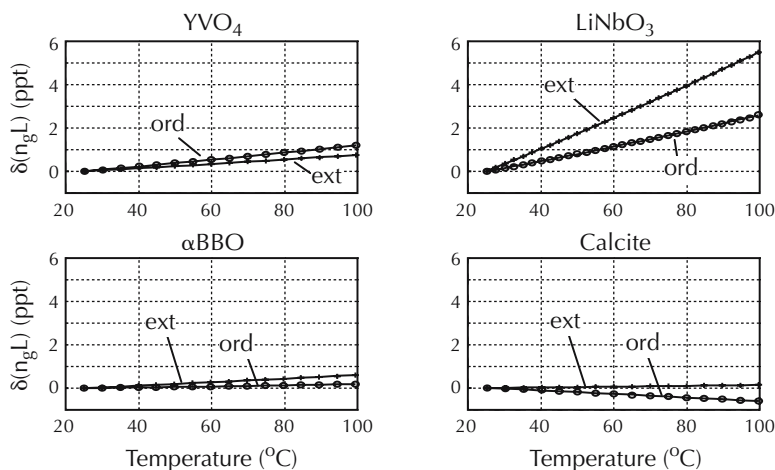


Fig. 4.10. Measured and quadratically fit temperature-dependent refractive index change for select crystals. The ordinary and extraordinary axes are separately measured. LiNbO₃ has a much larger thermal-optic effect than the other crystals. Calcite has a negative thermal-optic effect on the ordinary axis.

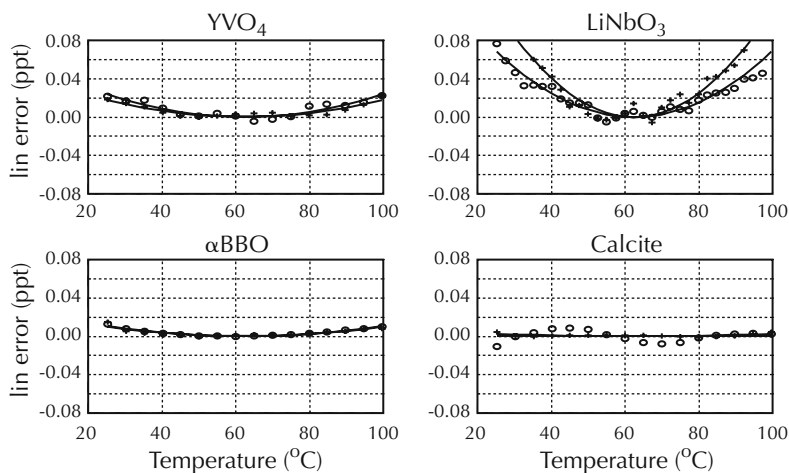


Fig. 4.11. Measured and estimated quadratic residual change in index, with the linear term removed. LiNbO₃ has a large quadratic shift. Calcite shows a spurious undulation along the ordinary axis.

Table 4.6. Group Index and Thermal-Optic Coefficients

Property	YVO ₄	LiNbO ₃	α -BBO	Calcite	Units
$n_{g,e}$	2.1918	2.1834	1.5296	1.4832	
$n_{g,o}$	1.9770	2.2656	1.6593	1.6593	
Δn_g	+0.2148	-0.0822	-0.1297	-0.1761	
$K_e^{(1)}$	4.9077	32.228	5.2729	1.3008	$\times 10^{-6}/K$
$K_o^{(1)}$	8.1226	15.351	1.7364	-4.9372	$\times 10^{-6}/K$
$K_{\Delta n}^{(1)}$	-24.682	-432.93	-39.971	-57.477	$\times 10^{-6}/K$
$K_e^{(2)}$	0.0112	0.0707	0.0069	0.0035	$\times 10^{-6}/K$
$K_o^{(2)}$	0.0149	0.0424	0.0061	-0.0009	$\times 10^{-6}/K$
$K_{\Delta n}^{(2)}$	-0.0224	-0.7093	-0.0033	-0.0380	$\times 10^{-6}/K$

Measurement range: 1515–1575 nm, 25°C–100°C.

$\Delta T = 0$ at $T = 62.5^\circ\text{C}$.

efficients were generated for $\Delta T = 0$ at $T_o = 62.5^\circ\text{C}$. Figure 4.10 shows the change of the group-index-length product as a function of temperature, the room-temperature group index being subtracted from the data to better show the variation. The solid lines are the quadratic curve fit. LiNbO₃ is clearly the most temperature-dependent material of the four. Calcite has a negative thermal-optic coefficient for its ordinary axis. Whether the negative coefficient comes from change in group index or length cannot be distinguished from this experiment.

Figures 4.11 shows the same data as in Fig. 4.10 but with the best-fit linear slope removed. This highlights the deviation from linear temperature dependence. Again, LiNbO₃ has the largest quadratic effect. Calcite shows an irregular undulation for the ordinary axis. The quadratic terms for YVO₄ and α -BBO are small but may be significant.

In some cases the thermal-optic coefficient K for the birefringence needs to be known, rather than the coefficient for either index. The birefringent coefficient for linear and quadratic terms is derived from

$$K_{\Delta n_g}^{(1,2)} = \frac{1}{\Delta n_g} \left(n_{g,e} K_e^{(1,2)} - n_{g,o} K_o^{(1,2)} \right) \quad (4.4.11)$$

Passive compensation of birefringent phase requires this combined coefficient.

4.4.5 Passive Temperature Compensation

In applications of birefringent filters, birefringent crystals cut as waveplates are used to generate a periodic filter response. The periodicity is generated by

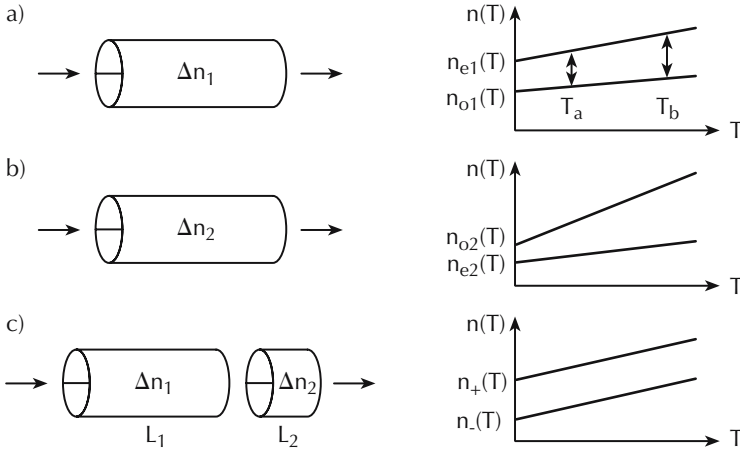


Fig. 4.12. Passive temperature compensation of birefringent phase. a) A birefringent crystal cut as a waveplate, line in face indicates extraordinary axis. Temperature dependence of refractive index for ordinary and extraordinary indices differs, as indicated on the right graph. b) Another birefringent crystal having different thermal-optic coefficients from a). c) Length ratio of two crystals is designed so that temperature dependencies of n_+ and n_- are matched. The birefringent phase of the cascade is temperature-compensated to first order.

the frequency-dependent polarization transformation of the waveplate rather than by the Fabry-Perot effect discussed in the preceding sections. Like the Fabry-Perot interferometers, the comb spectrum of a waveplate shifts in frequency due to the temperature dependence of the birefringence.

The birefringent phase φ is defined through the difference between extraordinary and ordinary indices as

$$\varphi = \frac{\omega}{c} (n_e - n_o) L \quad (4.4.12)$$

A change in temperature changes the phase, to first order, as

$$\Delta\varphi = \frac{\omega}{c} (n_e K_e - n_o K_o) L (\Delta T) \quad (4.4.13)$$

Consider for example a YVO_4 waveplate that is 10 mm long, and the temperature is changed by 75°C . Using the entries from Table 4.6 one finds the birefringent phase changes by $\Delta\varphi \simeq -2.6\lambda$, or over two waves. The comb spectrum shifts by more than two beats. For a LiNbO_3 waveplate of equal length the birefringent phase shifts $\Delta\varphi \simeq 17.1\lambda$.

Figure 4.12(a,b) illustrates this reason for the change. The ordinary and extraordinary indices generally have different thermal-optic coefficients, which in turn causes a temperature-dependent shift in birefringent phase.

To correct for first-order temperature dependence, two complementary crystals can be cascaded. As illustrated in Fig. 4.12(c), given the correct length

ratio the temperature dependence on one polarization axis can be adjusted to equal the temperature dependence on the orthogonal polarization axis. The term “polarization axis” is used here because the extraordinary axes of the two crystals may be aligned or crossed, depending on the materials, so the designation of extraordinary axes for the combination loses its meaning.

In practice, an athermal crystal pair is designed for a specific free-spectral range. The length ratio of the pair sets the athermalization and the length total sets the free-spectral range. The free-spectral range for a waveplate is determined from the differential group delay τ , where τ is defined as

$$\tau = \frac{\partial \varphi}{\partial \omega} = \frac{\Delta n_g L}{c} \quad (4.4.14)$$

and the waveplate free-spectral range is $\text{FSR} = 1/\tau$. Note that the FSR of a waveplate is different from a Fabry-Perot (4.3.20): the waveplate FSR is governed by its group-index difference and crystal length (see (3.6.35) on page 121); the Fabry-Perot FSR is governed by the group index and twice the cavity length. The combination of two crystals gives a total differential group delay of

$$\tau_s = \tau_1 + \tau_2 \quad (4.4.15a)$$

$$= (\Delta n_{g,1} L_1 \pm \Delta n_{g,2} L_2) / c \quad (4.4.15b)$$

where the $+$ and $-$ signs refer to parallel and perpendicular alignment, respectively, of the extraordinary axes of the two crystals.

The temperature dependence of the birefringent phase to first order is

$$\frac{\partial \varphi}{\partial T} = \frac{\omega}{c} \Delta n_g L K_{\Delta n_g}^{(1)} \quad (4.4.16)$$

Stripping unnecessary sub- and superscripts, the combined temperature dependence is cancelled when

$$\frac{\omega}{c} (\Delta n_1 L_1 K_1 \pm \Delta n_2 L_2 K_2) = 0 \quad (4.4.17)$$

where the $\pm$ sign has the same meaning as for (4.4.15b). Combining (4.4.15b) and (4.4.17), the length ratio is

$$L_2/L_1 = \mp \frac{\Delta n_1 K_1}{\Delta n_2 K_2} \quad (4.4.18)$$

and the length of the first crystal is

$$L_1 = \frac{c \tau_s}{\Delta n_1 \pm \Delta n_2 (L_2/L_1)} \quad (4.4.19)$$

The one necessary variable to attain a solution for any pair of crystals is the alignment or crossing of the extraordinary axes. Table 4.6 shows that

Table 4.7. Passive Temperature Compensation Combinations

Combination	L_1 (mm)	L_2 (mm)	$L_1 + L_2$ (mm)
$\text{YVO}_4 \parallel \text{LiNbO}_3$	14.801	2.205	17.005
$\text{YVO}_4 \parallel \alpha\text{-BBO}$	36.488	37.315	73.803
$\text{YVO}_4 \parallel \text{Calcite}$	24.461	12.812	37.273
$\alpha\text{-BBO} \perp \text{LiNbO}_3$	25.465	3.710	29.175
$\text{Calcite} \perp \text{LiNbO}_3$	19.630	5.583	25.213
$\alpha\text{-BBO} \perp \text{Calcite}$	75.890	38.870	114.761

$L_1 + L_2$ generates $\tau_s = \pm 10.0$ ps.

$\parallel$ and $\perp$ indicate aligned or crossed extraordinary axes, respectively.

the signs for the birefringent thermal-optic coefficients are the same for all crystals. However, YVO_4 is positive uniaxial and the rest are negative uniaxial. Referring to (4.4.17), if the birefringence of a crystal pair has the same sign the extraordinary axes must be crossed; otherwise they are aligned.

Table 4.7 tabulates the crystal lengths required for all six combinations of crystal pairs such that the pair generates $\tau_s = \pm 10$ ps. Clearly the YVO_4 and LiNbO_3 combination yields the shortest total length.

The residual phase error is calculated from the quadratic deviation. Similar to 4.4.13, the quadratic error is

$$\Delta\varphi = \frac{\omega}{2c} \left(\Delta n_1 L_1 K_1^{(2)} \pm \Delta n_2 L_2 K_2^{(2)} \right) (\Delta T)^2 \quad (4.4.20)$$

For the $\text{YVO}_4\text{-LiNbO}_3$ combination at $f = 194.1$ THz and $\Delta T = 37.5^\circ\text{C}$, $\Delta\varphi \simeq 0.026\lambda$. This is a several order-of-magnitude decrease in temperature dependence of birefringent phase as compared with either crystal separately.

4.5 Compound Crystals For Off-Axis Delay

Applications such as birefringent filters and polarization-mode dispersion (PMD) sources require the cascade of several crystals of equal length or having integral multiples of a unit length. The Lyot filter as adopted to interleavers is a classic example. But birefringent crystals are expensive and if the light beam is to pass through a cascade of equal-length crystals it would be better to fold the beam to pass through the same crystal several times. This has been proposed for both interleavers [13] and PMD generators [15]. For either component, after each pass of the crystal the polarization is rotated with a waveplate; the delay (crystal)/rotation (waveplate) sequence can be designed to make a birefringent filter.

There remains a problem, however. The problem is the beam must enter and exit the crystal, or temperature-compensated crystal pair, normal to the input face or else suffer double refraction. Double refraction is compounded by every pass and results in polarization-dependent loss. This is illustrated in Fig. 4.13(a). Here an off-normal beam is double-refracted by a first high-birefringent crystal. The two beams inside the crystal walkoff from one another and emerge offset. After polarization rotation from the intermediate waveplate, the two beams enter the second crystal and are each double refracted. Four beams emerge from the second crystal: however, the center two will overlap if the two crystal lengths are identical. Since the displaced beams will couple to a receiving collimator with different efficiencies, the concatenation generates PDL.

Double refraction for off-normal incidence onto a waveplate-cut crystal must be accepted because one polarization will see an effective index; the other, the ordinary index. This effect is treated in §3.6.2. One solution that fixes the walkoff problem but does not otherwise work is illustrated in Fig. 4.13(b). Here two equal-length crystals of the same material are placed with e -axes perpendicular to one another. The crystals are either both positive or negative uniaxial. The double refraction of the first crystal is corrected by the second crystal because extraordinary and ordinary rays are exchanged with one another. However, as shown in Fig. 4.13(c), the exchange used to cancel the net walkoff also exchanges the fast and slow axes; the delay imparted by the first crystal is cancelled by an equal and opposite delay from the second. In principle there is no net effect.

Inspection of Fig. 3.13 on page 114 shows that there is a solution for a birefringent delay having zero net walkoff with off-normal incidence. Any solution focuses on the Poynting vector directions, not the k -vector directions. One possible configuration is shown in Fig. 4.14(a): a positive uniaxial crystal is followed by a negative uniaxial crystal oriented such that the extraordinary axes are perpendicular [16]. In the first crystal the e -ray is refracted more than the ordinary ray because the crystal is positive uniaxial, Fig. 4.14(b). However, the e -Poynting vector is deflected by the refraction less than the ordinary Poynting vector. That is, the o -Poynting vector lies between the extraordinary Poynting and k -vectors. The e -Poynting vector splits from its k -vector because its linear polarization state is neither parallel nor perpendicular to the e -axis of the first crystal.

Entrance into the second crystal exchanges the designations of two rays. Also, both polarization states are either parallel or perpendicular to the e -axis, so there is no further splitting of the e -Poynting vector. Even though the o - and e -ray designations are flipped, the slow ray remains slow and the fast ray remains fast because the second crystal is negative uniaxial. Moreover, examination of the Poynting vectors shows that they will converge in the second crystal. The length ratio of the first and second crystal can be designed to impart a target delay and yield zero net walkoff.

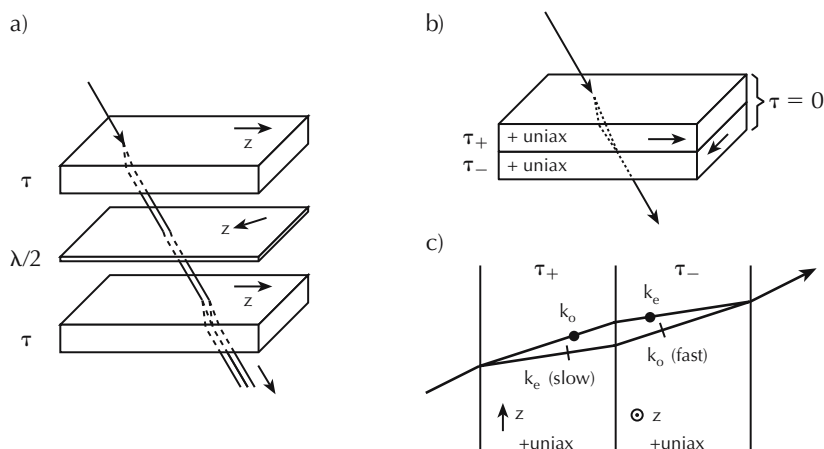


Fig. 4.13. Off-axis incidence angle. a) Off-axis passage through two delay crystals with an intermediate waveplate that rotates polarization (as in a filter) creates beam-splitting due to double refraction. The multiple output beams produce PDL. b) A “solution” that has zero net walkoff but no accrued delay. Two like crystals are placed perpendicular to one another. c) Ray trace of k -vectors. Fast exchanges with slow. No net delay.

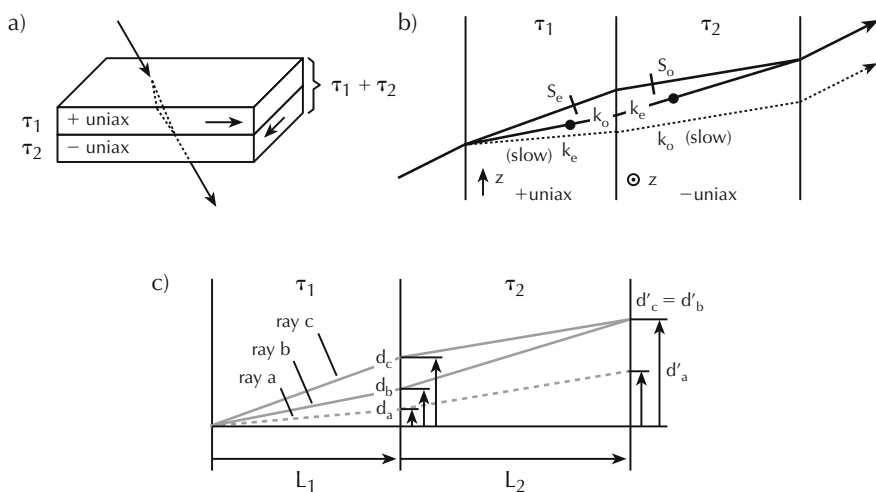


Fig. 4.14. An off-axis zero net walkoff crystal pair that accrues delay. a) Two different crystals with e -axes perpendicular, the first crystal is positive uniaxial and the second negative uniaxial. b) Ray trace. The e -Poynting vector splits from its k -vector in the first crystal, undergoing less deflection than the o -Poynting vector but at slower speed. In second crystal $e \leftrightarrow o$ but the slow path remains slow. c) Proper length ratio L_2/L_1 achieves zero net walkoff.

The length ratio for a two-crystal solution is calculated by equating the cumulative displacement of the two Poynting vectors. Figure 4.14(c) illustrates an a -ray, a b -ray, and a c -ray. The a -ray is the k -vector of the e -ray in the first crystal, the b -ray is the k - and Poynting vectors for the o -ray, and the c -ray is the Poynting vector of the e -ray. The b - and c -rays are to converge.

For small inclination angles all refraction expressions are linearized to good approximation. The displacements through the first crystal are

$$\begin{aligned} d_b(L_1) &= \theta_b^{(1)} L_1 \\ d_c(L_1) &= \left(\theta_a^{(1)} + \gamma \right) L_1 \end{aligned}$$

where $\theta_a^{(1)}$ and $\theta_b^{(1)}$ are the refraction angles and γ is the Poynting-vector tilt angle within the first crystal. The cumulative displacements through the second crystal are

$$\begin{aligned} d_b(L_1 + L_2) &= d_b(L_1) + \theta_b^{(2)} L_2 \\ d_c(L_1 + L_2) &= d_c(L_1) + \theta_a^{(2)} L_2 \end{aligned}$$

where $\theta_a^{(2)}$ and $\theta_b^{(2)}$ are the refraction angles in the second crystal. Zero net walkoff is achieved for $d_b(L_1 + L_2) = d_c(L_1 + L_2)$. This condition gives

$$\frac{L_2}{L_1} = \frac{\gamma - \left(\theta_b^{(1)} - \theta_a^{(1)} \right)}{\theta_b^{(2)} - \theta_a^{(2)}} \quad (4.5.1)$$

Taking the ambient index as one, the linearized refraction angles are

$$\theta_a^{(1)} = \theta_o/n_{e,1}, \quad \theta_b^{(1)} = \theta_o/n_{o,1}, \quad \theta_a^{(2)} = \theta_o/n_{e,2}, \quad \text{and} \quad \theta_b^{(2)} = \theta_o/n_{o,2}$$

The Poynting-vector tilt angle γ , that is, the angle at which the vector tilts away from its corresponding k -vector, comes from linearization of (3.6.15) on page 110 where $\theta \simeq 90^\circ$. As the negative tilt has already been accounted for, the linearized tilt angle is

$$\gamma = \frac{(n_{e,1}/n_{o,1})^2 - 1}{n_{e,1}} \theta_o$$

In terms of refractive indices, the length ratio is

$$\frac{L_2}{L_1} = \frac{n_{o,2}}{n_{o,1}} \frac{(n_{e,1}/n_{o,1} - 1)}{(n_{o,2}/n_{e,2} - 1)} \quad (4.5.2)$$

This ratio is positive when the first crystal is positive uniaxial and the second crystal is negative uniaxial. A YVO₄ and α -BBO crystal will satisfy this equation. One can verify that for this length ratio, the change in net displacement

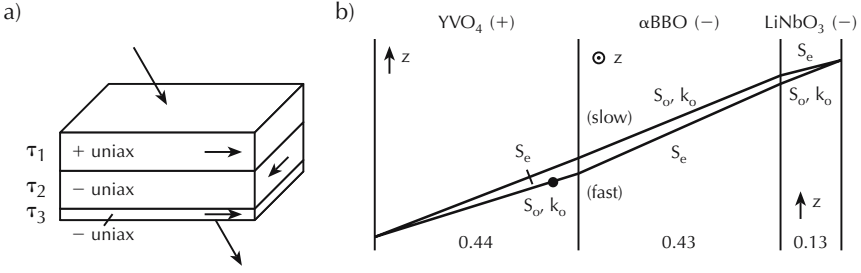


Fig. 4.15. Three crystal solution: accrued delay with zero net walkoff with inclined incidence and first-order temperature compensation. a) Specific crystal design. b) Ray trace of Poynting vectors. The Poynting vector is split from the e -ray k -vector in the first and third crystal.

for a change of incident angle is zero to first order. That is, net walkoff grows as second order with change in incident angle, making the compound crystal more tolerant to alignment.

An effective index for the crystal pair can be defined as $n_{\text{eff}} = \theta_o / \theta_{\text{eff}}$. The effective refraction angle is the ratio of total displacement to length, or $d_c = \theta_{\text{eff}} (L_1 + L_2)$. Combining terms gives

$$n_{\text{eff}} = \frac{L_2/L_1 + 1}{L_2/L_1 + n_{o,2}/n_{o,1}} n_{o,2} \quad (4.5.3)$$

The remaining problem is that the two crystals necessary to make the compound crystal are not necessarily temperature compensated. The one degree of freedom available, the length ratio, was used to enforce zero net walkoff. To make the compound crystal temperature insensitive as well a third crystal is necessary (Fig. 4.15(a)). Here YVO_4 , $\alpha\text{-BBO}$, and LiNbO_3 crystals stacked so that the e -axis of the $\alpha\text{-BBO}$ crystal is perpendicular to the other two axes give a solution.

A set of three linear equations can be solved to determine the three crystal lengths such that there is zero net walkoff, temperature is compensated to first order, and a target delay τ is achieved. In matrix form the equations are

$$\begin{pmatrix} \alpha_1 & -\alpha_2 & \alpha_3 \\ \Delta n_1 K_1 & -\Delta n_2 K_2 & \Delta n_3 K_3 \\ \Delta n_1 & -\Delta n_2 & \Delta n_3 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \tau c \end{pmatrix} \quad (4.5.4)$$

where K is the first-order temperature coefficient defined by (4.4.5a) (but for the refractive index, not group index), and the angle differences α are

$$\alpha_1 = \gamma - \left(\theta_b^{(1)} - \theta_a^{(1)} \right) = \frac{n_{e,1}}{n_{o,1}} \left(\frac{1}{n_{o,1}} - \frac{1}{n_{e,1}} \right) \theta_o \quad (4.5.5a)$$

$$\alpha_2 = \theta_b^{(2)} - \theta_b^{(1)} = \left(\frac{1}{n_{e,2}} - \frac{1}{n_{o,2}} \right) \theta_o \quad (4.5.5b)$$

$$\alpha_3 = \gamma - \left(\theta_b^{(3)} - \theta_a^{(3)} \right) = \frac{n_{e,3}}{n_{o,3}} \left(\frac{1}{n_{o,3}} - \frac{1}{n_{e,3}} \right) \theta_o \quad (4.5.5c)$$

The negative sign in the second column of (4.5.4) accounts for the perpendicular orientation of the α -BBO e -axis with respect to the other crystals.

Group index and thermal coefficients for YVO_4 , α -BBO, and LiNbO_3 are found in Table 4.6. The refractive indices were not measured for this table, but it was observed that the frequency dependence of the index was below an observable level. So the refractive indices are approximated by the group indices. Using the tabulated values, a YVO_4 crystal of length $L_{\text{yvo}} = 9.49$ mm, an α -BBO length of $L_{\text{abbo}} = 9.37$ mm, and a LiNbO_3 length of $L_{\text{ln}} = 2.84$ mm satisfies (4.5.4). The ray-trace of the Poynting vectors is shown in Fig. 4.15(b).

The practical drawback of temperature-compensated birefringent crystal sets, either on-axis or off-axis, is the widely disparate, highly anisotropic thermal expansion coefficients. This is seen in Tables 4.2 and 4.3. Coupling these materials with glasses and packaging alloys makes for a complex thermal design which stretches the ability to achieve athermal birefringent phase over a 70° operating range or more. The on-axis athermal crystal set of YVO_4 and LiNbO_3 has been used quite successfully, however, in the laboratory environment.

4.6 Polarization Retarders

Polarization retarders are essential parts used to build birefringent components such as isolators, circulators, interleavers, and polarization mode dispersion generators. Broadly speaking, a retarder is any optical element that transforms the state of polarization without loss. Birefringent crystals cut as waveplates, electro-optic materials, optically active materials, Faraday media, and total-internal reflection rhombs can all transform polarization. The distinguishing attribute is that two polarization components slip in phase within the medium, transforming its polarization state.

As a comment on nomenclature, retardation and birefringent phase represent the same physical effect. Both characterize a phase slip between two co-propagating waves having different wavelengths. The nuance is that retardation denotes the fractional phase slip when the overall slip is zero, one, or a few birefringent beats. As used in this text, birefringent phase is the total phase slip, fractional and integral, between two co-propagating plane waves of dissimilar wavelength.

The preponderance of this section analyzes birefringent waveplate retarders and select combinations. Highly multi-order waveplates and the thermal stabilization of birefringent phase were covered in §4.4. Total internal reflection retarders are covered last.

4.6.1 Half-Wave and Quarter-Wave Waveplates

The calculus of polarization transformation is extensively treated in Chapter 2 and the reader is particularly directed to §2.6. These tools are used below to write by inspection the transformation matrix operators and vector expressions for quarter- and half-wave waveplates.

A waveplate is made of birefringent material such as solid crystals, polyimide films, or liquid crystals, to cite a few. The waveplate technologies addressed in this section are solid crystals, for the formalism for polarization transformation is applicable to any technology. A waveplate made from birefringent crystal, typically a uniaxial crystal, is cut so that the extraordinary axis lies in the crystal face. Drawing from §3.6.4, the birefringent phase of a waveplate is defined by

$$\varphi = \frac{\omega}{c} (n_e - n_o) L \quad (4.6.1)$$

where L is the length of the crystal. Half- and quarter-wave waveplates generate the birefringent phases

$$\varphi_{\lambda/2} = (2n + 1)\pi \quad (4.6.2a)$$

$$\varphi_{\lambda/4} = \left(2n + \frac{1}{2}\right)\pi \quad (4.6.2b)$$

respectively, where the order n is a positive integer including zero. The length of an arbitrary order half-wave waveplate is

$$L_{\lambda/2} = \left(\frac{2n + 1}{2}\right) \frac{\lambda}{\Delta n} \quad (4.6.3)$$

where $\lambda/\Delta n$ is recognized as the birefringent beat length in the crystal. Clearly, as the order decreases so does the plate length. A true zero-order ($n = 0$) half-wave plate fabricated from crystalline quartz for $\lambda = 1545$ nm is about $92 \mu\text{m}$ thick, which requires special care to fabricate. The extra cost associated with the true zero-order plate is why low-order waveplates are attractive for many non-telecommunications applications. The polarization transformation of a waveplate of any retardation φ is given in Jones space by (2.6.19) on page 68 and in Stokes space by (2.6.25) on page 68, where the birefringent axis post-fix operators are resolved into matrix form via (2.6.29) on page 70.

Consider a waveplate with extraordinary axis rotated by θ to the horizontal in the plane perpendicular to propagation direction. The birefringent vector $\hat{r}$ in Stokes space is

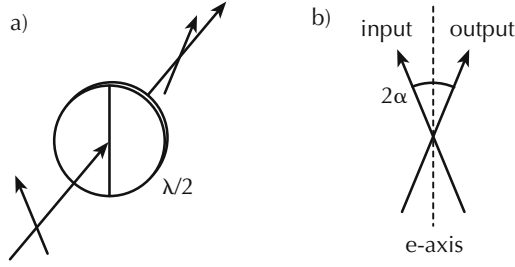


Fig. 4.16. Mirror image produced by a perfect half-wave waveplate. a) Incident linear polarization vector is mirror imaged about the extraordinary axis, producing a vector with the opposite tilt. b) Mirror image of a linear input state inclined to the extraordinary axis by α appears as a rotation by 2α .

$$\hat{r}(\theta) = \begin{pmatrix} \cos 2\theta \\ \sin 2\theta \\ 0 \end{pmatrix} \quad (4.6.4)$$

For a half-wave waveplate, the retardation is $\varphi = \pi$ and the Jones matrix operator is

$$U_{\lambda/2}(\theta) = -j \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \quad (4.6.5)$$

The -1 coefficient to the second diagonal term indicates a mirror image. The equivalent Stokes operator in vector form is

$$R_{\lambda/2} = 2(\hat{r}\hat{r}\cdot) - I \quad (4.6.6)$$

When resolved onto the basis implicit in (4.6.4), the Stokes matrix operator is

$$R_{\lambda/2}(\theta) = \begin{pmatrix} \cos 4\theta & \sin 4\theta & 0 \\ \sin 4\theta & -\cos 4\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (4.6.7)$$

Again the mirror image is apparent. A perfect half-wave waveplate generates the mirror image

$$(S_1, S_2, S_3) \longrightarrow (S_1, -S_2, -S_3) \quad (4.6.8)$$

along with rotation by 4θ about S_3 . The Stokes matrix operator imparts a fourfold multiple of the physical waveplate angle θ in its arguments. This is accounted for by first considering the $2\times$ multiple that results by going from Jones to Stokes space, and then the $2\times$ multiple generated by the mirror-image of the input state about the birefringent axis. Figure 4.16 illustrates the mirror image effect of a perfect half-wave waveplate and its apparent rotation about the birefringent axis. Figure 4.17(a) illustrates the Stokes space view of an $n = 0$ half-wave waveplate transformation. Polarization state a rotates to state b by precession about $\hat{r}$ by $\varphi = \pi$. Note that when the input state is linear (lies on the equator) the output is its orthogonal state.

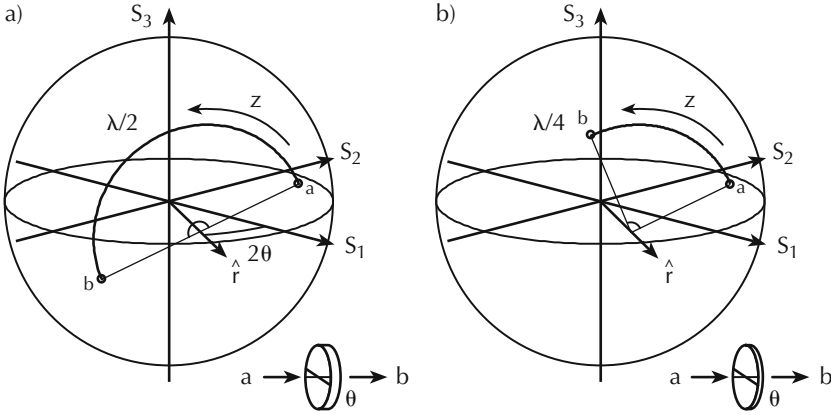


Fig. 4.17. Half-wave and quarter-wave waveplate transformations in Stokes spaces. a) Stokes picture of half-wave transformation. Birefringent axis $\hat{r}$ lies on the equator and is rotated by 2θ from S_1 . The birefringent axis corresponds to the extraordinary axis of the crystal. Input state a is transformed to b by $\varphi = \pi$ about $\hat{r}$. b) Stokes picture of quarter-wave transformation. Input state a rotates to b through $\varphi = \pi/2$.

For a quarter-wave waveplate rotated by θ , the retardation is $\varphi = \pi/2$ and the Jones matrix operator is

$$U_{\lambda/4}(\theta) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 - j \cos 2\theta & -j \sin 2\theta \\ -j \sin 2\theta & 1 + j \cos 2\theta \end{pmatrix} \quad (4.6.9)$$

The equivalent Stokes operator in vector form is

$$R_{\lambda/4} = (\hat{r}\hat{r}\cdot) + (\hat{r}\times) \quad (4.6.10)$$

When resolved onto the basis implicit in (4.6.4), the Stokes matrix operator is

$$R_{\lambda/4}(\theta) = \begin{pmatrix} \cos^2 2\theta & \sin 2\theta \cos 2\theta & \sin 2\theta \\ \sin 2\theta \cos 2\theta & \sin^2 2\theta & \cos 2\theta \\ -\sin 2\theta & \cos 2\theta & 0 \end{pmatrix} \quad (4.6.11)$$

Since no mirror image is derived from the quarter-wave plate, the arguments in $R_{\lambda/4}$ retain their $2\times$ multiple. The Stokes operator matrix is more complex as a result. Figure 4.17(b) illustrates the Stokes space view of an $n = 0$ quarter-wave waveplate transformation. Polarization state a rotates to state b via precession about $\hat{r}$ by $\varphi = \pi/2$.

Frequency Dependence

Birefringent phase (4.6.1) is directly proportional to optical frequency ω . The chromatic dependence is

$$\frac{d\varphi}{d\omega} = \varphi \left(\frac{1}{\omega} + \frac{1}{\Delta n_g} \frac{d(\Delta n_g)}{d\omega} \right) \quad (4.6.12)$$

Even excluding material dispersion, the chromatic dependence depends directly on the target value of φ :

$$\Delta\varphi = \frac{\Delta\omega}{\omega} \varphi \quad (4.6.13)$$

Waveplate order plays a role in the chromatic dependence. There is a direct tradeoff between the ease of fabrication versus the birefringent phase variation. For example, the frequency-dependent birefringent phase change for an n^{th} -order half-wave waveplate is

$$\Delta\varphi_{\lambda/2} = \frac{\Delta\omega}{\omega} (2n + 1) \pi \quad (4.6.14)$$

The change is a function of the retardation and the spectral bandwidth. Consider the spectral coverage for the C-band (4.1.5): $\mathcal{S}_C \simeq 2\%$. The change of birefringent phase across this band for three different waveplate orders is

$$\Delta\varphi_{\lambda/2} = \begin{cases} \pm 1.8^\circ & n = 0, \\ \pm 5.4^\circ & n = 1, \\ \pm 37.8^\circ & n = 10 \end{cases} \quad (4.6.15)$$

The change increment from zero order to first order alone imparts a threefold increase in chromatic dependence. Conversely, the bandwidth for a fixed $\Delta\varphi$ tolerance suffers a threefold decrease.

Depending on design requirements, waveplates in a component may have to be eliminated where possible to broaden the spectrum over which the component specifications are held. For example, in a circulator the extraordinary axes of the birefringent prisms can be cut in non-standard ways to align to the polarization axes rather than have a waveplate make the rotation. For other components where a waveplate is absolutely required, achromats made from waveplate combinations can in some cases be used.

4.6.2 Birefringent Waveplate Technologies

Fabrication of a waveplate requires high accuracy between the target retardation and the actual retardation of the part. Low-order waveplates are single- or double-side polished and the retardation is directly measured by interrupting the process. The waveplate is compared to a reference standard at the target wavelength and polishing continues until the waveplate is within tolerance.

So that there is minimum retardation change for each polish step, a very low birefringent material is used, such as crystalline quartz. Quartz has a birefringence of $\Delta n = 0.0084$ at $1.55 \mu\text{m}$ and associated birefringent beat length of $\Lambda \sim 184 \mu\text{m}$. A typical quartz waveplate is specified to within $\lambda/500$, which

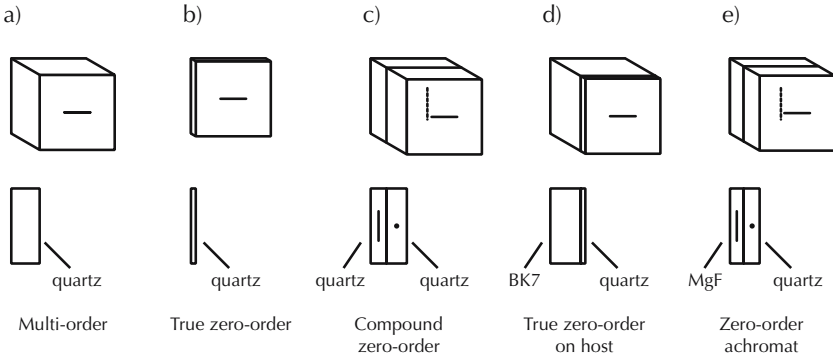


Fig. 4.18. Illustrations of solid-crystal waveplate technologies. a) Multi-order waveplate, single block of, e.g., quartz. b) True zero-order waveplate. Minimum thickness given the material birefringence to achieve required retardance. c) Compound zero-order waveplate makes for easier fabrication but poorer extinction ratio and angular aperture. d) Optically contacted true zero-order waveplate, very useful because quarter-wave waveplates of quartz, a low-birefringent material, are $45\ \mu\text{m}$ thick at zero order. e) Zero-order achromat, such as MgF_2 /crystalline quartz.

corresponds to a $\pm 0.2\ \mu\text{m}$ tolerance. Only with optical feedback is this tolerance possible.

Figure 4.18 illustrates five common types of waveplates. The easiest part to make is a multi-order waveplate (Fig. 4.18(a)). This waveplate has a thickness of $t = m\lambda/\Delta n + \delta t$, where m is the waveplate order, Δn is the birefringence, and δt is the minimum thickness required to achieve the target retardation. A half-wave waveplate at fourth order made with quartz is about $0.83\ \text{mm}$ thick. Such a part is certainly easy to handle. However, as discussed in the previous section, the bandwidth of such a high-order plate unsuitably low.

A true zero-order waveplate is the thinnest plate possible that meets the target retardation (Fig. 4.18(b)). The thickness of a true zero-order half-wave and quarter-wave waveplate made in crystalline quartz is $L_{\lambda/2} \simeq 92\ \mu\text{m}$ and $L_{\lambda/4} \simeq 46\ \mu\text{m}$ at $1.55\ \mu\text{m}$, respectively. It is hard to fixture and hold such thin plates, which until recently has made these plates expensive. However, optical contacting of quartz blanks to an optically flat fused silica block and one-side polishing of the blank eases the fixturing problem. The finished waveplate is removed by heating the block. Such waveplates can be very high quality and exhibit extinction ratios better than 55 dB.

A classic solution to the cost required to make a true zero-order plate is to make a compound zero-order plate (Fig. 4.18(c)). Here two thick quartz blocks are cemented or optically contacted with the extraordinary axes crossed. The difference in thickness determines the retardation. The two parts can be made arbitrarily thick as long as the difference is maintained. However, the compound zero-order waveplate is limited by the accuracy with which the birefringent axes can be crossed. Typical extinction ratios are no better than 27 dB.

The retardation produced by a compound zero-order plate is very sensitive to misalignment to the beam. The plate is zero-order not just at the center wavelength but when perpendicular to the beam. The angular aperture of the compound plate must be considered, and is much smaller than that of a true zero-order waveplate. Overlooking the effect of the angular aperture will lead to erroneous experimental results. Lastly, if the waveplate is to be rotated, the thickness of the compound waveplate will displace the beam if there is any wobble in the rotation. Programmable PMD sources are an example where the waveplates are mechanically rotated.

A very good option which is becoming more common today is to fabricate a true zero-order waveplate optically contacted to a permanent host such as BK7 (Fig. 4.18(d)). Typically a larger part is made and several waveplates with hosts are separated by dicing. Optical contacting is the process of bonding two highly polished glass surfaces together using van der Waals force and is the preferred method, as cement bonding causes undesired reflections at the interface. To strengthen the contact, the parts are annealed to make them inseparable. The hosted true zero-order waveplate allows the part to be handled and fixed into location with ease.

Finally, a bi-crystalline waveplate achromat can be made with materials having complementary wavelength dispersions. The MgF_2 /crystalline quartz achromat is one example (Fig. 4.18(e)). As both crystalline quartz and MgF_2 are positive uniaxial crystals, the birefringent axes are necessarily crossed. The difference in thickness-index product sets the net retardation [19].

4.6.3 Waveplate Combinations

Waveplate combinations can be used to create a variety of effects such as waveplate achromats or polarization controllers. Below are three important examples of waveplate combinations: the Evans phase shifter, the Pancharatnam achromat, and the Shirasaki achromat. Each combination uses multiple waveplates of various retardations and orientations to achieve a particular result. Not presented below but of some importance is the Koester stacked half-wave achromat [34]. Dynamic polarization control in elementary form is left to §4.6.4.

The Evans Phase Shifter

The Evans phase shifter [20] mechanically imparts a precession on an input polarization state, about an associated principal axis, that is equivalent to a change in frequency. Evans originally applied the phase shifter to tune a Lyot filter designed for astronomical observations. Other demonstrations have shown further application to astronomy [6], to interleavers [8], and to PMD sources [17].

The phase shifter, illustrated in Fig. 4.19, is a three-element combination of quarter-, half-, and quarter-wave waveplates. The equivalent birefringent

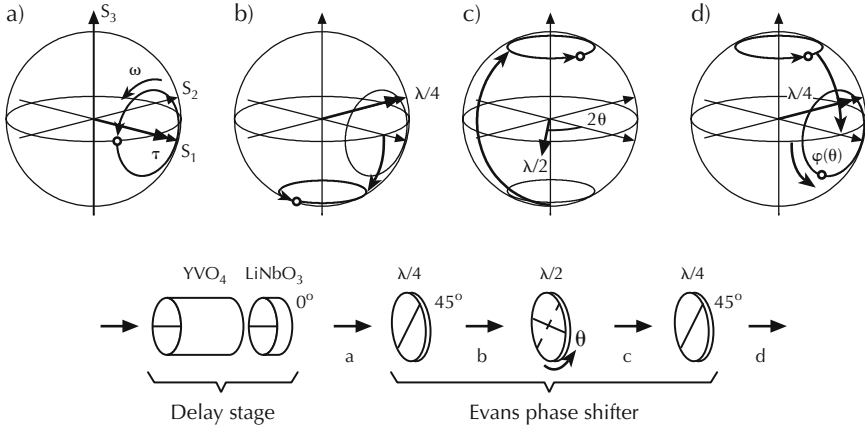


Fig. 4.19. Evans phase shifter. Quarter-, half-, quarter-wave waveplate combination, with outer plates fixed and center plate rotatable, tunes the birefringent phase of the adjacent principal waveplates. The quarter-wave plates are fixed at $+45^\circ$ with respect to the principal waveplate. In Stokes space: a) locus of output SOP from principal plates over frequency; open circle denotes one frequency. b) Quarter-wave transformation to lower pole. c) Half-wave transformation to upper pole and mirror image about birefringent axis. d) Quarter-wave transformation back to principal waveplate axis. The change in open-circle position is equivalent to a frequency change. The null position of the tuning plate is at -45° with respect to the principal axis.

axis of the combination is called the principal axis. The phase shifter can be located adjacent to a highly multi-order delay stage (such as the YVO_4 - LiNbO_3 stage, as illustrated) and will tune the birefringent phase of the stage when the principal axis of the phase shifter is aligned to the birefringent axis of the delay.

In the phase shifter, the two quarter-wave waveplates are aligned and their axes are further rotated by 45° with respect to the principal axis. The half-wave waveplate, called the tuning plate and located between the two quarter-wave waveplates, controls the birefringent phase of the cascade. The null position of the tuning plate is at -45° with respect to the principal axis. A rotation of the tuning plate is tantamount to a shift of the birefringent phase. Jones calculus shows that the quarter-, half-, quarter-waveplate cascade combines as

$$\frac{1}{2} \begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix} \begin{pmatrix} \cos 2\theta_\varphi & \sin 2\theta_\varphi \\ \sin 2\theta_\varphi & -\cos 2\theta_\varphi \end{pmatrix} \begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix} = \begin{pmatrix} e^{-j2\theta_\varphi} & 0 \\ 0 & -e^{j2\theta_\varphi} \end{pmatrix} \quad (4.6.16)$$

where the first and third Jones matrices describe the quarter-wave waveplates at $+45^\circ$ and the second Jones matrix describes the half-wave waveplate rotated by physical angle θ_φ . The resultant matrix shows a net birefringent phase plus a mirror image taken about the principal axis. The total birefringent phase of the delay and phase shifter is

$$\varphi(\omega, \theta_\varphi) = \omega\tau_s + (4\theta_\varphi - \pi) \quad (4.6.17)$$

The action of the Evans phase shifter in Stokes space is illustrated in Fig. 4.19. While the phase shifter can endlessly and continuously tune the birefringent phase of the cascade, there is no significant delay through the shifter. Endless rotation creates endless frequency shift (of the periodic spectrum) but without change in free-spectral range.

The Pancharatnam Achromat

Pancharatnam [5, 39] determined the conditions under which three waveplates can be combined so that the equivalent birefringent axis lies in the equatorial plane and the equivalent retardation is a prescribed value. His work can be reduced to the Evans phase shifter, which he briefly described and seems to have developed independently, but is more generally used to build achromatic retardation plates such as quarter-wave achromats. The Pancharatnam has a direct analogue to cascaded Mach-Zehnder interferometers uses in integrated optics to form achromatic waveguide-waveguide couplers [10]. The Pancharatnam and waveguide achromats were invented independently.

The Pancharatnam achromat is constructed with three waveplates, a first and last waveplate having equal retardation and extraordinary axis orientation, and an intermediate waveplate having a possibly different retardation and orientation. In this case, any choice of retardation and orientation values keeps the principal axis of the combination on the equator. There are two steps for the achromatic calculation: a first step derives the retardation and principal axis of the combination, and a second calculation determines the achromatic behavior.

As a departure from Pancharatnam's derivation, spin-vector calculus in Jones form is used here to determine the governing equations. For first and third waveplates having orientation $\hat{r}_1$ and retardation φ_1 , and a second waveplate having orientation $\hat{r}_2$ and retardation φ_2 , the Jones operators are

$$U_{1,2} = \cos(\varphi_{1,2}/2)I - j(\hat{r}_{1,2} \cdot \vec{\sigma}) \sin(\varphi_{1,2}/2) \quad (4.6.18)$$

As an aid to resolve the following operator product, the vector $\hat{r}_2$ is projected onto $\hat{r}_1$ and an orthogonal axis $\hat{r}_\perp$ as

$$\hat{r}_2 = \hat{r}_1 (\hat{r}_2 \cdot \hat{r}_1) + \hat{r}_\perp (\hat{r}_2 \cdot \hat{r}_\perp) \quad (4.6.19)$$

and, for the following, $\hat{r}_2 \cdot \hat{r}_1 = \cos 2\theta_{21}$, as is customary. Using the spin-vector identities in §2.5.4, the waveplate combination is written

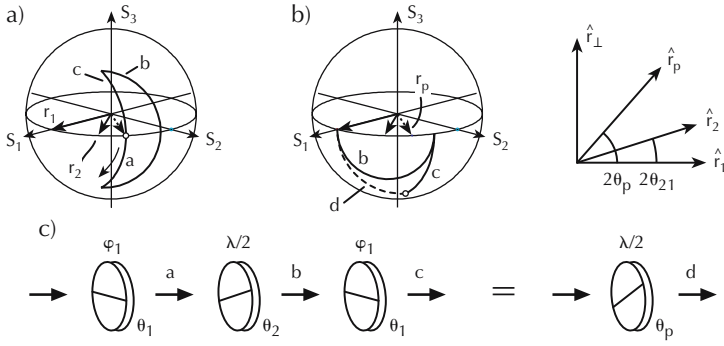


Fig. 4.20. Pancharatnam waveplate combination equivalent to single principal waveplate. a) Polarization evolution comparison, launched along eigen-axis of principal waveplate. Through combination, polarization first rotates along (a) about $\hat{r}_1$, then along (b) about $\hat{r}_2$, then returning along (c) about $\hat{r}_1$. b) Launch state along eigen-axis of $\hat{r}_1$. Through principal waveplate polarization rotates along (d) about $\hat{r}_p$. Through combination, polarization pirouettes about $\hat{r}_1$, rotates along (b) about $\hat{r}_2$, and rotates along (c) about $\hat{r}_1$. c) Equivalence between combination and principal waveplate; first and third waveplates of combination have same retardation and orientation.

$$\begin{aligned}
 U_1 U_2 U_1 = & \cos \varphi_1 \cos(\varphi_2/2) - \cos 2\theta_{21} \sin \varphi_1 \sin(\varphi_2/2) \\
 & - j(\hat{r}_1 \cdot \vec{\sigma}) \sin \varphi_1 \cos(\varphi_2/2) - j(\hat{r}_1 \cdot \vec{\sigma}) \cos 2\theta_{21} \cos \varphi_1 \sin(\varphi_2/2) \\
 & - j(\hat{r}_2 \cdot \hat{r}_\perp)(\hat{r}_\perp \cdot \vec{\sigma}) \sin(\varphi_2/2) \quad (4.6.20)
 \end{aligned}$$

Notice that there are no $\hat{s}_3$ components in (4.6.20); the principal axis of the combination lies in the equatorial plane.

The waveplate combination can be identified with a single principal waveplate U_p ,

$$U_p = \cos(\varphi_p/2)I - j(\hat{r}_p \cdot \vec{\sigma}) \sin(\varphi_p/2) \quad (4.6.21)$$

Making identification with (4.6.20), the principal retardation φ_p is

$$\cos(\varphi_p/2) = \cos \varphi_1 \cos(\varphi_2/2) - \cos 2\theta_{21} \sin \varphi_1 \sin(\varphi_2/2) \quad (4.6.22)$$

and the principal axis θ_p is

$$\cot(2\theta_p) = \csc(2\theta_{21}) (\sin \varphi_1 \cot(\varphi_2/2) + \cos 2\theta_{21} \cos \varphi_1) \quad (4.6.23)$$

where, in the latter case, the arc cotangent between the orthogonal $(\hat{r}_1 \cdot \vec{\sigma})$ and $(\hat{r}_\perp \cdot \vec{\sigma})$ axes was taken. Equations (4.6.22-4.6.23) are the two main results first derived by Pancharatnam.

The equivalence between a single waveplate and a Pancharatnam combination is illustrated in Fig. 4.20. There the polarization transformation through a principal waveplate and an equivalent combination of three plates, where the

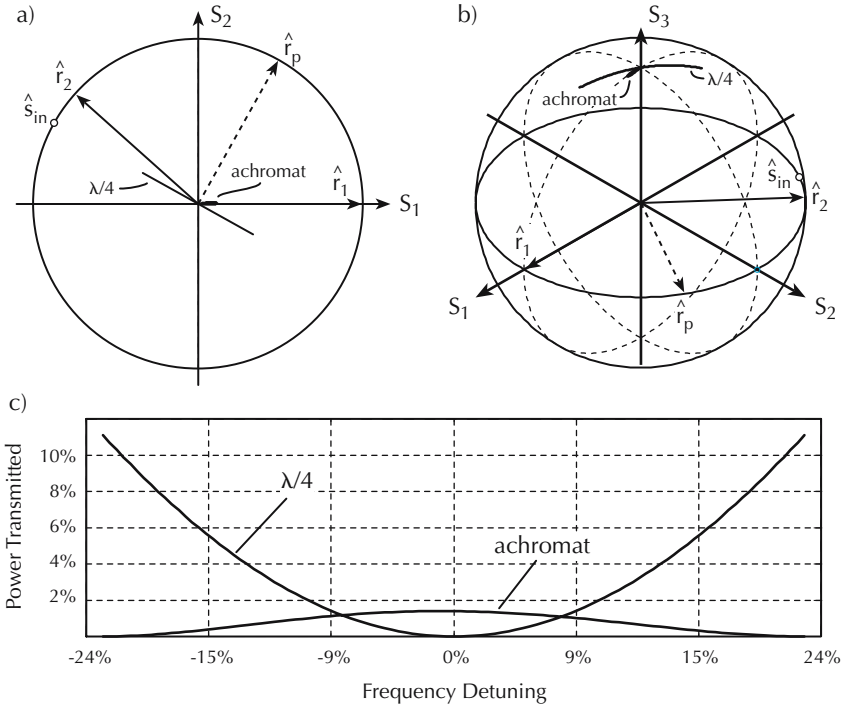


Fig. 4.21. Realization of a Pancharatnam achromat: $\varphi_p = \pi/2$, $\varphi_1 = 116.2^\circ$, $\varphi_2 = \pi$, $\theta_{21} = 69.1^\circ$ physical angle, $\theta_p = 30.3^\circ$ physical angle. a) Comparison of single $\lambda/4$ waveplate (principal axis dotted) and achromatic waveplate (extraordinary axes solid) over detuning range $\epsilon = \pm 25\%$, projected on Stokes space in plan view. Achromat shows tighter progression than single plate. b) Different view of same, input state rotated to top pole. c) Comparison of transmitted power through polarizer oriented along $\hat{s}_1$ for single plate and achromat.

center plate is half-wave, is illustrated. Note the distinctly different contours that are traced and yet the output states are the same in either case.

The Pancharatnam combination supports enough degrees of freedom to implement a waveplate combination that is less chromatically dependent than a single waveplate. This is called the Pancharatnam achromat. Consider a frequency deviation such that

$$\varphi \pm \delta\varphi = (1 \pm \epsilon)\varphi \quad (4.6.24)$$

The detuning will also be written as $\varphi^\pm$ until later expansion. The equations that define the solution require the same principal retardation and axis at $\varphi^\pm$:

$$\cos 2(\varphi_p/2) = \cos \varphi_1^+ \cos(\varphi_2^+/2) - \sin \varphi_1^+ \sin(\varphi_2^+/2) \cos 2\theta_{21} \quad (4.6.25a)$$

$$\cos 2(\varphi_p/2) = \cos \varphi_1^- \cos(\varphi_2^-/2) - \sin \varphi_1^- \sin(\varphi_2^-/2) \cos 2\theta_{21} \quad (4.6.25b)$$

$$\sin \varphi_1^+ \cot(\varphi_2^+/2) + \cos \varphi_1^+ \cos 2\theta_{21} = \cot(\varphi_2^-/2) \sin \varphi_1^- + \cos 2\theta_{21} \cos \varphi_1^- \quad (4.6.25c)$$

After some cumbersome manipulations and setting the center waveplate to half-wave retardation $\varphi_2 = \pi$, the retardation of the outer two plates satisfies the equation

$$\sin \epsilon \varphi_1 = \frac{\sin(\epsilon\pi/2)}{\cos(\varphi_p/2)} \sin \varphi_1 \quad (4.6.26)$$

The inclination of the center waveplate with respect to the outer plates is

$$\cos 2\theta_{21} = -\frac{\tan(\epsilon\pi/2)}{\tan(\epsilon\varphi_1)} \quad (4.6.27)$$

Lastly, the inclination of the principal axis is calculated from

$$\sin 2\theta_{21} \cot 2\theta_p = \sin(1 - \epsilon)\varphi_1 \tan(\epsilon\pi/2) + \cos(1 - \epsilon)\varphi_1 \cos 2\theta_{21} \quad (4.6.28)$$

Inspection of (4.6.26) shows that the possible achromatic bandwidth depends on the principal retardation. A half-wave retardation cannot be achromatic while a quarter-wave retardation exhibits the broadest possible range. For this reason the Pancharatnam achromatic is typically quarter-wave. Note, however, that (4.6.26–4.6.28) are derived for $\varphi_2 = \pi$; relaxation of this parameter allows more complexity and, in particular, a achromatic half-wave combination.

Figure 4.21(a,b) shows a Stokes-space comparison between a quarter-wave waveplate and an achromat designed to cover a detuning of $\epsilon = \pm 25\%$. As illustrated, a polarization state on the equator transformed through the achromat traces a tight loop about $\hat{s}_3$. In contrast, the same state through the quarter-wave waveplate traces a unidirectional arc through $\hat{s}_3$ as is expected from simple precession. Figure 4.21(c) shows the output states as transmitted through a polarizer. Again the bandwidth of the achromat is substantially broader than the single waveplate.

The Shirasaki Achromat

The Shirasaki achromat [55] compensates to first order the frequency dependence of a Faraday rotator (or optically active) waveplate for a particular input state of polarization. The input state is known in components such as optical isolators and circulators. A Faraday rotator waveplate precesses an input state of polarization about the $\pm \hat{s}_3$ axis, the sign determined by the relation between the magnetization vector and the propagation direction. Table 4.8 on page 207 lists Jones and Stokes operators for Faraday rotation.

One realization of the achromat, illustrated in Fig. 4.22(a), is constructed with a half-wave and then quarter-wave waveplate, followed by the Faraday

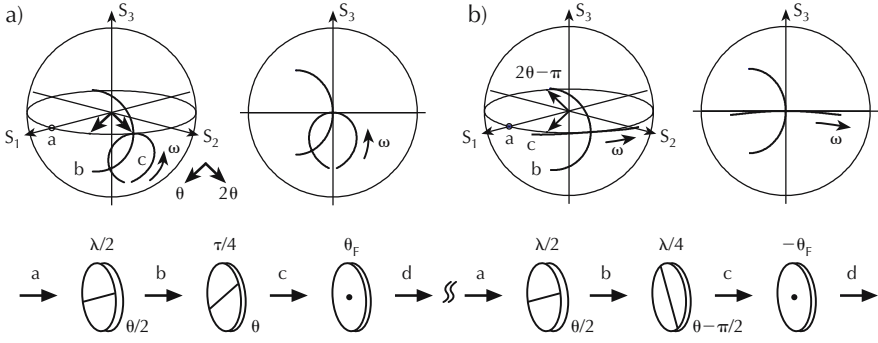


Fig. 4.22. Two realizations of the Shirasaki achromat: Half-wave waveplate followed by a quarter-wave waveplate, the combination preceding a Faraday rotator. a) Waveplates are oriented at $\theta/2$ and θ , where $\theta = 30^\circ$. To first order, the Stokes view shows frequency-dependent motion counter to that of a $+\hat{s}_3$ -oriented Faraday rotator. To second order the contour curvatures of the waveplates add, reducing the bandwidth. b) Quarter-wave waveplate rotated by -90° from a). Curvatures largely cancel and the bandwidth is increased. The transformation motion runs counter to a), requiring a reversed orientation of the Faraday rotator.

rotator. The extraordinary axis of the half-wave plate is rotated by $\theta/2$ from the horizontal while that of the quarter-wave plate is rotated by θ . The input state of polarization is expected to be linear and along the horizontal $+\hat{s}_1$. All three plates change retardation to first order with frequency; denote the changes as $\delta\varphi_F$, $\delta\varphi_{\lambda/4}$, and $\delta\varphi_{\lambda/2}$ for the Faraday, quarter-wave, and half-wave plates, respectively.

In Stokes space, the half-wave waveplate rotates the horizontal input polarization to another point on the equator, the angle of separation being 2θ . Considering small changes in frequency, the locus of polarization states forms a line perpendicular to the equator, to first order. The quarter-wave waveplate, whose birefringent axis is at 2θ , rotates the locus parallel to the equator. The achromat generates a state on the equator at 2θ that moves toward $+\hat{s}_1$ with increased frequency. This motion is counter to that of a Faraday rotator having its orientation along $+\hat{s}_3$, which rotates the polarization state toward $-\hat{s}_1$ with increased frequency. These two motions can cancel.

A rigorous analysis is easily done with spin-vector operators. Stokes operators are constructed for each plate and expanded to include first-order frequency deviation. The first-order operators are

$$R_F + \delta R_F = \hat{s}_3(\hat{s}_3 \cdot) + (\hat{s}_3 \times) + \delta\varphi_F(\hat{s}_3 \times \hat{s}_3 \times) \quad (4.6.29a)$$

$$R_{\lambda/4} + \delta R_{\lambda/4} = \hat{r}_4(\hat{r}_4 \cdot) + (\hat{r}_4 \times) + \delta\varphi_4(\hat{r}_4(\hat{r}_4 \cdot) - I) \quad (4.6.29b)$$

$$R_{\lambda/2} + \delta R_{\lambda/2} = 2\hat{r}_2(\hat{r}_2 \cdot) - I - \delta\varphi_2(\hat{r}_2 \times) \quad (4.6.29c)$$

where the birefringent vectors $\hat{r}_2$ and $\hat{r}_4$ lie in the equatorial plane. To first order, a frequency change is expanded as

$$\delta(R_F R_{\lambda/4} R_{\lambda/2}) \simeq (\delta R_F) R_{\lambda/4} R_{\lambda/2} + R_F (\delta R_{\lambda/4}) R_{\lambda/2} + R_F R_{\lambda/4} (\delta R_{\lambda/2}) \quad (4.6.30)$$

Interestingly, the second term on the right-hand side vanishes because $\hat{r}_4$ is aligned to the nominal polarization state produced by $R_{\lambda/2}$. When the operator (4.6.30) is applied to state $\hat{s}_1$ the contributing difference terms evaluate to

$$(\delta R_F) R_{\lambda/4} R_{\lambda/2} = \delta \varphi_F (\hat{s}_3 \times \hat{s}_3 \times) (\hat{s}_1 \cos 2\theta + \hat{s}_2 \sin 2\theta) \quad (4.6.31a)$$

$$R_F R_{\lambda/4} (\delta R_{\lambda/2}) = \delta \varphi_{\lambda/2} \sin \theta (\hat{s}_3 (\hat{s}_3 \cdot) + (\hat{s}_3 \times)) (\hat{s}_1 \sin 2\theta - \hat{s}_2 \cos 2\theta) \quad (4.6.31b)$$

By completing the vector products and setting the result to zero, a simple relation between frequency deviations of the Faraday and half-wave plates is found

$$\delta \varphi_F = \delta \varphi_{\lambda/2} \sin 2\theta \quad (4.6.32)$$

This relation makes physical sense because the precession rates need to be matched. The half-wave waveplate has a $\sin \theta$ multiplier because the radius of the precession circle depends on the angle between the input state and the extraordinary axis. In light of (4.6.13), the Stokes angle 2θ is defined by

$$\sin 2\theta = \frac{\varphi_F}{\varphi_{\lambda/2}} \quad (4.6.33)$$

In the case of an isolator, $\varphi_F = \pi/2$. Accordingly, the physical angles for the waveplates are $\theta_{\lambda/2} = 15^\circ$ and $\theta_{\lambda/4} = 30^\circ$. To maximize the second-order bandwidth, the waveplates should be made true zero-order.

Another consideration is necessary when analyzing the achromat to second order. As illustrated in Fig. 4.22(a) the curvature of states can add, which reduces the achromatic bandwidth. As an alternative, Fig. 4.22(b) illustrates the same achromat with the quarter-wave waveplate rotated by a 90° , which is tantamount to exchange of the ordinary and extraordinary axes. In this case the curvature of states imparted by the half-wave and quarter-wave waveplates subtract, providing a very broad band over which the combination is achromatic. When the quarter-wave plate is oriented like this, the state motion with increasing frequency is opposite that of the origin configuration. The Faraday rotator must be accordingly reversed. Whether the quarter-wave waveplate is to be rotated an additional 90° depends on θ : if $\theta_{\lambda/4} < 45^\circ$ then the additional rotation is necessary, otherwise it is not.

4.6.4 Elementary Polarization Control

A polarization controller dynamically changes its transformation effect to either track an incoming state of polarization or alter the output state polarization in a programmed way. Although a polarization controller can be arbitrarily complex with enough waveplates, the fundamental question is how few waveplates are needed for certain transformations.

A particularly useful control is “arbitrary-to-arbitrary,” wherein an arbitrary input polarization state is mapped to an arbitrary output state. Such a controller offers complete control. One subset is “linear-to-arbitrary,” which is useful when the input comes from a laser. A strictly endless controller is built with parts that are themselves endlessly rotatable, like a birefringent-crystal waveplate mounted on a rotary stage. Other controller types provide an apparently endless transformation but do so either by “unwinding” or digitally “flipping” the elementary parts. As an example, liquid crystal birefringent elements have a limited voltage range, which inhibits endless control of any one element. The same applies for fiber squeezers that induce birefringence through stress. Theoretically, unwinding or digital flipping can create endless polarization control, but the algorithms rely on certainty of the element retardations, which in practice is rarely the case. An extensive review of polarization control can be found in [60].

The elementary controllers considered in this section are built with a cascade of birefringent waveplates. These plates are physical pieces that have a fixed retardation and variable angle, although the birefringent axis is constrained to the equatorial plane. A common alternative in the industry is the lithium-niobate electro-optic polarization controller [26, 59]. Both birefringent and electro-optic controllers are strictly endless, and the electro-optic controllers can actuate in the megahertz range or higher. Also, unlike the birefringent waveplates, the bias voltage on an electro-optic controller can be changed as well as the orientation voltage, which changes the retardation. A polarization controller that combines effects of waveplate rotation and retardation modulation is called a hybrid controller.

Before doing an analysis of waveplate combinations, it should be obvious that a single fixed-retardation waveplate element cannot make arbitrary transformations. The only possible transformation type through a single waveplate is when the cross-product between the input state and desired output state lies on the equator, so that the birefringent axis can point in that direction, and the dot-product between the states is equal to the retardation of the plate.

Figure 4.23 illustrates transformations through single quarter- and half-wave waveplates over a full revolution of the plates for a fixed input state. Figures 4.23(a,b) illustrate the “bow-tie” pattern created by a quarter-wave waveplate at a fixed frequency, while Figures 4.23(c,d) show the constant-latitude pattern created by a half-wave waveplate. In particular note that a quarter-wave waveplate can transform a circular state to a linear state and back, while a half-wave waveplate maps linear states to linear and circular states to circular.

$\lambda/4, \lambda/4$ Arbitrary-To-Arbitrary Control

Two independently adjustable quarter-wave waveplates are the minimum requirement for arbitrary-to-arbitrary polarization transformation. To show that two quarter-wave plates with variable extraordinary axis inclination

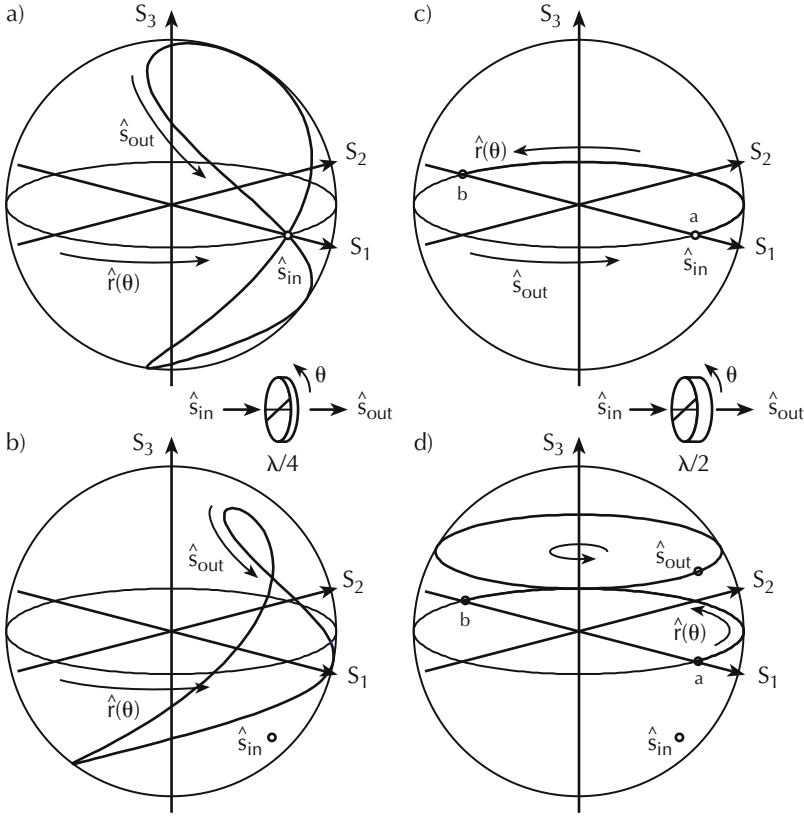


Fig. 4.23. Polarization control for single quarter- and half-wave waveplates. Trajectories show output locus for a fixed input over full revolution of the respective birefringent waveplate (shown in inset). a) “Bow-tie” locus traced by a quarter-wave plate for horizontal linear input polarization. b) Distorted bow-tie for elliptical input polarization. c) Line-of-latitude locus traced by a half-wave plate for horizontal (or any) linear input polarization. d) Elevated line-of-latitude for elliptical input polarization. Note the eccentricity does not change.

(Fig. 4.24(a)) can map any arbitrary polarization state to another arbitrary state, it is sufficient to show that orthogonal input states, such as along $\hat{s}_1$, $\hat{s}_2$, and $\hat{s}_3$, can each be mapped anywhere in Stokes space. An arbitrary input state can then be composed of these orthogonal states without violation of the mapping.

Recall that the Stokes operator for a quarter-wave plate is

$$R_{\lambda/4} = (\hat{r}\hat{r}\cdot) + (\hat{r}\times) \quad (4.6.34)$$

Two quarter waveplates in cascade generates the operator

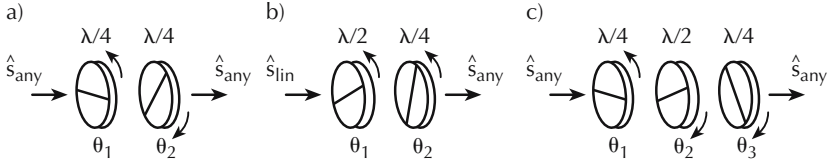


Fig. 4.24. Illustration of three birefringent waveplate polarization controllers. Each plate can be rotated independently. a) Quarter-, quarter-wave waveplate combination. Can map arbitrary-to-arbitrary polarization states. b) Half-, quarter-wave waveplate combination. Can map linear-to-arbitrary states. c) Quarter-, half-, quarter-wave waveplate combination. Can map arbitrary-to-arbitrary states.

$$\begin{aligned}
 R_2 R_1 &= \hat{r}_2 (\hat{r}_2 \cdot \hat{r}_1) (\hat{r}_1 \cdot) + \hat{r}_2 (\hat{r}_2 \cdot \hat{r}_1 \times) + \hat{r}_2 \times \hat{r}_1 (\hat{r}_1 \cdot) + \hat{r}_2 \times \hat{r}_1 \times \\
 &= \hat{r}_2 [\cos \theta_{21} (\hat{r}_1 \cdot) - \sin \theta_{21} (\hat{s}_3 \cdot)] + (\hat{r}_2 \times \hat{r}_1 \times) - \hat{s}_3 \sin \theta_{21} (\hat{r}_1 \cdot) \quad (4.6.35)
 \end{aligned}$$

Operating on three orthogonal input states, the output states are

$$R_2 R_1 \hat{s} = \begin{cases} \hat{r}_2 (\cos \theta_{21} \cos \theta_{1s}) + \hat{r}_\perp (\sin \theta_{1s}) - \hat{s}_3 (\sin \theta_{21} \cos \theta_{1s}) & \hat{s} = \hat{s}_1 \\ \hat{r}_2 (\cos \theta_{21} \sin \theta_{1s}) - \hat{r}_\perp (\cos \theta_{1s}) - \hat{s}_3 (\sin \theta_{21} \sin \theta_{1s}) & \hat{s} = \hat{s}_2 \\ -\hat{r}_2 (\sin \theta_{21}) - \hat{s}_3 (\cos \theta_{21}) & \hat{s} = \hat{s}_3 \end{cases} \quad (4.6.36)$$

where θ_{1s} is the angle between the associated vector and $\hat{s}_1$, and where, as a variation on (4.6.19),

$$\hat{r}_1 = \hat{r}_2 (\hat{r}_1 \cdot \hat{r}_2) + \hat{r}_\perp (\hat{r}_1 \cdot \hat{r}_\perp) \quad (4.6.37)$$

By definition, $\hat{r}_\perp$ is a perpendicular axis in the equatorial plane such that $\hat{r}_2 \cdot \hat{r}_\perp = 0$ and $\hat{r}_2 \times \hat{r}_\perp = \hat{s}_3$.

Consider the transformation of $\hat{s}_1$. The coefficients to $\hat{r}_2$, $\hat{r}_\perp$, and $\hat{s}_3$ can each independently span the range $[-1, 1]$, provided a judicious choice of θ_{21} and θ_{1s} is made. Therefore $R_2 R_1$ can map $\hat{s}_1$ to anywhere in Stokes space. The same argument applies to transformations of $\hat{s}_2$ and $\hat{s}_3$. In the latter case note that the direction of $\hat{r}_2$ is arbitrary: the linear component of the output polarization can be arbitrarily oriented.

Since an arbitrary input state is a linear combination of projections onto orthogonal axes, the linearity of $R_2 R_1$ ensures that any such input state can be mapped to an arbitrary output state.

$\lambda/2, \lambda/4$ Linear-To-Arbitrary Control

A half-wave plate followed by a quarter-wave plate (Fig. 4.24(b)), can transform any linear input state to an arbitrary output state. To show this, recall that the Stokes operator for a half-wave plate is

$$R_{\lambda/2} = 2(\hat{r}\hat{r}\cdot) - I \quad (4.6.38)$$

The half-wave (R'_2), quarter-wave (R_1) cascade generates the operator

$$\begin{aligned} R'_2 R_1 &= 2\hat{r}_2(\hat{r}_2 \cdot \hat{r}_1)(\hat{r}_1 \cdot) - \hat{r}_1(\hat{r}_1 \cdot) + 2\hat{r}_2(\hat{r}_2 \cdot \hat{r}_1 \times) - \hat{r}_1 \times \\ &= 2\hat{r}_2 [\cos \theta_{21}(\hat{r}_1 \cdot) - \sin \theta_{21}(\hat{s}_3 \cdot)] - \hat{r}_1(\hat{r}_1 \cdot) - (\hat{r}_1 \times) \end{aligned} \quad (4.6.39)$$

Operating on three orthogonal input states, the output states are

$$R'_2 R_1 \hat{s} = \begin{cases} \hat{r}_2 (\cos \theta_{21} \cos \theta_{1s}) - \hat{r}_\perp (\sin \theta_{21} \cos \theta_{1s}) - \hat{s}_3 (\sin \theta_{1s}) & \hat{s} = \hat{s}_1 \\ \hat{r}_2 (\cos \theta_{21} \sin \theta_{1s}) - \hat{r}_\perp (\sin \theta_{21} \sin \theta_{1s}) - \hat{s}_3 (\cos \theta_{1s}) & \hat{s} = \hat{s}_2 \\ -3\hat{r}_2 (\sin \theta_{21}) + \hat{r}_\perp (\cos \theta_{21}) & \hat{s} = \hat{s}_3 \end{cases} \quad (4.6.40)$$

As in the case of cascaded quarter-wave waveplates, input states $\hat{s}_1$ and $\hat{s}_2$ can be mapped arbitrarily. However, not so for a launch along $\hat{s}_3$. In that case the half-, quarter-wave pair only maps to the equator; any remaining component along $\hat{s}_3$ vanishes. This makes physical sense because a circular state that transits a half-wave plate is rotated to the orthogonal circular state. The quarter-wave plate in turn rotates the resultant circular state to the equator.

The half-wave, quarter-wave cascade transforms any linear state to an arbitrary state. In the opposite direction, any arbitrary state can be mapped to a linear state. This may be useful when the light is subsequently analyzed by a polarizer.

$\lambda/4, \lambda/2, \lambda/4$ Arbitrary-To-Arbitrary Control

The same procedure is used to analyze the cascade of quarter-, half-, quarter-wave waveplates. To aid with the reductions, $\hat{r}_3 \cdot \hat{r}_{3\perp} = 0$ and $\hat{r}_3 \times \hat{r}_{3\perp} = \hat{s}_3$ define the vector $\hat{r}_{3\perp}$. The concatenated Stokes operator is

$$\begin{aligned} R_3 R'_2 R_1 &= 2\hat{r}_3(\hat{r}_3 \cdot \hat{r}_2)(\cos \theta_{21}(\hat{r}_1 \cdot) - \sin \theta_{21}(\hat{s}_3 \cdot)) \\ &\quad - \hat{r}_3(\hat{r}_3 \cdot \hat{r}_1)(\hat{r}_1 \cdot) - \hat{r}_3(\hat{r}_3 \cdot \hat{r}_1 \times) \\ &\quad + 2(\hat{r}_3 \times \hat{r}_2)(\cos \theta_{21}(\hat{r}_1 \cdot) - \sin \theta_{21}(\hat{s}_3 \cdot)) \\ &\quad - (\hat{r}_3 \times \hat{r}_1)(\hat{r}_1 \cdot) - (\hat{r}_3 \times \hat{r}_1 \times) \end{aligned} \quad (4.6.41)$$

Transformation of the three orthogonal input states generates the following expressions:

$$\begin{aligned} R_3 R'_2 R_1 \hat{s}_1 &= \hat{r}_3 \cos(\theta_{32} - \theta_{21}) \cos \theta_{1s} - \hat{r}_{3\perp} \sin \theta_{1s} - \hat{s}_3 \sin(\theta_{32} - \theta_{21}) \cos \theta_{1s} \\ R_3 R'_2 R_1 \hat{s}_2 &= \hat{r}_3 \cos(\theta_{32} - \theta_{21}) \sin \theta_{1s} + \hat{r}_{3\perp} \cos \theta_{1s} - \hat{s}_3 \sin(\theta_{32} - \theta_{21}) \sin \theta_{1s} \\ R_3 R'_2 R_1 \hat{s}_3 &= \hat{r}_3 \sin(\theta_{32} - \theta_{21}) + \hat{s}_3 \cos(\theta_{32} - \theta_{21}) \end{aligned}$$

As was the case with two quarter-wave waveplates, the output polarization for each orthogonal input is mapped arbitrarily in Stokes space provided a judicious choice of waveplate angles.

In comparison to (4.6.36), the present transformation makes the angle substitution $\theta_{21} \mapsto (\theta_{32} - \theta_{21})$. The addition of the intermediate half-wave plate provides a degree of freedom not available for the quarter-wave pair, wherein the value of $(\theta_{32} - \theta_{21})$ can be changed without simultaneous change of $\hat{r}_3$, $\sin \theta_{1s}$, or both. For this reason the quarter-, half-, quarter-wave combination is often preferred over the lone pair.

4.6.5 TIR Polarization Retarders

A total-internal reflection retarder is nearly achromatic and accordingly has practical application in components and instruments. As analyzed in §3.5.6, total internal reflection retards light due to the difference of evanescent field decay between TE and TM plane waves. While the retardation of a waveplate is directly proportional to frequency (4.6.1), the frequency dependence of TIR retardation is governed by the material dispersion alone. Indeed, differentiation of the retardance expression (3.5.47) on page 103 with respect to frequency yields

$$\frac{d\varphi}{d\omega} = \frac{2u}{u^2 + 1} \frac{1}{n(\omega) (n^2(\omega) \sin^2 \theta - 1)} \frac{dn(\omega)}{d\omega} \quad (4.6.42)$$

where

$$u = \frac{\cos \theta \sqrt{n^2(\omega) \sin^2 \theta - 1}}{n(\omega) \sin^2 \theta} \quad (4.6.43)$$

Selection of a low-dispersion material will produce a highly achromatic retarder.

A common TIR retarder is the Fresnel rhomb, illustrated in Fig. 4.25(a). The Fresnel rhomb is designed to convert linearly polarized light of the proper inclination to circular polarization after two reflections while maintaining the output co-linear with the input. The three design parameters are the retardance φ per TIR, the rhombohedral angle θ , and the material index n . Association of physical space to Stokes space is made by referencing the Stokes axis $\hat{s}_1$ to the TE direction on the plane of incidence. Since $\hat{s}_1$ is also associated with the positive eigenvector of the Jones operator U , the retardance expression

$$\varphi = 2 \tan^{-1} u \quad (4.6.44)$$

follows a right-hand precession rule about $\hat{s}_1$. Accordingly, linearly polarized light aligned to $+\hat{s}_2$ is transformed to circular polarization $\hat{s}_3$ when $\varphi = \pi/2$.

The practical design of a Fresnel rhomb should minimize the retardance sensitivities to frequency and incident angle. The frequency sensitivity is given above, and the angular sensitivity of total internal reflection retardance is

$$\frac{d\varphi}{d\theta} = \frac{2u}{u^2 + 1} \frac{2 - (n^2 + 1) \sin^2 \theta}{\sin \theta \cos \theta (n^2 \sin^2 \theta - 1)} \quad (4.6.45)$$

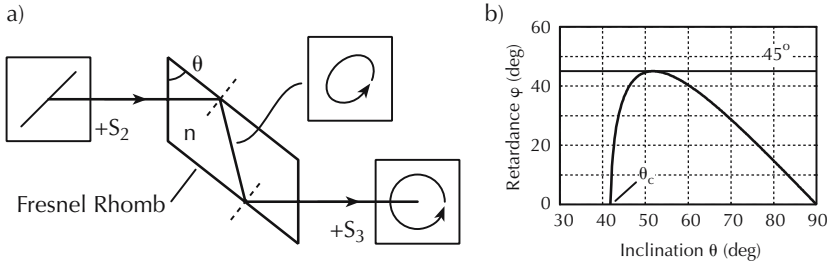


Fig. 4.25. A Fresnel rhomb can transform linear 45° polarization into circular polarization over a bandwidth limited only by material dispersion. a) Illustration of the rhomb, where two total-internal reflections impart a combined quarter-wave shift. b) Retardance of the rhomb as a function of apex angle θ , where $n = 1.497$. Retardance is zero at the critical angle and glancing angle. The retardance is 45° at $\theta = 51.8^\circ$ while the first-order retardance sensitivity to input angle, by design, vanishes.

The angular sensitivity vanishes for

$$\sin \theta_o = \sqrt{\frac{2}{n^2 + 1}} \quad (4.6.46)$$

The angle θ_o also yields the maximum retardance for a given index n . Substitution of (4.6.46) into (4.6.43) gives the maximum u values for a given n :

$$u_{\max} = \frac{n^2 - 1}{2n} \quad (4.6.47)$$

For example, a lead-doped glass having index $n \sim 1.8$ generates a retardance per TIR of $\phi \sim 64^\circ$. As $\phi = 90^\circ$ is necessary for linear to circular conversion, the Fresnel rhomb typically uses two reflections to accumulate the full $\pi/2$ retardance.

To minimize the angular sensitivity while $\phi = \pi/4$, the value of u is determined from (4.6.44) and the associated index n is calculated from (4.6.47). Figure 4.25(b) plots the retardance for a single reflection as a function of angle for this solution. The angular sensitivity vanishes just at the point of eighth-wave shift.

Beyond the elementary analysis presented here, two complete studies of rhomb sensitivities can be found in [4, 42]. Also, reference [9] applies TIR prisms to isolators and circulators to greatly extend their bandwidth; however the implementations are not particularly practical.

Separate from Fresnel rhombs, high-sensitivity magneto-optic sensors can use turning prisms to complete an optical circuit around a conductor [40]. When the light transits one or more unsaturated iron-garnet Faraday rotator elements located in proximity to the conductor, the Faraday rotation is proportional to the current-induced magnetic field. For such sensors, retardance

generated by the prisms reduces the small-signal sensitivity [41]. While the retardance per reflection can be reduced by bringing the incidence angle close to the critical angle, as can be seen in Fig. 4.25(b), the error sensitivities become impractically large. To overcome this limitation, a specially designed thin-film coating can be applied to the hypotenuse of the prism to reduce the retardance while remaining away from the critical angle. Reference [43] cites a design where the retardance was reduced to 1° at $1.3\ \mu\text{m}$ and the retardance remained within 6° for a $\pm 5^\circ$ angular error.

4.7 Single and Compound Prisms

Prisms are a cornerstone of birefringent optical components. Simple isotropic prisms are used as turning prisms to bend the light 90° or 180° within a component, or as straightening prisms to compensate for the angular divergence between two beams emergent from a dual-fiber collimator. Birefringent prism pairs are used in isolators to create polarization diversity. Compound prisms such as the Wollaston, Rochon, and Kaifa prisms are used in many circular designs for polarization diversity and angular compensation for dual-fiber collimators. The prisms studied in this section are designed in the small-angle limit, where Snell's law may be linearly approximated. A broad range of prisms is discussed in [24].

The isotropic isosceles prism illustrated in Fig. 4.26(a) has an apex angle α and refractive index n which, in general, is a function of wavelength. For a given angle of incidence θ_{in} on one face of the prism, the deflection angle β is

$$\beta = \theta_1 - \alpha + \sin^{-1} \left(\sin \alpha \sqrt{n^2 - \sin^2 \theta_1} - \cos \alpha \sin \theta_1 \right) \quad (4.7.1)$$

The angle of minimum deviation is that θ_1 , or θ_m , which minimizes β . While the expression is complicated in the general case, minimum deviation requires symmetry between the input and output angles: $\theta_1 = \theta_4$.

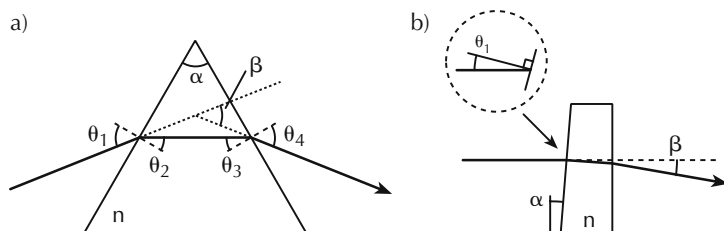


Fig. 4.26. Isosceles and small-angle prisms. a) Isosceles prism with apex angle α and refractive index n . Prism deflects input beam by angle β . b) Small-angle prism with near-normal incidence. For small angles $\beta = (n - 1)\alpha$. This shape is also called a wedge prism.

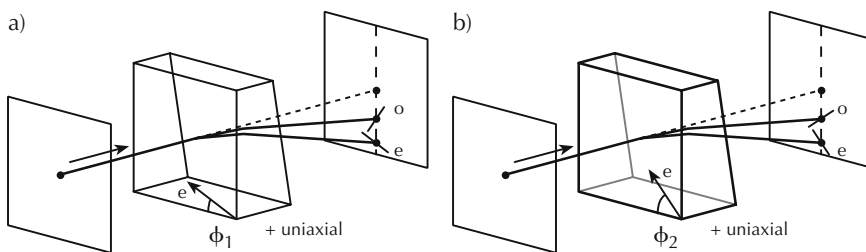


Fig. 4.27. Birefringent prisms having extraordinary axis perpendicular to input beam. The deflection of the beam is polarization dependent. For a (+) uniaxial crystal the polarization aligned to the extraordinary axis is deflected more than that aligned to the ordinary axis. The output polarization states are aligned to the ordinary and extraordinary axes of the prism. Deflection and output polarization states are independent. a) e axis tilted at angle ϕ_1 . b) e axis tilted at angle ϕ_2 .

For prisms with small apex angles and near-normal incidence (Fig. 4.26(b)), (4.7.1) is linearized to yield the deflection of a small-angle prism,

$$\beta \simeq (n - 1) \alpha \quad (4.7.2)$$

The input angle for minimum deviation in this case is $\theta_m = \alpha n$.

The birefringent prism is a prism made of birefringent crystalline material, typically uniaxial material. Birefringent prisms for isolators and most circulators are small angle prisms. A birefringent prism refracts like an isotropic prism and obeys (4.7.1), but the refractive index depends on the input polarization. In turn, the deflection is polarization dependent. Figure 4.27 illustrates two birefringent prisms made of the same material and having the same apex angle. The difference between the two prisms is the orientation of the extraordinary axis. For positive uniaxial material, the e -ray deflects more than the o -ray. The linear polarization states of the two output beams are aligned to the ordinary and extraordinary axes of the prism. The e -axis orientation determines the output polarization orientation but does not contribute to the deflection.

The following studies detail birefringent prism pairs in Wollaston, Rochon, Kaifa, and Shirasaki configurations.

4.7.1 Wollaston and Rochon Prisms

The Wollaston and Rochon compound prisms are birefringent prism pairs that angularly separate orthogonal linear polarization states. The Wollaston type uses two prisms with the same apex angle and material, and the extraordinary axes are crossed (Fig. 4.28(a)). The line of contact between the two parts is the hypotenuse of the prisms, and the input and output faces are parallel to one another. The prism is generally oriented perpendicular to, or with a small tilt to, the input beam. Two beams emerge from the compound prism,

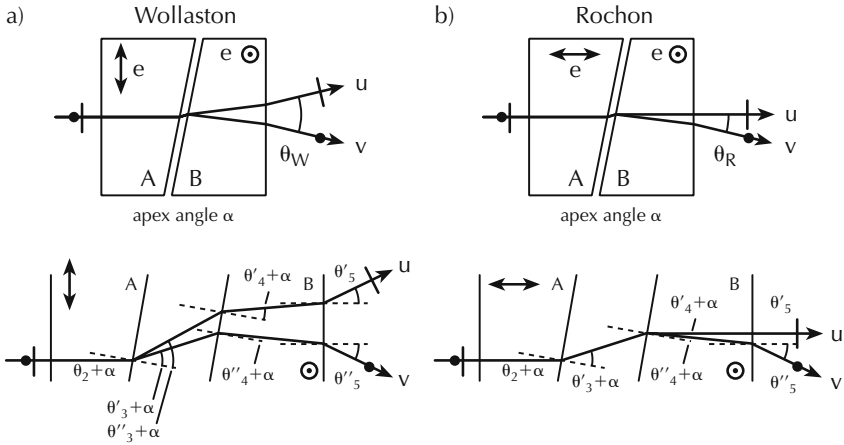


Fig. 4.28. Wollaston and Rochon compound prisms separate an input beam based on its polarization. The Wollaston prism symmetrically separates the orthogonal states while the Rochon prism deflects only the state aligned with the e -axis in prism B. Below shows ray-trace of orthogonal polarization components.

the u -beam following the u -path and the v -beam following the v -path. The output angles are calculated from Snell's equation applied to each interface. From left to right, the equations for the u -path are

$$\sin \theta_1 = n_e \sin \theta'_2 \quad (4.7.3a)$$

$$n_e \sin (\theta'_2 + \alpha) = n_g \sin (\theta'_3 + \alpha) \quad (4.7.3b)$$

$$n_g \sin (\theta'_3 + \alpha) = n_o \sin (\theta'_4 + \alpha) \quad (4.7.3c)$$

$$n_o \sin \theta'_4 = \sin \theta'_5 \quad (4.7.3d)$$

where n_g is the index in the gap. For the v -path the equations are

$$\sin \theta_1 = n_o \sin \theta''_2 \quad (4.7.4a)$$

$$n_o \sin (\theta''_2 + \alpha) = n_g \sin (\theta''_3 + \alpha) \quad (4.7.4b)$$

$$n_g \sin (\theta''_3 + \alpha) = n_e \sin (\theta''_4 + \alpha) \quad (4.7.4c)$$

$$n_e \sin \theta''_4 = \sin \theta''_5 \quad (4.7.4d)$$

Taking incident angles as small but allowing the apex angle to be significant,¹ the two output deflections are related to the birefringence and apex angle as

$$\theta'_5 - \theta_1 \simeq (n_e - n_o) \tan \alpha, \quad \theta''_5 - \theta_1 \simeq -(n_e - n_o) \tan \alpha \quad (4.7.5)$$

The Wollaston deflection θ_W is the full angle between the outputs, which is

¹ The approximation is $\sin(\theta + \alpha) \simeq \theta \cos \alpha + \sin \alpha$.

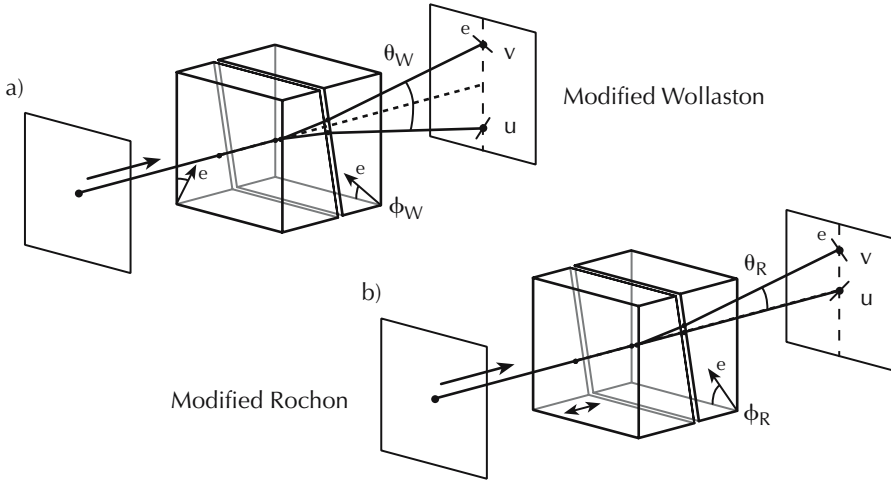


Fig. 4.29. Modified Wollaston and modified Rochon prisms. a) Modified Wollaston tilts the e -axes of the birefringent prisms while maintaining a 90° separation. The output polarization states are aligned to the e - and o -axes of the second prism. b) Modified Rochon prism tilts the e -axis of the second prism.

$$\theta_W = 2(n_e - n_o) \tan \alpha \quad (4.7.6)$$

As an example, to compensate for a 3° full-angle divergence between beams emergent from a dual-fiber collimator, using YVO_4 material which has a birefringence at $1.55 \mu\text{m}$ of $\Delta n = 0.2039$, the required apex angle is $\alpha \simeq 14.4^\circ$.

The Rochon compound prism Fig. 4.28(b) is like the Wollaston in shape but the e -axis of prism A is oriented along the propagation axis. In this way both u - and v -path polarizations see the ordinary index and experience the same refraction into the gap. Moreover, in prism B the u -path also sees the ordinary index which in turn imparts zero net deflection. Only the v -path is deflected at the output. The Rochon deflection θ_R is the full angle between the outputs, which is

$$\theta_R = (n_e - n_o) \tan \alpha \quad (4.7.7)$$

The Rochon has half the deflecting power of the Wollaston, but has the advantage of keeping one beam parallel to the input.

Path balancing is another key difference between the Wollaston and Rochon. For the Wollaston, in prism A the u -path has the extraordinary polarization while the v -path has the ordinary. In prism B these associations are reversed. As long as the prism lengths are the same the two paths are temporally balanced and one expects no appreciable PMD. In contrast, there is no temporal difference imparted by prism A of the Rochon, but there is by prism B. A single Rochon is not temporally balanced. The differential-group delays τ accumulated in the Wollaston and Rochon prisms are

$$\tau_W \simeq \frac{(n_e - n_o)(L_A - L_B)}{c} \quad (4.7.8a)$$

$$\tau_R \simeq \frac{(n_e - n_o)L_B}{c} \quad (4.7.8b)$$

where L_A and L_B are the lengths of the first and second prisms in each pair, and the small-angle limit is used. Using YVO_4 with a length $L_B = 2$ mm, the delay from a Rochon prism is $\tau \sim 1.3$ ps. This is an appreciable imbalance in the context of current component PMD specifications.

A variation of the Wollaston and Rochon compound prisms of significant practical importance is illustrated in Fig. 4.29 [56, 61]. The modified Wollaston compound prism changes the cut of the e -axes of prisms A and B while maintaining a 90° difference between them. The modified Rochon compound prism changes the e -axis cut in prism B. The modified Wollaston and modified Rochon prisms impart the same polarization-dependent deflections as their standard counterparts, but the linear states of polarization are rotated to align with the ordinary and extraordinary axes in prism B. Tilting of the extraordinary axes in this way adds a degree of freedom in the polarization-evolution schemes used in isolators and circulators.

4.7.2 Kaifa Prism

The Kaifa prism is a hybrid of the Wollaston and Rochon prisms and is illustrated in Fig. 4.30 [2]. The Kaifa prism serves two functions at once: the displacement of one polarization from the other, and the deflection of the two polarizations. The prism can be designed with no differential-group delay.

The compound prism is made from two birefringent prisms. Unlike the preceding prisms, the extraordinary axis in prism A is cut at angle α_{BC} to the longitudinal axis to produce Poynting vector walkoff along the u -path. For normal incidence, the k -vectors of the e - and o -rays remain coincident. At the hypotenuse interface the v -path follows the same path as in the Wollaston and Rochon prisms, while the u -path experiences a deflection that is between zero and $\theta_W/2$.

The Kaifa deflection θ_K is determined from ray tracing. The u -path follows

$$\sin \theta_1 = n_{\text{eff}} \sin \theta'_2 \quad (4.7.9a)$$

$$n_{\text{eff}} \sin (\theta'_2 + \alpha) = n_g \sin (\theta'_3 + \alpha) \quad (4.7.9b)$$

$$n_g \sin (\theta'_3 + \alpha) = n_o \sin (\theta'_4 + \alpha) \quad (4.7.9c)$$

$$n_o \sin \theta'_4 = \sin \theta'_5 \quad (4.7.9d)$$

where n_{eff} is determined from the birefringence and extraordinary axis angle α_{BC} (cf. (3.6.26) on page 117):

$$n_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2 \alpha_{BC} + n_o^2 \cos^2 \alpha_{BC}}} \quad (4.7.10)$$

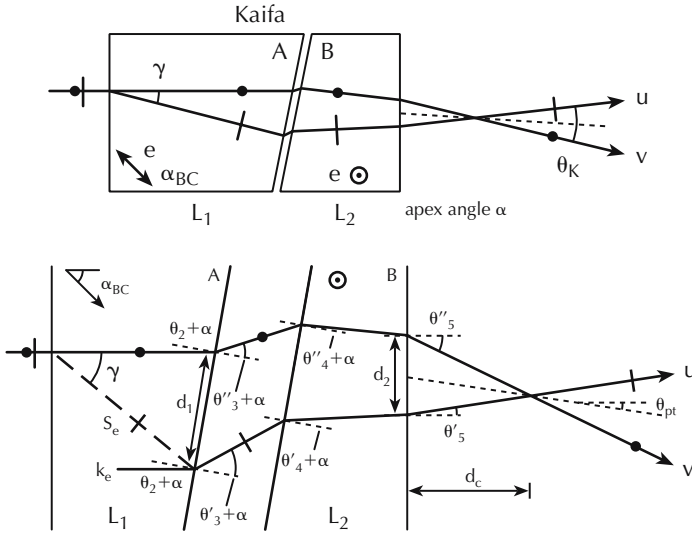


Fig. 4.30. The Kaifa prism is a hybrid of the Wollaston and Rochon prisms. Due to inclination of the extraordinary axis in prism A the Poynting vector of the extraordinary ray walks away from the ordinary ray. Prism B deflects both rays, but due to the intermediate refraction angle from prism A, the angle of the u -path output from prism B lies between that of the Wollaston and Rochon prisms.

Note that if the incident angle is not $\theta_1 = 0$ then (4.7.9a) must be replaced with (3.6.23) on page 115. For small incident angles, the deflection along the u -path is then

$$\theta'_5 - \theta_1 \simeq (n_{\text{eff}} - n_o) \tan \alpha \quad (4.7.11)$$

The v -path follows (4.7.4) with deflection (4.7.5). Accordingly, the Kaifa compound prism deflection angle is

$$\theta_K = (n_e + n_{\text{eff}} - 2n_o) \tan \alpha \quad (4.7.12)$$

The pointing direction, the center line of the two paths, is

$$\theta_{\text{pt}} = -\frac{1}{2} (n_e - n_{\text{eff}}) \tan \alpha \quad (4.7.13)$$

When $\alpha_{BC} = 0^\circ$ the Kaifa prism reverts to a Rochon prism with $n_{\text{eff}} = n_o$. Likewise, when $\alpha_{BC} = 90^\circ$ the prism reverts to a Wollaston prism with $n_{\text{eff}} = n_e$. In either of these two extremes there is no walkoff of the extraordinary path from the ordinary path.

The optical path lengths can be balanced in the Kaifa prism. The accumulated phase along the u - and v -paths is

$$\varphi_u = \frac{\omega}{c} (n_{\text{eff}} L_1 + n_o L_2) \quad (4.7.14a)$$

$$\varphi_v = \frac{\omega}{c} (n_o L_1 + n_e L_2) \quad (4.7.14b)$$

The requisite prism length ratio that balances the phase and thereby eliminates differential-group delay is

$$\frac{L_2}{L_1} = \frac{n_{\text{eff}} - n_o}{n_e - n_o} \quad (4.7.15)$$

The length L_1 of the first prism determines the displacement of the e -ray. The walkoff angle γ of the Poynting vector is governed by (cf. (3.6.15) on page 110)

$$\tan \gamma = -\frac{(n_e^2 - n_o^2) \sin \alpha_{\text{BC}} \cos \alpha_{\text{BC}}}{n_e^2 \cos^2 \alpha_{\text{BC}} + n_o^2 \sin^2 \alpha_{\text{BC}}} \quad (4.7.16)$$

The displacement d_1 at the end of prism A is

$$d_1 \simeq L_1 \tan \gamma \quad (4.7.17)$$

The change in displacement after transiting prism B is

$$(d_1 - d_2) \simeq L_2 \tan (\theta'_4 - \theta''_4) \quad (4.7.18)$$

Given the displacement at the end face of prism B and the full deflection angle, the distance to the crossing point of the u - and v -paths is approximately

$$d_c \simeq \frac{d_2}{\theta_K} \quad (4.7.19)$$

Expansion of the respective terms yields

$$d_c \simeq \left(\frac{\tan \gamma - \left(\frac{n_{\text{eff}} - n_o}{n_e - n_o} \right) \left(\frac{n_{\text{eff}}}{n_o} - \frac{n_o}{n_e} \right) \tan \alpha}{(n_e + n_{\text{eff}} - 2n_o) \tan \alpha} \right) L_1 \quad (4.7.20)$$

The Kaifa prism is useful in some circulator as well as interleaver applications because of the simplicity of its displacement and deflection behavior.

4.7.3 Shirasaki Prism

The Shirasaki compound prism illustrated in Fig. 4.31 is a birefringent prism pair designed to split an input light ray into orthogonal polarization components that run parallel to one another [54, 57]. The birefringent prisms are designed such that one polarization component undergoes total internal reflection while at the same interface the remaining polarization component is transmitted through the Brewster angle. The Shirasaki prism is a variation on the Glan-Taylor prism but directs the outputs to run parallel.

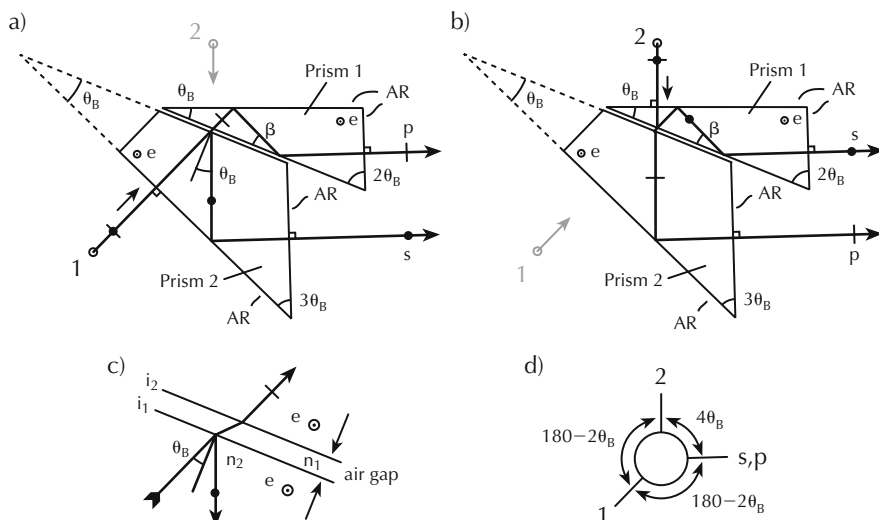


Fig. 4.31. The Shirasaki compound prism. Two birefringent prisms cut as illustrated separate a light beam into orthogonal linear components. At the air-gap interface between the two prisms one polarization is totally internally reflected while the other is transmitted through Brewster's angle. The extraordinary axis is aligned perpendicular to the page, and the output surfaces are AR coated. a) Input through port 1 separates polarizations onto two parallel paths, the p polarization runs along the top path. b) Input through port 2 also separates polarization, the p polarization runs along the bottom path. c) Optical path at air-gap interface. d) Angular orientation of four ports.

As illustrated in Fig. 4.31(a,b), there are four ports to the compound prism. Ports 1 and 2 have their polarization components resolved by the prism and diverted to run parallel at the output ports. Whether the p polarization component runs along the upper or lower output path depends on the input port, and likewise for the s polarization. The birefringent prisms have their extraordinary axes cut perpendicular to the plane in which the light travels. The compound prism is easily designed for positive uniaxial crystals.

The characteristic action of the compound prism lies at interface i_1 (Fig. 4.31(c)). Here one polarization component experiences total internal reflection, which is determined by the condition

$$n_2 \sin \theta_2 \geq n_1 \quad (4.7.21)$$

The other polarization is to experience total transmission through the Brewster angle. Recalling that the Brewster condition requires the reflected and refracted light to lie at right angles, or $\theta_1 + \theta_2 = \pi/2$, the reflection coefficient (3.5.39) on page 99 vanishes when

$$\tan \theta_2 = n_1/n_2 \quad (4.7.22)$$

With this condition satisfied, one writes $\theta_2 = \theta_B$, where here θ_B is the internal Brewster angle. For an isotropic material (4.7.21) and (4.7.22) cannot be simultaneously satisfied. However, they can be simultaneously satisfied for a uniaxial birefringent material. Consider, without loss of generality, a positive uniaxial material such that TIR occurs for the extraordinary index. In this case, the governing expressions are

$$n_e \sin \theta_2 \geq n_1 \quad (4.7.23a)$$

$$n_o \sin \theta_2 = n_1 \cos \theta_2 \quad (4.7.23b)$$

Combination of the square of these relations generates the condition that

$$n_e^2 - n_o^2 \geq n_1^2 \quad (4.7.24)$$

When the gap index is air, $n_1 = 1$ and (4.7.24) is satisfied. Rutile and YVO₄ satisfy this condition.

A consideration with this compound prism is the reflection coefficient when the prism index changes, as with temperature. With a temperature change, the input angle does not change but the index does. If one writes (3.5.39) as $\Gamma_{\text{TM}} = f/g$, then to first order about the Brewster angle,

$$\frac{d\Gamma_{\text{TM}}}{dn_2} \simeq \frac{df}{g} \quad (4.7.25)$$

Taking the differential and then substituting in Brewster's expression (4.7.22), reflectivity change from zero is

$$\frac{d\Gamma_{\text{TM}}}{dn_2} \simeq \frac{(n_2/n_1)^2 - 1}{2n_2} \quad (4.7.26)$$

Consider that the prisms are made of YVO₄. At room temperature the reflection of the ordinary ray is zero. For a 40° temperature change, given a thermal-optic coefficient of $dn_o/dT \sim 8.6 \times 10^{-6}$, the reflection changes to $|\Gamma_{\text{TM}}|^2 \sim -70$ dB. The Shirasaki prism was used to demonstrate a polarization-independent circulator having high extinction [35, 57].

Table 4.8. Summary of Waveplate Vector and Matrix Operators

Waveplate	Jones expressions	Stokes expressions
Quarter-wave		
$\{U, R\}$ Operators	$U_{\lambda/4} = \frac{1}{\sqrt{2}} (I - j(\hat{r} \cdot \vec{\sigma}))$	$R_{\lambda/4} = (\hat{r}\hat{r}\cdot) + (\hat{r}\times)$
$\{U, R\}_{\lambda/4}(\theta)$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 - j \cos 2\theta & -j \sin 2\theta \\ -j \sin 2\theta & 1 + j \cos 2\theta \end{pmatrix}$	$\begin{pmatrix} \cos^2 2\theta & \sin 2\theta \cos 2\theta & \sin 2\theta \\ \sin 2\theta \cos 2\theta & \sin^2 2\theta & \cos 2\theta \\ -\sin 2\theta & \cos 2\theta & 0 \end{pmatrix}$
$\{U, R\}_{\lambda/4}(45^\circ)$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -j \\ -j & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$
Half-wave		
$\{U, R\}$ Operators	$U_{\lambda/2} = -j(\hat{r} \cdot \vec{\sigma})$	$R_{\lambda/2} = 2(\hat{r}\hat{r}\cdot) - I$
$\{U, R\}_{\lambda/2}(\theta)$	$-j \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$	$\begin{pmatrix} \cos 4\theta & \sin 4\theta & 0 \\ \sin 4\theta & -\cos 4\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$
$\{U, R\}_{\lambda/2}(45^\circ)$	$\frac{j}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
Faraday rotator^(a)		
$\{U, R\}$ Operators	$U_F = \cos \theta_F I \mp j \sigma_3 \sin \theta_F$	$R_F = \cos 2\theta_F + (1 - \cos 2\theta_F)(\sigma_3 \sigma_3 \cdot) \pm \sin 2\theta_F (\sigma_3 \times)$
$\{U, R\}_F(\theta_F)$	$\begin{pmatrix} \cos \theta_F & \mp \sin \theta_F \\ \pm \sin \theta_F & \cos \theta_F \end{pmatrix}$	$\begin{pmatrix} \cos 2\theta_F & \mp \sin 2\theta_F & 0 \\ \pm \sin 2\theta_F & \cos 2\theta_F & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$\{U, R\}_F(45^\circ)$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \mp 1 \\ \pm 1 & 1 \end{pmatrix}$	$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \mp 1 & 0 \\ \pm 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$
^(a) The (+) and (−) signs refer to the relation of the magnetization vector and Faraday rotation direction of the material. Once a sign is set, it is fixed for both forward and backward propagation.		

References

1. M. Ariei, N. Takeda, Y. Tagami, and K. Shirai, "Magneto-optic garnet," U.S. Patent 4,932,760, June 12, 1990.
2. V. Au-Yeung, Q.-D. Gao, and X. L. Wang, "Optical circulator," U.S. Patent 6,331,912, Dec. 18, 2001.
3. M. Bass, Ed., *Handbook of Optics: Volume II*. New York: McGraw-Hill, Inc., 1995.
4. J. M. Bennett, "A critical evaluation of rhomb-type quarterwave retarders," *Applied Optics*, vol. 9, pp. 2123–2129, 1970.
5. B. H. Billings, Ed., *Selected Papers on Polarization*. Bellingham, Washington: SPIE Optical Engineering Press, 1990, vol. MS 23, SPIE Milestone Series.
6. J. Bland-Hawthorn, W. V. Breugel, P. R. Gillingham, I. K. Baldry, and D. H. Jones, "A tunable lyot filter at prime focus: A method for tracing supercluster scales at $z = 1$," *The Astrophysical Journal*, vol. 563, pp. 611–628, Dec. 2001.
7. C. D. Brandle, V. J. Fratello, and S. J. Licht, "Article comprising a magneto-optic material having low magnetic moment," U.S. Patent 5,608,570, Mar. 4, 1997.
8. C. F. Buhner, "Four waveplate dual tuner for birefringent filters and multiplexers," *Applied Optics*, vol. 26, no. 17, pp. 3628–3632, 1987.
9. —, "Quasi-achromatic optical isolators and circulators using prisms with total internal fresnel reflection," U.S. Patent 4,991,938, Feb. 12, 1991.
10. S. Cao, J. Chen, J. N. Damask, C. Doerr, L. Guiziou, G. Harvey, Y. Hibino, H. Li, S. Suzuki, K.-Y. Wu, and P. Xie, "Interleaver technology: Comparisons and applications requirements," *Journal of Lightwave Technology*, vol. 22, no. 1, pp. 281–289, Jan. 2004.
11. "Carpenter invar 36 alloy," Carpenter Technology Corporation, Wyomissing, Pennsylvania, 1990, edition date 08/01/1990.
12. "Kovar alloy," Carpenter Technology Corporation, Wyomissing, Pennsylvania, 1990, edition date 10/01/1990.
13. J.-H. Chen, K.-W. Chang, K. Tai, H.-W. Mao, and Y. Yin, "Apparatus capable of operating as interleaver/deinterleavers for filters," U.S. Patent 6,333,816, Dec. 25, 2001.
14. "Lithium niobate optical crystals," Crystal Technology, Inc., Palo Alto, CA, 1999. [Online]. Available: <http://www.crystaltechnology.com/LN-Optical-Crystals.pdf>
15. J. N. Damask, "Polarization mode dispersion generator," U.S. Patent 2002/0012487 A1, Jan. 31, 2002.
16. —, "Composite birefringent crystal and filter," U.S. Patent 6,577,445, June 10, 2003.
17. J. N. Damask, P. R. Myers, G. J. Simer, and A. Boschi, "Methods to construct programmable PMD sources, Part II: Instrument demonstrations," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1006–1013, Apr. 2004.
18. E. Desurvire, *Erbium-Doped Fiber Amplifiers, Principles and Applications*. Hoboken, New Jersey: Wiley-Interscience, 2002.
19. S. M. Etzel, A. H. Rose, and C. M. Wang, "Dispersion of the temperature dependence of the retardance in SiO_2 and MgF_2 ," *Applied Optics*, vol. 39, no. 31, pp. 5796–5800, Nov. 2000.
20. J. W. Evans, "The birefringent filter," *Journal of the Optical Society of America*, vol. 39, no. 3, pp. 229–242, 1949.

21. V. J. Fratello and R. Wolfe, *Handbook of Thin Film Devices, Vol. 4: Magnetic Thin Film Devices*. San Diego: Academic Press, 2001, ch. Epitaxial Garnet Films for Nonreciprocal Magneto-Optic Devices, pp. 93–141.
22. C. E. Gaebe, “Optical isolator and alignment method,” U.S. Patent 5,737,349, Apr. 7, 1999.
23. D. J. Gauthier, P. Narum, and R. W. Boyd, “Simple, compact, high-performance permanent-magnet faraday isolator,” *Optics Letters*, vol. 11, no. 10, pp. 623–625, 1986.
24. E. Hecht, *Optics*, 2nd ed. Reading, Massachusetts: Addison-Wesley Publishing Company, 1987.
25. A. J. Heiney and D. K. Wilson, “Optical isolators employing oppositely signed faraday rotating materials,” U.S. Patent 5,087,984, Feb. 11, 1992.
26. F. Heismann, “Analysis of a reset-free polarization controller for fast automatic polarization stablition in fiber-optic transmission systems,” *Journal of Lighthwave Technology*, vol. 12, no. 4, pp. 690–699, Apr. 1994.
27. K. Hiramatsu, K. Shirai, and N. Takeda, “Low magnet-saturation bismuth-substituted rare-earth iron garnet single crystal film,” U.S. Patent 6,031,654, Feb. 29, 2000.
28. “Hoya glass catalog,” Hoya, Incorporated, 2004.
29. *Optical Interfaces for Multichannel Systems with Optical Amplifiers*, International Telecommunication Union Std. ITU-T G.692, Oct. 1998.
30. *Spectral Grids for WDM Applications: DWDM Frequency Grid*, International Telecommunication Union Std. ITU-T G.694.1, June 2002.
31. “PbMoO₄ data sheet,” Isomet Corporation, Springfield, Virginia, 2003. [Online]. Available: <http://www.isomet.com/>
32. “Casix product catalog 2003,” JDSU, Inc., Canada, 2003. [Online]. Available: <http://www.casix.com/crystals/birefringentcrystal.htm>
33. D. Jones, S. Diddams, J. Ranka, A. Stentz, R. Windeler, J. Hall, and S. Cundiff, “Carrier-envelope phase control of femtosecond modelocked lasers and direct optical frequency synthesis,” *Science*, vol. 288, pp. 635–639, 2000.
34. C. J. Koester, “Achromatic combinations of half-wave plates,” *Journal of the Optical Society of America*, vol. 49(4), pp. 405–409, Apr. 1959.
35. H. Kuwahara, “Optical circulator,” U.S. Patent 4,650,289, Mar. 17, 1987.
36. M. Legre, M. Wegmuller, and N. Gisin, “Investigation of the ratio between phase and group birefringence in optical single-mode fibers,” *Journal of Lightwave Technology*, vol. 21, no. 12, pp. 3374–3378, Dec. 2003.
37. U. Morgner, R. Ell, G. Metzler, T. R. Schibli, F. X. Kartner, J. G. Fujimoto, H. A. Haus, and E. P. Ippen, “Nonlinear optics with phase-controlled pulses in the sub-two-cycle regime,” *Physical Review Letters*, vol. 86, no. 24, pp. 5462–5465, 2001.
38. “Ohara glass catalog,” Ohara, Incorporated, Kanagawa, Japan, 2004. [Online]. Available: <http://www.oharacorp.com/swf/catalog.html>
39. S. Pancharatnam, “Achromatic combinations of birefringent plates,” *Proc. Indian Acad. Sci.*, vol. A41, pp. 137–144, 1955.
40. K. B. Rochford, A. H. Rose, and G. Day, “Magneto-optic sensors based on iron garnets,” *IEEE Transactions on Magnetics*, vol. 32, no. 5, pp. 4113–4117, 1996.
41. K. B. Rochford, A. H. Rose, M. N. Deeter, and G. W. Day, “Faraday effect current sensor with improved sensitivity-bandwidth product,” *Optics Letters*, vol. 19, no. 22, p. 1903, Nov. 1994.

42. K. B. Rochford, A. H. Rose, P. A. Williams, C. M. Wang, I. G. Clarke, P. D. Hale, and G. W. Day, "Design and performance of a stable linear retarder," *Applied Optics*, vol. 36, no. 26, pp. 6458–6465, Sept. 1997.
43. K. B. Rochford, A. Rose, M. Deeter, and G. Day, "Faraday effect current sensor with improved sensitivity-bandwidth product," in *Tenth Int'l Conference on Fiber Optic Sensors (SPIE 2360)*, B. Culshaw and J. Jones, Eds., Glasgow, Scotland, Oct. 1994, pp. 32–35.
44. "Optical glass description of properties," Schott Glass, Mainz, Germany, Sept. 2000, version 1.2. [Online]. Available: <http://www.us.schott.com/sgt/english/products/catalogs.html>
45. "Optical glass properties," Schott Glass, Mainz, Germany, Sept. 2000, version 1.2. [Online]. Available: <http://www.us.schott.com/sgt/english/products/catalogs.html>
46. "Synthetic fused silica," Schott Lithotec AG, Mainz, Germany, 2001. [Online]. Available: http://www.schott.com/lithotec/english/products/fused_silica.html
47. K. Shirai, K. Ishikura, and N. Takeda, "Bismuth-substituted rare earth iron garnet single crystal," U.S. Patent 5,565,131, Oct. 15, 1996.
48. —, "Low saturated magnetic field bismuth-substituted rare earth iron garnet single crystal and its use," U.S. Patent 5,512,193, Aug. 30, 1996.
49. K. Shirai, M. Sumitani, N. Takeda, and M. Ariei, "Optical isolator," U.S. Patent 5,278,853, Jan. 11, 1994.
50. K. Shirai, T. Takano, N. Takeda, and M. Ariei, "Polarization-independent optical isolator," U.S. Patent 5,345,329, Sept. 6, 1994.
51. K. Shirai and N. Takeda, "Faraday rotator," U.S. Patent 5,535,046, July 9, 1996.
52. —, "Bismuth-substituted rare earth iron garnet single crystal film," U.S. Patent 5,925,474, July 20, 1999.
53. K. Shirai, N. Takeda, and K. Hiramatsu, "Faraday rotator having a rectangular shaped hysteresis," U.S. Patent 5,898,516, Apr. 27, 1999.
54. M. Shirasaki, "Prism polarizer," U.S. Patent 4,392,722, July 12, 1983.
55. —, "Polarization rotation compensator and optical isolator using the same," U.S. Patent 4,712,880, Dec. 15, 1987.
56. M. Shirasaki and K. Asama, "Compact optical isolator for fibers using birefringent wedges," *Applied Optics*, vol. 21, no. 23, pp. 4296–4299, 1982.
57. M. Shirasaki, H. Kuwahara, and T. Obokata, "Compact polarization-independent optical circulator," *Applied Optics*, vol. 20, no. 15, pp. 2683–2687, Aug. 1981.
58. *Generic Reliability Assurance Requirements for Passive Optical Components*, Telcordia Technologies Std. GR-1221-CORE, 1999. [Online]. Available: <http://telecom-info.telcordia.com/site-cgi/ido/index.html>
59. S. Thaniyavarn, "Wavelength-independent polarization converter," U.S. Patent 4,691,984, Sept. 8, 1987.
60. G. R. Walker and N. G. Walker, "Polarization control for coherent communications," *Journal of Lighthwave Technology*, vol. 8, no. 3, pp. 438–458, Mar. 1990.
61. P. Xie and Y. Huang, "Compact polarization insensitive circulators with simplified structure and low polarization mode dispersion," U.S. Patent 6,052,228, Apr. 18, 2000.

Collimator Technologies

Fiber-optic collimation and focusing assemblies, together known as collimators, are used to launch a beam of light from an optical fiber into free space and then to capture that light and refocus it into the same or another fiber. Collimator technologies are necessary whenever the gap introduced by optical component cores exceeds several hundred microns. For example, the numerical aperture of Corning single-mode fiber (smf-28) is N.A. $\simeq 0.14$, which corresponds to a 16° full-angle beam expansion cone. Any fiber-to-fiber separation over a distance of several hundred microns introduces severe insertion loss into the system. A collimator assembly adds a lens between the fiber termination and free space region to allow the light to travel tens or hundreds of millimeters without severe loss penalties.

Collimators are essential elements for micro-optic packaging of isolators, circulators, interleavers, thin-film filters, free-space dispersion compensators, and free-space polarization-based components. The lengths of these exemplar components require end-to-end beam coupling that ranges from 5 mm to 150 mm. A collimator assembly and lensing design can be optimized for any particular gap, but in general the larger the gap, the greater are the sensitivities to wavefront distortion.

There are two axes in the collimator technology selection matrix, one axis being the choice of lens and the other axis the choice of assembly. The two predominant lenses are Graded-Index, or GRIN, lenses, and shaped lenses, often called C-lenses (Fig. 5.1). A third type of lens that has some applications for micro-optic components are Gradium lenses, but these lenses are not ubiquitous. The two predominant collimator assemblies are air-gap collimators and fused collimators. Air-gap collimators have a physical air gap between the fiber termination and the lens entry surface; the adjacent surfaces are AR coated to maximize transmission. Fused collimators fusion-splice the fiber directly to an index-matched lens. A third assembly type, used early in the development chronology and now obsolete, is the epoxy-joint collimator where the fiber and lens were directly attached with a bead of epoxy in the light path.

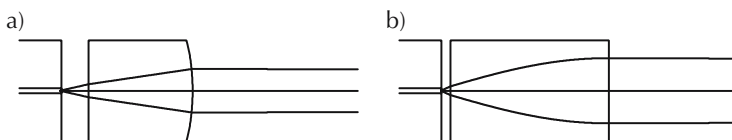


Fig. 5.1. Collimation of light from a fiber core by a) a shaped lens, and b) a GRIN lens. The wave-front curvature is eliminated by the curved surface of the shaped lens and progressively by the lateral index gradient of the GRIN lens.

Across the technology selection matrix there is a common set of optical, mechanical, and manufacturing-related design goals; these design goals are captured in Table 5.1. Insertion loss is typically reported by manufacturers as loss through a collimator pair. The high return loss is necessary to avoid coherent interference in a long chain of components. Pointing accuracy identifies the angular offset between the collimator housing and the axis of the exiting beam. WDM and spike power handling refers, respectively, to continuous operation and to the collimator's ability to tolerate a large surge of optical power, such as when an amplifier chain is cut or when a signal is first applied to the chain.

The mechanical goals include the means to attach and fix the collimator assembly to a micro-optic package. Most packages must be hermitically sealed to pass the lifetime requirements for passive components [21]. Accordingly, these packages are made of metal and all seams are soldered or welded. The collimator must conform to these hermeticity requirements as well. A collimator must maintain its integrity over a wide temperature range, which often includes a shipping range and an operational range. The pull tolerance refers to the handleability of the part and in particular to the avoidance of separation between the fiber pigtail and the lens/ferrule assembly. Lastly, as miniaturization is a demand on all micro-optic components, the collimator assembly must be small.

Finally, any collimator design must be manufacturable with high yield at low cost. The design must be, in a word: simple. As of this writing, most collimators are manually assembled. There is some push toward automation of the assembly process, but it is not clear that the tolerances achieved from automation are either required, superior, or cost effective.

The following sections provide theoretical and practical detail regarding collimator optics and technologies. The purpose of the theoretical analyses is to enable back-of-the-envelope estimates of component layout, dimensions, and lens selection. All truly accurate designs must be computed with commercially available ray-trace and beam propagation software. These numerical tools capture the effects of aberrations and provide the necessary level of tolerancing required for a robust design. The purpose of the technical overview is to highlight the central issues and challenges concerning collimator assemblies.

Table 5.1. Common Collimator Design Goals

Category	Specification	Goal
Optical	Low insertion loss	$< 0.25 \text{ dB}^{(a)}$
	High return loss	$> 60 \text{ dB}$
	Pointing accuracy	$\leq 1^\circ$
	WDM power handling	$> 100 \text{ mW}$
	Spike power handling	$> 1 \text{ W}$
Mechanical	Housing fixity	Solder or weld
	Temperature stability	$-20 \text{ to } +60 \text{ C}$
	Pull tolerance	5 N
	Small diameter	$< 4 \text{ mm}$
Manufacturing	High reliability	Pass Telcordia qualifications ^(b)
	High yield	$> 90 \%$
	Simple process	
	Low cost	

^(a)Measured through collimator pair. ^(b)Telcordia GR-1221-CORE [21].

5.1 Collimator Assemblies

This section details the assembly technologies of collimators. There are three assembly types: the epoxy-joint, the air-gap, and the fused-joint collimators. All assemblies seek to fix the position between a polished end of one or more single-mode fibers and a micro-optic lens. Moreover, all assemblies must be fixable to the metal component housing.

Epoxy-Joint Collimators

The epoxy-joint assembly, illustrated in Fig. 5.2, is the earliest type of integrated package though now obsolete. The three key elements are the fiber ferrule, the lens, and the sleeve. The ferrule is a quartz cylinder specially manufactured so that a wet chemical etch opens a capillary tube lengthwise down the center. One or more unjacketed single-mode fibers are inserted through the tube and epoxied into place with heat-curing epoxy. A typical heat-curing epoxy is 353ND manufactured by Epoxy Technology, Inc., in Billerica, MA. Typically a one-hour cure at 85°C completely fixes the fiber and ferrule together. The fiber end(s) are then clipped and the end face is polished so that the fiber and ferrule terminate on the same plane.

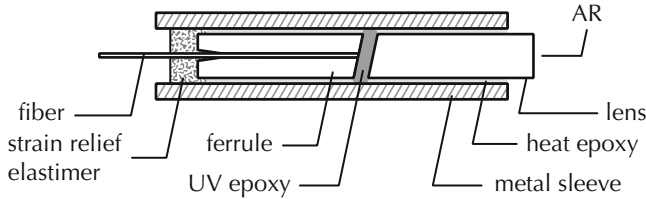


Fig. 5.2. Epoxy-joint collimator. Fiber is threaded through ferrule and fixed with heat-curing epoxy. Ferrule and fiber end face then polished at an angle ($6\text{--}8^\circ$). Ferrule and lens (with angle-polished facet) are aligned and set with UV-curing epoxy. The assembly is inserted into a metal sleeve (the sleeve may or may not cover the ferrule/lens joint) and fixed with heat-curing epoxy. A strain-relief elastomer is added around the exposed fiber to increase pull tolerance. Final assembly is soldered to micro-optic package.

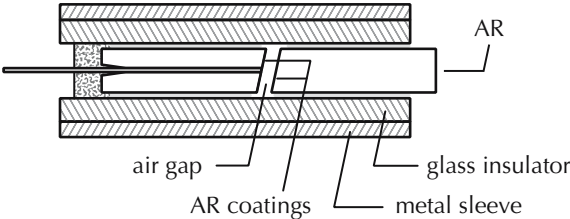


Fig. 5.3. Air-gap collimator. Lens and ferrule are aligned within a glass insulator sleeve. Inner facets of lens and ferrule are polished at an angle ($6\text{--}8^\circ$) and subsequently AR coated to limit back reflection. The gap is adjusted for optimal position and then tacked with UV epoxy around the perimeter of the assembly but not within the gap. Assembly is further fixed with heat-curing epoxy. Assembly is then loaded into metal sleeve and fixed with heat-curing epoxy. Final assembly is soldered to micro-optic package.

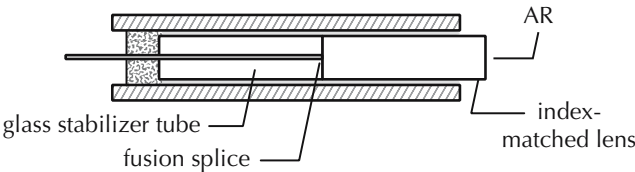


Fig. 5.4. Fusion-joint collimator. One or two fibers are directed fused to an index-matched lens. The fiber is threaded through a glass stabilizer tube for mechanical integrity. No angled facets nor AR coatings are required. The assembly is loaded into a metal sleeve and fixed with heat-curing epoxy. Final assembly is preferably laser-welded to micro-optic package.

All early collimator assemblies used GRIN lenses because of their small size and availability. For this generation and the ones to follow, the outer face of the lens is anti-reflection coated to increase transmission and reduce back reflection. The pitch of the GRIN lens (defined by (5.2.44) on page 232) is generally selected as $P = 0.23$ [18, 24], where a pitch of $P = 0.25$ is the theoretical choice for a collimating lens. The small reduction of the pitch serves two purposes. First, it is a practical step necessary to allow for a small gap between the front face of the ferrule and the back face of the lens. A quarter-pitch lens requires the fiber end face to be positioned directly on the back face of the lens. Second, in recognition that the fiber core is not a point source, rays emitted from the edge of the core are not over collimated with a $P = 0.23$ lens.

A ferrule end face that is perpendicular to the core creates high Fresnel reflection which leads to unacceptable back reflection into the fiber. To reduce the back reflection the ferrule is polished with a tilt angle, typically in the range of $6\text{--}8^\circ$. The back face of the GRIN lens is likewise polished. The early designs used the same angle of polish for the ferrule and lens, not accounting for an unintended Fabry-Perot cavity nor the refraction difference due to the different indices of the fiber and lens.

The epoxy-joint collimator assembly fixes the ferrule to the lens with a UV epoxy. Prior to the UV shot, the ferrule is positioned with positioning stages to the appropriate location behind the lens. Reference [24] cites the OG154 UV-curing epoxy from Epoxy Technology. It is interesting to record the indices at $1.55\text{ }\mu\text{m}$ for SMF-28 fiber $n_{\text{eff}} = 1.4682$ [9], for OG154 $n \simeq 1.545$, and for GRIN lens $n_o \sim 1.57 - 1.61$ [15]. The UV epoxy does a better job of index matching the fiber to the lens than would air, but the residual difference creates excess loss. After the UV tack, a heat-curing epoxy is painted around the joint and the assembly is heat cured. As a final step a metal sleeve is slipped around the assembly and heat-cured into place. Early sleeves were stainless steel [18] with a gold plating for better soldering ability to a metal housing.

The problems with epoxy-joint collimators are many. First, the UV epoxy in the optical path severely degrades the reliability of the component, especially under the damp-heat tests specified by [21]. UV epoxies in general have low humidity resistance as compared with heat-curing epoxies. In particular, the epoxy-joint collimator is reported to tolerance the 85–85 test (85% humidity at 85°C) for only 350 hours [25]. Second, the attachment of the ferrule and lens directly to the metal sleeve does not provide enough heat resistance when the collimator is subsequently soldered to a component housing. All early attachments were done by hand and overheating the part was simple. The epoxy in the optical path cannot tolerate severe heating and breaks down, both darkening the lightpath and impairing its bonding strength. Third, the epoxy that is in the optical path exhibits an expansion coefficient and a temperature-dependent refractive index. An insertion loss variation of 0.12 dB is reported over the temperature range of $0\text{--}80^\circ\text{C}$ [25]. Finally, the power-handling ability

of any epoxy is weak. For a mode-field diameter of $10\text{ }\mu\text{m}$ at the fiber core, the power density is $\sim 12.5\text{ MW/m}^2/\text{mW}$. Even in CW operation the epoxy is exposed to a large power density. Moreover, power transients that generate spikes will irreversibly darken the material. In view of the design goals stated in Table 5.1, few conditions are achieved with the epoxy-joint technology.

Air-Gap Collimators

Figure 5.3 illustrates the second generation of collimator assembly, an assembly that has few drawbacks and is ubiquitously used today. The three achievements of the air-gap assembly are that the in-path epoxy is removed, the index mismatch between fiber and lens to air is mitigated with AR coatings, and the overheating from soldering is reduced using an intermediate quartz sleeve. Additionally, some augmentations are made taking advantage of better technology and understanding. In particular, the lens selection today includes both GRIN and shaped lenses, and the polished facets of the ferrule and lens can be tuned to account for the fiber/lens index difference.

To assemble an air-gap collimator, an angle-polished and AR-end-coated lens is inserted into a quartz sleeve and bonded into place with a UV tack around the perimeter. The UV exposure lasts only seconds. The likewise-prepared ferrule is inserted into the other end of the quartz sleeve and adjusted to the correct location. The correct location is determined by actively monitoring the beam collimation or spot size while illuminating a target with laser light that is transmitted through the fiber and lens. The ferrule is tacked into place with UV epoxy applied around the perimeter. The sub-assembly is then coated with a heat-curing epoxy and cured. This sub-assembly is then inserted into a gold-plated metal sleeve and fixed into place with heat-curing epoxy. To finish the component and increase pull strength on the fiber end, a bead of strain-relief elastomer is added around the exposed fiber and cured in place.

The quartz sleeve has two advantages. First, the sleeve simplifies the alignment process by eliminating all degrees of freedom except the gap between the parts and the azimuth angle. Second, the quartz adds a heat-resistant layer between the metal housing and the ferrule/lens joint. Accordingly, the soldering process degrades the collimator epoxies less, especially as no UV epoxy is relied on as the primary bonding agent.

In relation to the aforementioned design goals of any collimator, the air-gap collimator answers most demands with good performance. Table 5.2 shows a favorable comparison of the air-gap technology to the epoxy-joint technology (as well as the fused-joint technology, to be detailed shortly). Also, the air-gap column favorably reads against the design goals on Table 5.1. As an example, the temperature dependence of the insertion loss is less than 0.05 dB over $0\text{--}80^\circ\text{C}$ and the air-gap collimators can tolerate the 85–85 damp-heat stress test for at least 2500 hours [25].

An important characteristic of epoxy-joint and air-gap collimators is that the pointing direction of the output beam is generally not parallel to the mechanical axis of the housing. This is because of the angle polish of the ferrule. As shown in Fig. 5.13 for a GRIN lens and Fig. 5.12 for a shaped lens, the output beam is deflected due to the offset of the input beam from the centerline of the lens. This effect is unavoidable, and in fact exacerbated for long-working-distance lenses that require large beam expansion before collimation. The specification of pointing accuracy addresses this effect. Solutions have been proposed such as laterally offsetting the ferrule from the centerline of the lens [17], adding a wedge between the ferrule and lens [19], or adjusting the polish angle of the lens to account for its index so that the refracted beam is straightened albeit offset nonetheless [5]. The last solution is defined by (5.3.1) on page 235. The first two solutions make manufacture more difficult either from creating an asymmetric part or adding another element. The third solution is probably the best to mitigate yet not eliminate the problem. However, it should be recognized that a zero pointing error enhances back reflection. Depending on the focal length and type of lens, it seems that the optimal pointing deflection is between $0.5\text{--}1.0^\circ$.

For reasons of manufacturability there is a natural division between use of GRIN and shaped lenses. The GRIN lens is suitable for short working distances because the beam does not require much expansion to counter diffraction. However, larger working distances requires a larger beam which in turn requires a larger lens diameter. The manufacturing process of a GRIN lens, typically a chemical vapor deposition process, cannot sufficiently control the quadratic curvature of the index profile nor introduce too much sag [14], which together limits the practical size of the lens. The shaped lens is in fact suitable for long working distances because the outside surface can be ground or molded to sufficient tolerances, and in fact the surface can approximate an asphere to mitigate spherical aberration. It is the short-working-distance-shaped lens that is difficult to manufacture because of the requisite small radius of curvature. The choice of lens therefore is roughly divided depending on application: short-reach collimators use GRIN lenses and long-reach collimators use aspheric lenses.

Fused-Joint Collimators

Figure 5.4 illustrates the third generation of collimator assembly, the fused-joint collimator. Unlike the preceding collimator assemblies, the fiber is not attached to a ferrule but directly fused to the lens. A glass stabilizer tube through which the fiber is threaded is used to add mechanical integrity. The advantages of a fused joint over the air-gap with AR coatings is the former's power handling ability and environmental stability. It is reported that the fused-joint collimator design can handle 10 W of optical power [13], a factor of 20 increase over the power handling ability of AR coatings. And clearly the direct fusion eliminates issues of gap change with temperature and lifetime.

The barriers to high performance of a fused-joint collimator are the ability to fuse the fiber to the lens and the index matching of the fiber mode to the lens material.

Regarding the fusion process there are several barriers: the melt points of the lens and fiber must be similar; the temperature coefficients of the lens and fiber must be similar; and enough heat must be added to the lens, having a large thermal mass, to cross the glass transition temperature while not melting away the much smaller fiber. The melting point of a typical GRIN lens is 500–600°C and that of quartz glass fiber is 1700°C. Either the GRIN material set must be changed or a shaped lens made from quartz must be used to avoid deformation. Both concepts have been demonstrated [1–4]. As with the melt point difference, similar material systems must be used for fiber and lens to avoid cracking and fracture of the fused joint during cooling.

Finally, a clever way to deliver the heat without deformation of either part is necessary for the fusion splice. One proposal is to add an intermediate coupling rod that is larger in diameter than the fiber yet smaller than the lens [22], but this adds more steps and more parts. The more elegant demonstration is with direct fiber to lens fusion using a CO₂ laser [2–4]. In the reported process, a fiber is threaded through a narrow slot cut in a mirror tailored for high reflectivity of CO₂ radiation. The mirror is tilted to 45° with respect to the back face of the lens. The laser beam is focused to the back face and heats only in proximity of the fiber placement. After a short pre-heat step the lens and fiber are fused. The fusion happens within a seconds, so active alignment is not required during the fixity. Moreover, the high temperature eliminates all contaminants yet the low amount of heat allows the front face of the lens to be AR coated in advance. Optimization of the process includes the creation of a thin (2 μm) melt layer on the lens to overcome any residual Fresnel reflection. Measurement of return loss can validate the quality of the joint.

The second barrier is the index matching of the fiber to the lens. A shaped lens can be made of fused silica without difficulty. This is the basis of some products [13]. However, a large effort has gone into the fabrication of index-matched GRIN lenses [1] with good success. These lenses are made with a plasma-enhanced chemical vapor deposition (PECVD) technology using a Ge:SiO₂ material system. The PECVD process controls the radial uniformity through rotation of the preform rod and controls the longitudinal uniformity through programmed up-and-down motion of the plasma along the rod length. After deposition the rod is consolidated at 2000°C and collapses to a radius of 0.3–1.2 mm, depending on application.

There are two more characteristics that need to be addressed for fused-joint collimators. First, even though the fiber is directly attached to the lens, the fiber must be slightly offset from center to avoid back-reflection from the front lens surface. A typical offset is 5 μm . With this offset the spot of the slightly reflected beam is offset from the fiber core. The resultant pointing deflection is typically 0.5°. Second, as all epoxy is removed from the fiber-

Table 5.2. Collimator Technology Comparison

Specification	Epoxy joint	AR-coated air gap	Fusion joint
Optical			
Low insertion loss	med ^(a)	< 0.2 dB @ 20 mm < 0.5 dB @ 150 mm ^(b,c)	< 0.2 dB @ 20 mm < 1.0 dB @ 150 mm ^(d)
High return loss	low ^(a)	> 60 dB ^(b,c)	> 55 dB, > 75 dB ^(d,e)
Pointing accuracy	med ^(a)	1.0 deg ^(c)	0.5 deg ^(d)
WDM power handling	low ^(a)	0.5 W ^(b,c)	10 W ^(d)
Spike power handling	low ^(a)	n/a	> 10 W ^(d)
Mechanical			
Temperature stability	low ^(a)	0 to +70 C ^(b,c)	-20 to +60 C ^(d)
Pull tolerance	med ^(a)	5 N ^(b,c)	high ^(d)
Pass Telcordia	low ^(a)	high ^(f)	high
Housing fixity	lens, ferrule direct attachment to housing	intermediate glass tube between lens, ferrule and housing	lens, ferrule attached directly to housing in anticipation of laser welding
Manufacturing			
UV cures	in optical path	tack	no
Heat cures	in housing	in housing	in housing
Fusion splicing	no	no	yes
AR coatings	lens output face	ferrule, two lens faces	lens output face
Active alignment	yes	yes	no

^(a) US 6,148,126 [24].

^(b) Koncent single-fiber collimator [11].

^(c) Koncent long-working-distance collimator [11].

^(d) LightPath Generation 3 collimator [13].

^(e) LightPath [3].

^(f) Casix Reliability Report [6, 7].

to-lens attachment, the insulating glass sleeve used in the air-gap collimator assembly may be eliminated for the fused-joint assembly. This reduces the diameter of the collimator which increases part density.

The third column of Table 5.2 includes specifications to compare with air-gap and epoxy collimators. While the fused-joint collimator looks attractive for a few of the aforementioned reasons, the reported performance of fused-joint and air-gap collimators is about the same save the power-handling ability.

5.2 Gaussian Optics

Gaussian optics is an analytic formalism that is a reasonable alternative to numerical ray tracing techniques when a component design is being roughed out. Gaussian optics addresses the adiabatic expansion of an optical beam as it propagates in isotropic media, covers the focusing properties of shaped and

GRIN lenses, and can estimate the mode coupling from one fiber to another. However, gaussian optics does not include lens aberration theory nor the true mode profiles of fiber-guided modes. In particular, the profile of a guided mode in single-mode fiber is a Bessel function, but that profile is approximated as a gaussian beam with a certain diffraction angle. Once the gaussian-mode envelope function is derived, subsequent augmentation produces the ABCD ray tracing matrices, which are useful for analytically tracing a paraxial ray through a cascade of optical elements.

In Chapter 1 plane wave solutions were sought for the wave equation (1.1.6) of the electric field. These solutions are valid when the source of those fields can be considered at infinity. In contrast, a mode that emerges from a fiber into free space has an aperture at the fiber/free-space boundary. The emerging field has a spherical phase front across its leading edge. As discussed in §1.2 on page 8 the vector potential is required to find field solutions from point sources or, in this case, approximations of point sources. In particular, the vector and scalar wave equations (1.2.9) govern the field evolution.

As a starting point, consider a vector potential trial solution

$$\mathbf{A}(\mathbf{r}, t) = \hat{\mathbf{n}}\psi(x, y, z) e^{j\omega t} \quad (5.2.1)$$

where $\hat{\mathbf{n}}$ is a unit vector denoting a single pointing direction and ψ is a scalar field absent the fast oscillation $\exp(j\omega t)$ term. Substitution of (5.2.1) into (1.2.9a) yields a time-harmonic scalar wave equation

$$(\nabla^2 + k^2) \psi = 0 \quad (5.2.2)$$

where the wavenumber k is defined as usual: $k^2 = \omega^2 \mu_o \varepsilon_o$. To this point the implications of the trial solution (5.2.1) have been exact. The next step establishes the paraxial approximation where change of the field is predominantly along the direction of propagation and only small changes occur in the transverse direction. The k vector is accordingly approximated as $k \simeq k_z$, or

$$\begin{aligned} k_z &= \sqrt{k^2 - (k_x^2 + k_y^2)} \\ &\simeq k - \frac{k_x^2 + k_y^2}{2k} \end{aligned} \quad (5.2.3)$$

The above approximation is called the paraxial limit of the k -vector. An envelope function u is defined after removing the fast $\exp(-jkz)$ dependence of the field:

$$\psi = u(x, y, z) e^{-jkz} \quad (5.2.4)$$

Substitution of (5.2.4) into (5.2.2) generates the paraxial wave equation

$$\nabla_T^2 u - 2jk \frac{\partial}{\partial z} u = 0 \quad (5.2.5)$$

where the approximation $\partial^2 u / \partial z^2 \ll k(\partial u / \partial z)$ eliminates terms of order $\partial^2 u / \partial z^2$.

Construction of formal solutions to (5.2.5) requires Fresnel and Fraunhofer diffraction theorems. It serves the present purposes to propose a trial solution [23] and determine *a priori* unknown constants through substitution into (5.2.5). For a fundamental gaussian mode, rather than higher-order Hermite gaussian modes, the trial solution to the scalar envelope is

$$u = u_o \exp \left[-j \left(p(z) + k(x^2 + y^2)/2q(z) \right) \right] \quad (5.2.6)$$

For the interested reader, the component $p(z)$ comes from the Fraunhofer diffraction integral and the component $k(x^2 + y^2)/2q(z)$ comes from the Fresnel kernel [10, 12]. Substitution of (5.2.6) into (5.2.5) generates the parametric equation

$$-2k \left[p'(z) + \frac{j}{q} \right] + \frac{k^2(x^2 + y^2)}{q^2(z)} \left[q'(z) - 1 \right] = 0$$

where primes denote differentiation with respect to z . Solutions to this parametric equation are

$$p' = -\frac{j}{q}, \quad \text{and} \quad q' = 1$$

Integration of $q' = 1$ generates $q(z) = z + c_a$, where c_a is a constant of integration yet to be determined. With this general solution, $p'(z)$ is integrated. The two general solutions are

$$p(z) = -j \ln(z + c_a) + c_b, \quad \text{and} \quad q(z) = z + c_a$$

Choosing c_a as real only produces an offset of $q(z)$. Alternatively, choosing c_a as imaginary introduces an orthogonal coordinate that can be used to scale the field solutions to the initial aperture. With the solution

$$q(z) = z + jb \quad (5.2.7)$$

the parameter $p(z)$ is determined as

$$p(z) = -\tan^{-1} \left(\frac{z}{b} \right) - j \ln \sqrt{1 + \left(\frac{z}{b} \right)^2}$$

Pulling these parameters together, the unnormalized envelope solution is

$$\begin{aligned} u(x, y, z) = & \frac{u_o}{\sqrt{1 + (z/b)^2}} \exp(j \tan^{-1}(z/b)) \times \\ & \exp \left(-j \frac{kz(x^2 + y^2)}{2(z^2 + b^2)} \right) \exp \left(-\frac{kb(x^2 + y^2)}{2(z^2 + b^2)} \right) \end{aligned} \quad (5.2.8)$$

The expression for b is determined by identifying the e^{-2} power decay of the mode at $z = 0$ (which is e^{-1} decay of the field):

$$u(x, y, 0) = u_o \exp \left(-\frac{k(x^2 + y^2)}{2b} \right) \quad (5.2.9)$$

where in the transverse direction

$$\frac{k(x^2 + y^2)}{2b} = 1$$

or

$$b = \frac{k w_o^2}{2} = \frac{\pi w_o^2}{\lambda} \quad (5.2.10)$$

The parameter b is called the confocal parameter or the Rayleigh length and is related to the minimum beam radius w_o according to the above equation. The unit of b is length. In order to complete the trial envelope solution the field must be normalized. Normalization can be calculated at any point z since there is no loss or gain in the system. Normalization of $u(x, y, z)$ such that

$$\iint_{-\infty}^{\infty} |u(x, y, 0)|^2 = 1$$

yields $u_o = \sqrt{k/\pi b}$. Pulling together all the pieces, the envelope solution to the scalar wave equation is

$$u(x, y, z) = \sqrt{\frac{2}{\pi w^2}} \exp(j\phi) \exp \left(-\frac{x^2 + y^2}{w^2} \right) \exp \left(-j \frac{k(x^2 + y^2)}{2R} \right) \quad (5.2.11)$$

where the mode parameters are defined as

$$w^2(z) = w_o^2 \left(1 + \frac{z^2}{b^2} \right) \quad (5.2.12)$$

$$\frac{1}{R(z)} = \frac{z}{z^2 + b^2} \quad (5.2.13)$$

$$\tan \phi = \frac{z}{b} \quad (5.2.14)$$

Before the various definitions are discussed in detail, note that the single parameter $q(z)$ completely determines the behavior of the beam. This will be important in the following sections.

Symbol $w(z)$ is the e^{-2} waist radius (in power) along z ; $R(z)$ is the radius of the phase-front, or field, curvature of the mode along z ; and ϕ is the common phase of the mode (Fig. 5.5). These parameters have two solutions in the extreme. First, at $z = 0$, the waist does not get smaller than diameter $2w_o$. This is in contrast to a ray-optic model, where paraxial rays cross on a focal plane, indicating a zero beam waist. A gaussian mode always has a non-zero minimum waist. Moreover, the radius of the phase-front R is infinity at $z = 0$. There is no phase curvature, the beam is a plane wave at this position, and the mode is in focus on the plane. Second, in the far field, the waist and phase-front radius approach asymptotic limits: they both grow linearly with z . In

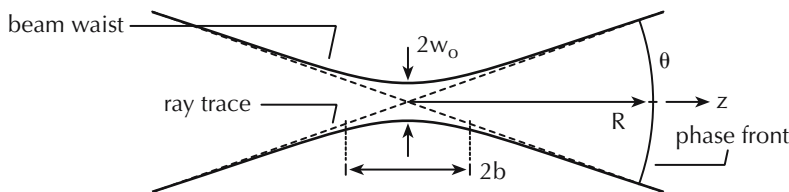


Fig. 5.5. A gaussian mode passing through a waist minimum. The field curvature is governed by $R(z)$ and the e^{-2} beam waist is governed by $w(z)$. In the far field the adiabatic expansion of the beam waist falls within the diffraction angle $\theta \simeq \lambda/\pi w_o$. The smaller the minimum waist w_o the larger the beam divergence. Also, the depth of focus, $2b$, is the length between points where the beam waist is $\sqrt{2}w_o$.

particular, the beam waist expands in a cone whose half-angle is called the diffraction angle. The diffraction angle θ is calculated from $w(z)/z$, or

$$\tan \theta = \frac{\lambda}{\pi w_o} \quad (5.2.15)$$

Clearly the diffraction angle increases as the minimum beam waist, or the aperture a collimated beam passes through, decreases. The level of approximation used herein is that the aperture be at least a wavelength large.

The numerical aperture also describes the half-angle of the cone within which a beam of light adiabatically expands, but the diffraction angle θ and numerical aperture are not precisely the same. The numerical aperture is defined as

$$\text{N.A.} = n \sin \theta_{\text{na}} \quad (5.2.16)$$

where θ_{na} is the angle the marginal ray in a ray-optic formalism takes when propagating away from a point source. The diffraction angle is defined for a beam waist at e^{-2} , where 13.5% of the optical power is accordingly excluded. The marginal ray for a sufficiently large collecting lens covers “all” the optical power and would therefore trace a larger angle. For example, Corning reports a numerical aperture of its SMF-28 fiber of 0.14 and a mode-field diameter of $10.4 \pm 0.8 \mu\text{m}$ [9]. These two quantities are directly measured using the procedure referenced in [8]. Asserting $w_o \sim 5.2 \mu\text{m}$ yields a diffraction angle of $\theta = 0.094$ rad, which is less than the measured numerical aperture. Conversely, asserting $\theta = 0.14$ rad yields a beam diameter of $7.0 \mu\text{m}$. Similar results are obtained either way, but the point has been made that the diffraction angle and numerical aperture are similar but not the same.

Another parameter derived from (5.2.12) is the depth of focus. The depth of focus is the full length about a waist minimum where the waist crosses through $\sqrt{2}w_o$. In fact, the depth of focus is just twice the confocal parameter: $2b$. Of importance is that the depth of focus grows as the square of the minimum beam waist (cf. (5.2.10)). A doubling of the minimum waist quadruples the depth of focus, allowing the optical “throw” between lenses to become large.

To conclude this section, the functional form of the electric field profile is derived from the vector potential. With the vector potential defined by (5.2.1), the scalar potential in time-harmonic form is found from (1.2.8):

$$\Phi = \frac{j}{\omega\mu_o\varepsilon_o} \nabla \cdot \mathbf{A} \quad (5.2.17)$$

The electric field (1.2.5) in time-harmonic form is therefore

$$\mathbf{E} = -j\omega \left(\mathbf{A} + \frac{1}{\omega^2\mu_o\varepsilon_o} \nabla \nabla \cdot \mathbf{A} \right) \quad (5.2.18)$$

Suppose for instance that $\hat{\mathbf{n}} = \hat{x}$. The electric field (5.2.18) has a primary vector component in the $\hat{x}$ direction, but also a weak component in the longitudinal, or $\hat{z}$, direction. This is in contrast to a plane wave, where the electric field components line only the a plane perpendicular to the direction of propagation. For gaussian beams, with a spherical phase front as illustrated in Fig. 5.5, the electric field at a point off of the z -axis clearly has a longitudinal component.

5.2.1 q Transformation and ABCD Matrices

The gaussian beam analysis of the preceding section provides a functional form to determine parameters such as beam expansion, minimum beam waist, and field curvature. The next step addresses the behavior of a gaussian beam through a lens, a dielectric block, and more generally through a cascade of on-axis optical elements. The following analysis develops what is called the ABCD matrix formalism to calculate the behavior of a gaussian beam through a system. The reader should note that the ABCD formalism has limitations, including the absence of wavefront aberrations and the inability to directly calculate passage through a wedge. Following the warning at the introduction of the preceding section, ABCD matrices, like gaussian beam analysis, provides only an analytic framework with which to rough out a design; a ray tracing and beam propagation software program should be used to confirm and tolerance any manufactured component design.

Recall that the $q(z)$ parameter completely describes a gaussian beam. Accordingly, it is sufficient to track the transformation of q from one position to another, through one interface to another. The q parameter is defined by (5.2.7). The inverse q parameter can easily be related to the beam waist w and phase-front radius R :

$$\frac{1}{q} = \frac{1}{R} - j\frac{2}{kw^2} \quad (5.2.19)$$

First consider propagation over length d . The q parameter is transformed to q' as

$$q' = q + d$$

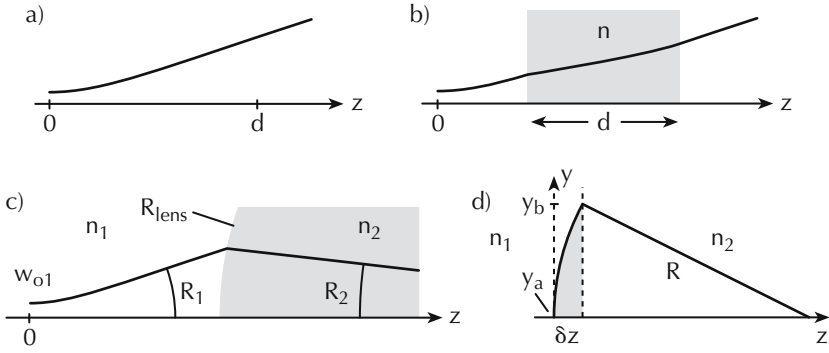


Fig. 5.6. Gaussian mode transformations. a) Adiabatic expansion over length d . b) Passage through a dielectric block of index n and length d having input and output faces perpendicular to the beam direction. c) Passage through a spherical surface into dielectric medium of index n_2 from medium of index n_1 . d) Geometry for phase difference calculation.

The beam semi-profile is illustrated in Fig. 5.6(a).

Next consider refraction through a dielectric block having index n and length d (Fig. 5.6(b)). There are two flat interfaces, both perpendicular to the direction of propagation. At the first interface the initial wavenumber k is changed to nk . At the second interface nk changes back to k . The change in wavenumber needs to be accounted for in the gaussian envelope equation (5.2.11). Consider the input interface. First, the confocal parameter changes within the medium: $b \rightarrow nb$. Second, the mode waist is continuous across the interface, so $w(z_-) = w(z_+)$. To maintain the same mode waist while accommodating the change of the confocal parameter, z must transform to nz . The q parameter thus transforms as

$$q' = nq \quad (5.2.20)$$

The resulting change of the mode envelope function within the medium produces a reduction of the field curvature. Within the medium the expansion of the mode is diminished by the factor n . However, expansion continues unabated because no curved surface or transverse index gradient has been encountered. Finally, the change at the output face is reversed from that at the input face, leading to the q transformation:

$$q' = q/n \quad (5.2.21)$$

Next consider the passage through a spherical lens surface from index n_1 to index n_2 (Fig. 5.6(c,d)). The surface is oriented such that the z -axis crosses its center. As detailed in 5.6(d), the phases accrued at lateral coordinates y_a and y_b while traveling along z are

$$\phi_a = kn_2\delta z, \text{ and } \phi_b = kn_1\delta z$$

Note that the wavenumber k is written for vacuum. The surface, being spherical, is described by $x^2 + y^2 + z^2 = R_s^2$, where R_s is the physical curvature. The difference in phase between an axial ray and a ray off-axis is given by the equation

$$\delta\phi = k(x^2 + y^2) \frac{(n_2 - n_1)}{2R_s} \quad (5.2.22)$$

The differential phase delay of the surface imparts a curvature on the gaussian mode, leaving a new field curvature R' . That curvature is

$$\frac{1}{R'} = \frac{1}{R} - \frac{n_2 - n_1}{R_s} \quad (5.2.23)$$

Notice that in the absence of a refractive index discontinuity at the surface there is no change of field curvature. In order to arrive at the correct q transformation, care must be taken to recognize that the mode first travels in n_1 and passes to n_2 . Applying (5.2.20-5.2.21) and (5.2.23) yields

$$\frac{1}{q'} = \left(\frac{n_1}{q} - \frac{(n_2 - n_1)}{R_s} \right) \frac{1}{n_2}$$

All of the above transformations are instances of the bilinear transformation. A bilinear transformation is a transformation in the complex plane to coordinate q' from coordinate q via

$$q' = \frac{Aq + B}{Cq + D} \quad (5.2.24)$$

where coefficients $A - D$ are real numbers. The four bilinear coefficients can be grouped in 2×2 matrix form. This matrix is called the ABCD matrix. In the case of uninterrupted propagation, the ABCD matrix is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \quad (5.2.25)$$

For transformations through a flat interface into and out of a dielectric medium having index n , the ABCD matrices are

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/n \end{pmatrix}, \text{ and } \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix}, \quad (5.2.26)$$

respectively. Finally, transformation through a spherical surface produces

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{(n_2 - n_1)}{n_2 R_s} & \frac{n_1}{n_2} \end{pmatrix} \quad (5.2.27)$$

The matrix operation is carried out as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} q_a \\ q_b \end{pmatrix} = \begin{pmatrix} q'_a \\ q'_b \end{pmatrix}$$

where subsequently $q' = q'_a/q'_b$.

Needless to say, armed with a matrix formalism, an arbitrary cascade of spherical lens surfaces, planar index discontinuities, and free-space propagation can be calculated and the resultant q parameter determined. This formalism will be used in the following to calculate collimator to collimator designs.

5.2.2 ABCD Ray Tracing

To the current level of approximation, ray tracing is derived from the gaussian mode in the limit of infinite beam waist. In this limit, there is no field curvature and the waves are planar. The q parameter becomes

$$\lim_{w \rightarrow \infty} \frac{1}{q} = \frac{1}{L}$$

where the symbol L replaces R since the phase-front radius is infinite. Rather, L is the length of the ray between two planes along z . Note that the q parameter is now purely real. Even with the elimination of the field curvature, the paraxial limit remains. In particular

$$\begin{aligned} \sin \theta &= y/L \simeq \theta & \cos \theta &= z/L \simeq 1 \\ \tan \theta &= y/z \simeq \theta & y' &= \Delta y / \Delta z \end{aligned}$$

With these approximations, the ray length from one boundary plane to another is

$$L = \frac{y}{y'} \tag{5.2.28}$$

The bilinear q transformation can also be rewritten for the plane wave where L replaces q :

$$q' = \frac{Aq + B}{Cq + D} \implies L' = \frac{AL + B}{CL + D}$$

or, using (5.2.28),

$$L' = \frac{Ay + By'}{Cy + Dy'} \tag{5.2.29}$$

With this transition to ray optics from gaussian optics in the paraxial limit, the use of the ABCD matrices (5.2.25–5.2.27) remains valid. Figure 5.7 illustrates a ray trace from an object to an image through a thick lens.

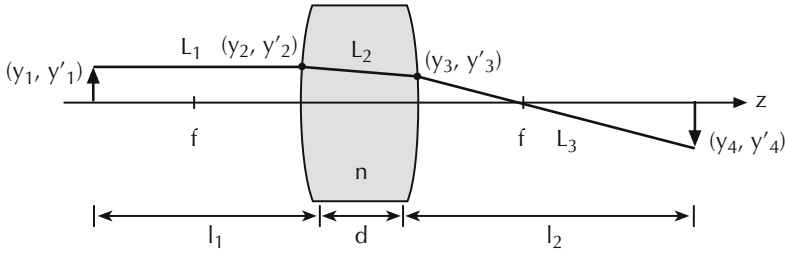


Fig. 5.7. Ray trace from object to image through a thick lens. The length of the ray $L_{1...3}$ between each boundary plane is indicated. The coordinate (y, y') at the intersection of the ray with each boundary plane completely describes the trace of the ray.

5.2.3 Action of a Single Lens

Figure 5.8 illustrates a thin symmetric-convex lens of index n with an object positioned at A' (distance l_1 from the lens) and an image at C' (distance l_2 from lens). The object and image heights are y_1 and y_2 , respectively. The ray trace concatenation from object to image is

$$\bar{\bar{A}} = \begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{(n-1)}{R_{\text{lens}}} & n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{(n-1)}{nR_{\text{lens}}} & \frac{1}{n} \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ 0 & 1 \end{pmatrix} \quad (5.2.30)$$

The center two matrices transform through the first and second spherical surfaces. Combined, they yield the focal length of the lens:

$$\begin{pmatrix} 1 & 0 \\ -\frac{(n-1)}{R_{\text{lens}}} & n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{(n-1)}{nR_{\text{lens}}} & \frac{1}{n} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}$$

where the inverse focal length is defined as

$$\frac{1}{f} = \frac{2(n-1)}{R_{\text{lens}}} \quad (5.2.31)$$

Thus the focal length is related to the lens curvature. As the lens was defined as a symmetric-convex lens, both surfaces refract the ray. The optical power $\mathcal{O}$ of a single spherical surface is defined as

$$\mathcal{O} = \frac{(n_2 - n_1)}{R_{\text{lens}}} \quad (5.2.32)$$

Optical power is the ability of an interface to change the field curvature of a gaussian beam. A flat interface has an infinite radius and therefore no optical power. A curved interface with zero index discontinuity likewise has no optical

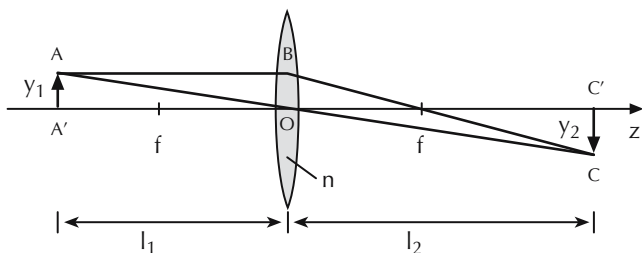


Fig. 5.8. Ray trace from object to image through a thin lens. Lengths l_1 and l_2 are related by the focal length of the lens. The focal length f is $f = R/2n$.

power. The lens surfaces described in this and the preceding sections do have optical power, and in anticipation of the derivation of the GRIN lens equation below, a flat interface having an lateral index gradient also has optical power.

Returning to the concatenation (5.2.30), the product is

$$\bar{\bar{A}} = \begin{pmatrix} 1 - l_2/f & l_1 + l_2 - l_1 l_2/f \\ -1/f & 1 - l_1/f \end{pmatrix} \quad (5.2.33)$$

On the image plane, all rays that emanate from the object converge to the same point at the image. This is illustrated in Fig. 5.8, where chief ray $\overline{AOC}$ and paraxial ray $\overline{ABC}$ both converge at point C . As this is a thin lens, the chief ray transits through the center of the lens, where there is no curvature, without a change of its direction. The paraxial ray travels straight to the lens and then alters slope so as to pass through the focal point f . In either case, the final position is independent of the ray slope y' at the object. Therefore $B = 0$. Application of this condition on (5.2.33) generates the formula for a thin lens:

$$\frac{1}{l_1} + \frac{1}{l_2} = \frac{1}{f} \quad (5.2.34)$$

In terms of the q parameter, in the object plane $1/R = 0$ and $q = -jb$. The matrix (5.2.33) yields

$$\frac{1}{q'} = \frac{1}{R'} - j\frac{1}{b'} \quad (5.2.35)$$

where

$$R' = \frac{f}{l_1/l_2}, \quad \text{and} \quad b' = \left(\frac{l_2}{l_1}\right)^2 b \quad (5.2.36)$$

Consider three cases: the object is placed behind, at, and in front of the front focal plane, the front focal plane being on the side of the object. In the first instance, the lens focuses the object onto the image plane. The image plane is a finite distance l_2 from the lens and the image is magnified by

$$M = \left| \frac{l_2}{l_1} \right| \quad (5.2.37)$$

The magnification can be greater or less than unity, depending on the relative positions of object and lens. Likewise, the gaussian beam waist is magnified as indicated by b' in (5.2.36): $M = w'_o/w_o$.

In the second instance, $l_1 = f$ and the lens collimates the light from the object. In the paraxial limit, all rays passing through the lens subsequently run parallel to one another. The ABCD matrix (5.2.33) for this condition yields only one meaningful relation, which is the inclination angle of the collimated beam as a function of offset on the back focal plane:

$$\theta_{\text{pt}} = -\frac{y}{f} \quad (5.2.38)$$

where, in anticipation of what is to follow, the inclination angle is denoted θ_{pt} for the pointing direction. A simple lens transforms positional offset on the focal plane to angle of the collimated beam. This transformation property is the basis of much Fourier optic filtering work as well as some telecommunications components [20].

The final instance is when $l_1 < f$. In this case the curvature of the lens is not sufficient to counter the adiabatic expansion of the mode; the mode will continue to expand. However, a virtual image is formed on the same side of the lens as the object, at location $-l_2$. To an observer on the far side of the lens, the object will appear located at the virtual image position.

5.2.4 Action of a GRIN Lens

A GRIN lens substitutes the optical power of a curved surface across an index discontinuity with the optical power of an index gradient in the plane perpendicular to the propagation axis. A canonical index profile is

$$n(x, y) = n_o \left(1 - \frac{x^2 + y^2}{2p^2} \right) \quad (5.2.39)$$

where n_o is the axial index of the lens and p is the curvature of the index gradient having units of length. The phase retardation across an infinitesimal slab Δz , determined in the same way (5.2.22) was derived, is

$$\delta\phi = k \frac{(x^2 + y^2)}{2p^2} n_o \Delta z$$

The focal length of this infinitesimal slab is identified as

$$\frac{1}{f_\Delta} = \frac{n_o \Delta z}{p^2}$$

A single slab is described by a cascade of three ABCD matrices: a first matrix that propagates $\Delta z/2n_o$, a second matrix that focuses with focal length f_Δ , and a final matrix that again propagates $\Delta z/2n_o$. Keeping terms to order $(\Delta z)^2$, the resulting concatenation is

$$\bar{\bar{A}} = \begin{pmatrix} 1 - \frac{\Delta z^2}{2p^2} & \frac{\Delta z}{n_o} \\ -\frac{n_o \Delta z}{p^2} & 1 - \frac{\Delta z^2}{2p^2} \end{pmatrix} \quad (5.2.40)$$

The determinant of (5.2.40) is

$$\det(\bar{\bar{A}}) = 1 + \left(\frac{\Delta z}{\sqrt{2}p} \right)^4$$

Only in the limit $\Delta z \rightarrow 0$ is the concatenation matrix $\bar{\bar{A}}$ unitary.

A grin lens having length L is divided into n sections of length Δz so that $L = n\Delta z$. The limits $\Delta z \rightarrow 0$ and $n \rightarrow \infty$ are then taken. In order to carry out these limits, matrix (5.2.40) is separated into its eigenvectors and eigenvalues as in (2.6.34). When $\bar{\bar{A}}$ is cascaded n times, the inner eigenvector matrices telescope and leave only

$$\bar{\bar{A}}^n = V \Lambda^n V^\dagger \quad (5.2.41)$$

where V is a square matrix whose columns are the eigenvectors of $\bar{\bar{A}}$ and Λ is a diagonal matrix of the corresponding eigenvalues. The eigenvectors and eigenvalues of $\bar{\bar{A}}$ are

$$V_\pm = \begin{pmatrix} \pm \frac{jp}{n_o} \\ 1 \end{pmatrix}, \quad \text{and} \quad \lambda_\pm = \left(1 - \frac{\Delta z^2}{2p^2} \right) \mp j \left(\frac{\Delta z}{p} \right)$$

Notice that Δz enters only in the eigenvalues and not the vectors; only the eigenvalues get integrated so that Δz will become length L . Retaining only terms of order Δz in $\lambda_\pm$ and recalling the limit expression for the exponential function, the cascaded eigenvalues take the form

$$\begin{aligned} \lim_{n \rightarrow \infty} (\lambda_\pm)^n &= \lim_{n \rightarrow \infty} \left(1 \mp j \frac{L}{np} \right)^n \\ &= \exp(\mp j \theta_g) \end{aligned}$$

where the GRIN lens angle is

$$\theta_g = \frac{L}{p} \quad (5.2.42)$$

Denoting $\bar{\bar{A}}^n$ as $\bar{\bar{A}}_L$, the ABCD matrix of a GRIN lens, derived from (5.2.41), is

$$\bar{\bar{A}}_L = \begin{pmatrix} \cos \theta_g & \frac{p}{n_o} \sin \theta_g \\ -\frac{n_o}{p} \sin \theta_g & \cos \theta_g \end{pmatrix} \quad (5.2.43)$$

This is an unusual governing equation because of its periodicity. Every $\theta_g = 2\pi$ yields a replication of the object in real space. Half of this length generates an

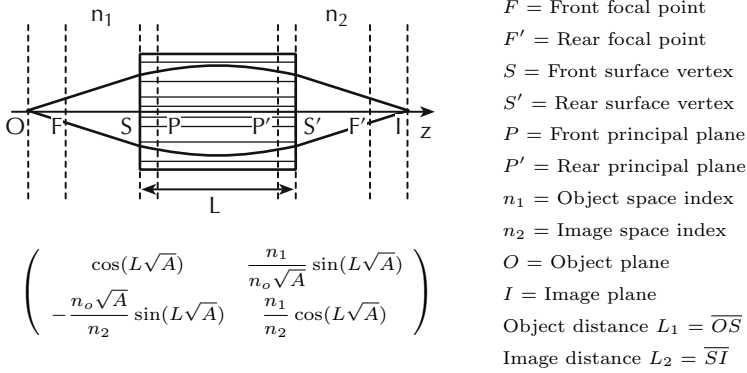


Fig. 5.9. GRIN lens reference planes and definitions (after [15]). Scale-factor $\sqrt{A}$ is the index gradient constant and has units of m^{-1} .

object inversion. An one-quarter of this length will collimate a point source. Along a long GRIN rod the light rays are bound and trace a periodic longitudinal pattern.

In order to maintain consistency with the industry norm, the symbol $\sqrt{A}$ replaces p and is called the index gradient constant, having units of m^{-1} . Figure 5.9 shows the matrix form with industry notation. The pitch of a GRIN lens P is defined as

$$P = \frac{L\sqrt{A}}{2\pi} \quad (5.2.44)$$

When a GRIN rod is cut to a quarter pitch, $P = 0.25$, and the rod is placed in air, the focal length of the lens is

$$\frac{1}{f_{1/4}} = n_o\sqrt{A} \quad (5.2.45)$$

This relation makes it clear that the stronger the index gradient, the higher the effective curvature of the lens. As a matter of common practice, a GRIN collimator has a pitch of $P = 0.23$; the purpose of this reduced pitch is to pull the front focal plane away from the physical face of the rod so that a fiber ferrule can be located in proximity to the lens without touching it. Table 5.3 provides the paraxial design equations [15] for a GRIN lens immersed in differing media. The locations of the reference planes are indicated in Fig. 5.9.

5.2.5 Some Limitations of the ABCD Matrix

Care must be taken when applying the ABCD matrix formalism to practical systems. The formalism is derived for a gaussian mode as it travels through a cascade of refractive indices and surfaces having optical power. Changes in ray direction originate from a change in the modal wavenumber k and

Table 5.3. GRIN Lens Paraxial Equations

Parameter	Function
Front focal length	$\overline{FS} = \frac{n_1 \cos(L\sqrt{A})}{n_o \sqrt{A} \sin(L\sqrt{A})}$
Effective front focal length	$\overline{FP} = \frac{n_1}{n_o \sqrt{A} \sin(L\sqrt{A})}$
Rear focal length	$\overline{S'F'} = \frac{n_2 \cos(L\sqrt{A})}{n_o \sqrt{A} \sin(L\sqrt{A})}$
Effective rear focal length	$\overline{P'F'} = \frac{n_2}{n_o \sqrt{A} \sin(L\sqrt{A})}$
Front principal distance	$\overline{SP} = \frac{n_1 1 - \cos(L\sqrt{A}) }{n_o \sqrt{A} \sin(L\sqrt{A})}$
Rear principal distance	$\overline{P'S'} = \frac{-n_2 1 - \cos(L\sqrt{A}) }{n_o \sqrt{A} \sin(L\sqrt{A})}$
Magnification	$M = \frac{n_1}{n_1 \cos(L\sqrt{A}) - n_o L \sqrt{A} \sin(L\sqrt{A})}$
Angular magnification	$M_a = \frac{n_1 \cos(L\sqrt{A}) - n_o L \sqrt{A} \sin(L\sqrt{A})}{n_2}$

changes in the field curvature. Absent is the ability to change direction of the mode through a wedge or a prism. According to (5.2.29), the ABCD matrix transforms (y_a, y'_a) to (y_b, y'_b) . One is generally more accustomed to transformations on coordinates (y, z) . In the latter case rotation matrices can be applied to rotate the coordinate system through flat, inclined refractive surfaces. But this does not apply to the ABCD formalism.

As an example, consider the propagation over length d through air and through a medium with index n . The initial ray comes from an axially located point source, i.e. $y_a = 0$. The two cases are

$$\begin{pmatrix} dy'_a \\ y'_a \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ y'_a \end{pmatrix} \quad (5.2.46a)$$

$$\begin{pmatrix} dy'_a/n \\ y'_a \end{pmatrix} = \begin{pmatrix} 1 & d/n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ y'_a \end{pmatrix} \quad (5.2.46b)$$

Using Snell's law, one expects the light ray to refract into the medium and thereby change its inclination. However, (5.2.46) does not directly indicate that the ray inclination has changed: the y' entry in both output vectors is the same. The way refraction is handled in the ABCD formalism is to compress the offset component to y/n from y .

This limitation poses problems for the analysis of most collimators. As discussed in §5.1 on page 213, epoxy-joint and air-gap collimator assemblies incline the ferrule and lens facets to minimize back reflection. How is the inclination to be handled? Ordinarily Snell's law is applied at the interface

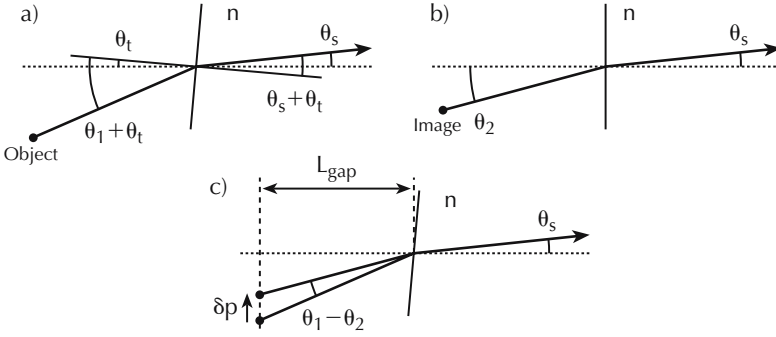


Fig. 5.10. Image location correction for a gaussian mode through an inclined face. a) Object point inclined by $\theta_1 + \theta_t$ to interface normal, the interface being tilted by θ_t . Refraction angle is θ_s . b) Image angle θ_2 such that θ_s is the same as in a). c) Over fixed gap L_{gap} image point is offset δp from object point.

and a new paraxial angle y' is determined. However, that is incompatible with the ABCD formalism. To fix this problem, an image point is created as a complement to the original object point. When a ray emanating from the image point transits a perpendicular interface (having the same refractive step as the inclined face), the refraction angle with respect to the horizontal (or other fixed reference) is the same as that created by a ray from the object point which transits the inclined face. On the refracted-ray side, one cannot distinguish whether the ray comes from the object or image point.

Figure 5.10 illustrates the required analysis. All calculations are taken in the paraxial limit of small angles. In Figs. 5.10(a,b) the refracted angles are $(\theta_1 + \theta_t) = n(\theta_s + \theta_t)$ and $\theta_2 = n\theta_s$, respectively, where θ_t is the tilt angle of the facet. The relation between θ_1 and θ_2 for a fixed θ_s is

$$\theta_2 = \theta_1 - (n - 1)\theta_t \quad (5.2.47)$$

The position offset over a fixed gap length L_{gap} is

$$\delta p = L_{\text{gap}}(\theta_2 - \theta_1) \quad (5.2.48)$$

This offset is illustrated in Fig. 5.10(c). Armed with these position and angular corrections, the ABCD formalism may be applied to inclined-facet collimators.

5.3 Select Collimators Analyzed with the ABCD Matrix

Four collimator examples are presented to highlight the analyses of the preceding sections. The first example is shown in Fig. 5.12 and uses a shaped lens. The optical data are detailed in the caption. This collimator is an air-gap collimator where the ferrule and lens rod are angle polished. As proposed

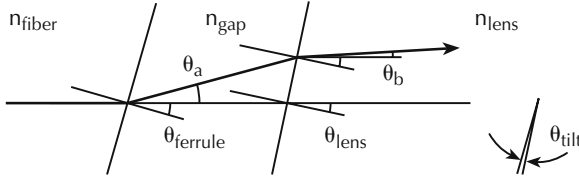


Fig. 5.11. Ray trace to determine tilt angle θ_t such that the reentrant beam runs parallel to the mode in the fiber given that the fiber and lens refractive indices differ.

in [5], the angle of the lens can be tailored with respect to the ferrule angle so that the beam that enters the lens runs parallel to the axis of the fiber even when the lens and fiber refractive indices differ. This detail is important when trying to minimize the pointing error of an air-gap collimator and can be applied to either a GRIN or shaped lens. Figure 5.11 illustrates the relevant calculation. The known parameters are the fiber and lens indices and the ferrule facet angle. In the small angle limit, the angle of the beam as it emerges from the ferrule, with respect to the fiber axis (horizontal), is

$$\theta_a = (n_{\text{fiber}} - 1)\theta_{\text{ferrule}}$$

where the index of the air gap is taken as unity. The angle of the beam after refraction into the lens with respect to the horizontal is

$$\theta_b = (\theta_a + \theta_{\text{lens}})/n_{\text{lens}} - \theta_{\text{lens}}$$

The goal is to have the reentrant beam run parallel to the fiber axis: $\theta_b = 0$. The lens facet angle with respect to the ferrule facet angle is therefore

$$\theta_{\text{lens}} = \left(\frac{n_{\text{fiber}} - 1}{n_{\text{lens}} - 1} \right) \theta_{\text{ferrule}} \quad (5.3.1)$$

The tilt angle θ_t , the angle difference between ferrule and lens, is

$$\theta_t = \left(\frac{n_{\text{lens}} - n_{\text{fiber}}}{n_{\text{lens}} - 1} \right) \theta_{\text{ferrule}} \quad (5.3.2)$$

For the example shown in Fig. 5.12, the ferrule and lens angles are 8° and 6.7° , respectively. The refraction from the object point through the wedge facet and the refraction from the image point through the planar facet produce the same ray trace. The image ray trace, which accounts for the faceting of the lens and calculated via (5.2.47-5.2.48), is shown in both figures and detailed in the lower figure. The pointing direction of the collimated beam is downward due to the upward offset of the central ray accrued in the gap. While the collimator in the figure shows a substantial gap between ferrule and shaped lens, it is clear that relocation of the ferrule immediately behind the lens facet and re-optimization of the lens focal length can minimize, albeit not eliminate, the pointing error.

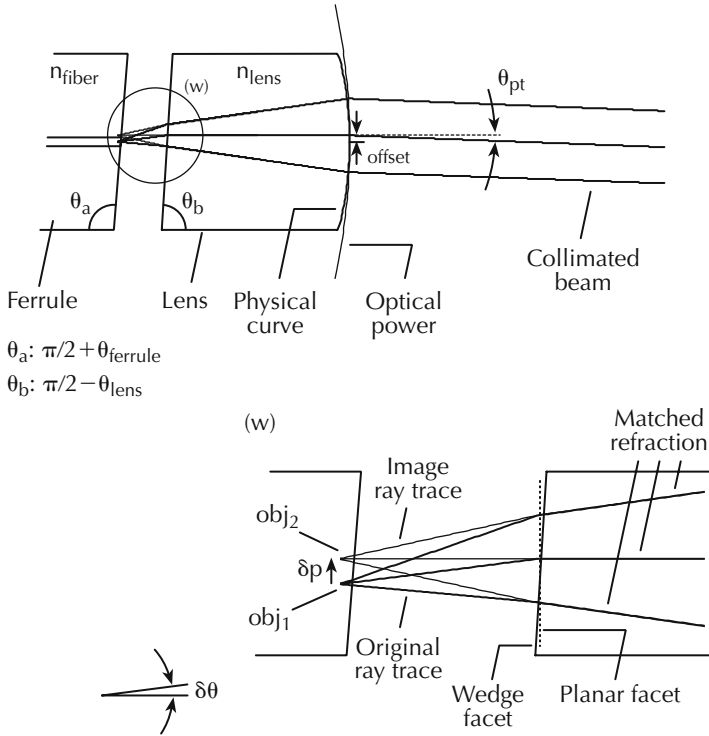


Fig. 5.12. Scale drawing of central and paraxial rays emergent from an SMF-28 fiber and captured by a shaped lens. The lens collimates the beam. The surface of optical power is superimposed over the physical curvature of the lens. The lens facet is designed to straighten the central ray after refraction. The parameters are: $n_{\text{fiber}} = 1.46$, $\theta_{\text{ferrule}} = 8.0^\circ$, $\text{N.A.} = 0.11$, $n_{\text{lens}} = 1.55$, $L_{\text{lens}} = 2.0$ mm, $R_{\text{lens}} = 1.0$ mm, $\theta_{\text{lens}} = 6.7^\circ$, $L_{\text{gap}} = 0.53$ mm, $\text{offset} = 0.0$ mm, $\text{image offset} = +0.033$ mm, $\theta_{\text{pt}} = -1.07^\circ$, vertical scale = $2\times$.

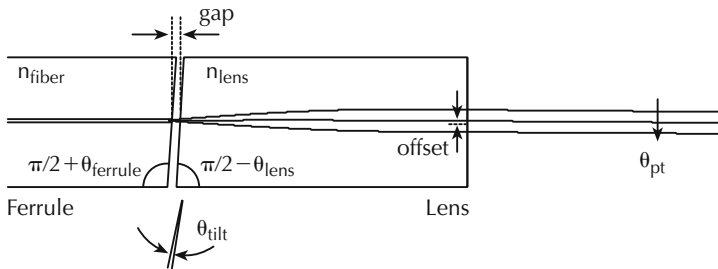


Fig. 5.13. Scale drawing of central and paraxial rays emergent from an SMF-28 fiber and captured by an NSG SLS2 lens [16]. The parameters are: $n_{\text{fiber}} = 1.46$, $\theta_{\text{ferrule}} = 8.0^\circ$, $\text{N.A.} = 0.11$, $n_{\text{lens}} = 1.5503$, $L_{\text{lens}} = 6.10$ mm, $P = 0.23$, $\sqrt{A} = 0.237$ mm $^{-1}$, $\text{EFL} = 2.743$ mm, $\theta_{\text{lens}} = 6.0^\circ$, $L_{\text{gap}} = 0.343$ mm, $\text{offset} = 0$ mm, $\text{image offset} = +0.020$ mm, $\theta_{\text{pt}} = -0.25^\circ$, vertical scale = $2\times$.

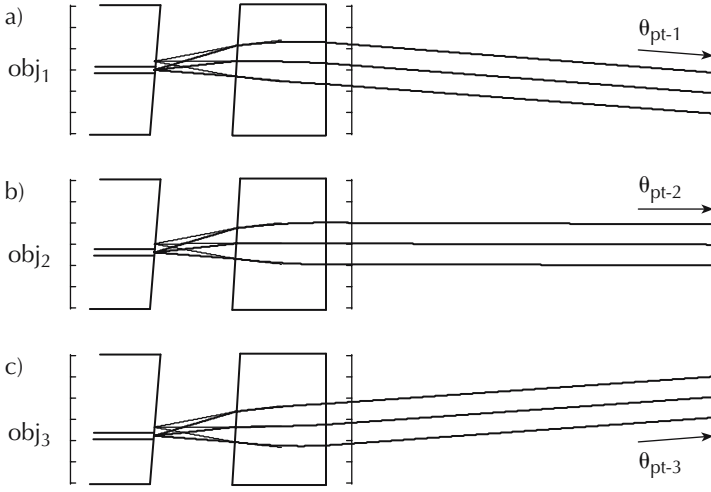


Fig. 5.14. Scale drawing of central and paraxial rays emergent from an SMF-28 fiber and captured by a long-reach GRIN lens [17]. Lateral offset of the ferrule can reduce the pointing error. The parameters are: $n_{\text{fiber}} = 1.46$, $\theta_{\text{ferrule}} = 8.0^\circ$, N.A. = 0.11, $n_{\text{lens}} = 1.5902$, $L_{\text{lens}} = 2.324$ mm, $P = 0.119$, $\sqrt{A} = 0.322$ mm $^{-1}$, EFL = 2.870 mm, $\theta_{\text{lens}} = 6.0^\circ$, $L_{\text{gap}} = 2.10$ mm, offset = 0, -125, -250 μm , image offset = +0.130 mm, $\theta_{\text{pt}} = -2.59, +0.10, +2.39^\circ$, vertical scale = $9.5\times$.

The second example is shown in Fig. 5.13 and uses a GRIN lens. The optical data are detailed in the caption and the design follows that of [16]. The ferrule and lens facets are angle polished to minimize back reflection and while the angles differ they do not exactly straighten the central ray. As with the shaped lens, the image method was used to calculate the beam refraction through the GRIN facet. The GRIN lens changes the wavefront curvature continuously throughout the body of the rod until collimation is achieved; the pitch of this lens is $P = 0.23$. Even with the small gap there is a downward pointing direction due to the lateral offset of the beam.

The third example is shown in Fig. 5.14 and uses a long-reach GRIN lens. The example follows that of [17] and is a study of pointing direction as a function of lateral ferrule offset. The optical data are detailed in the caption. A long working distance lens requires a large beam expansion to overcome diffraction. In this example the expansion occurs predominantly in the air gap. Accordingly, the ray bundle that impinges on the back face of the lens is substantially offset resulting in a large pointing error. Progressive offset by 125 μm steps shows how the pointing direction changes, the optimal offset being $\simeq 125$ μm .

Finally, the forth example is that of a dual-fiber collimator. Dual-fiber collimators are essential components for micro-optic devices because one lens acts to collimate light from and focus light onto two separate fibers. A dual

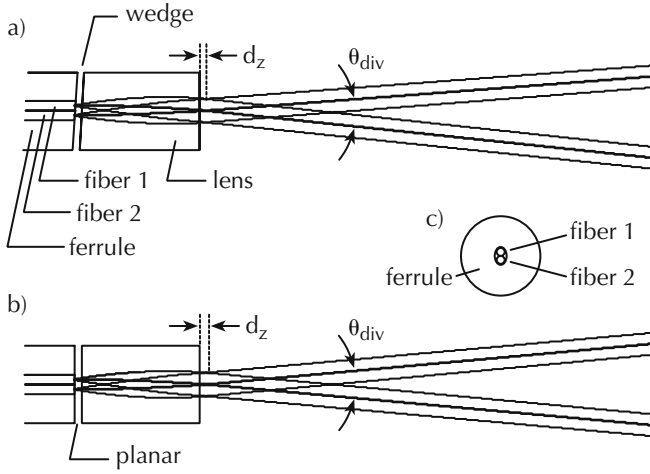


Fig. 5.15. Scale drawings of central and paraxial rays emergent from two SMF-28 fibers positioned in a dual-fiber ferrule (inset) and captured by an NSG SLS2 lens [16]. a) Fiber cores not in plane, b) Fiber cores in plane. The parameters for a) are: $n_{\text{fiber}} = 1.46$, $\theta_{\text{ferrule}} = 8.0^\circ$, N.A. = 0.11, $n_{\text{lens}} = 1.5503$, $L_{\text{lens}} = 6.10$ mm, $P = 0.23$, $\sqrt{A} = 0.237$ mm $^{-1}$, EFL = 2.743 mm, $\theta_{\text{lens}} = 6.0^\circ$, $L_{\text{gap}} = 0.343$ mm, offset = ± 125 μ m, $\theta_{\text{pt}} = -3.03^\circ$, $+2.20^\circ$, $\theta_{\text{div}} = 5.23^\circ$. The parameters for b) are the same except: $\theta_{\text{pt}} = \mp 2.61^\circ$, $\theta_{\text{div}} = 5.22^\circ$.

fiber collimator is the most compact way to fit two fibers into a small package. Two separate collimators use the spatial offset of the lenses to distinguish one port from another. A single dual-fiber collimator uses the divergence angle to distinguish between ports. Many elegant schemes for circulator, interleaver, and thin-film filter architectures have been developed to take advantage of this component.

Figure 5.15 shows a dual fiber collimator using a GRIN lens. A shaped lens could be used as well. The inset illustrates the face of the two fibers inserted into the ferrule. The fibers are stripped of their jacket so that the separation is twice that of the fiber radius. For SMF-28 fiber the core-to-core separation is 250 μ m. For an air-gap collimator there is a choice on how to orient the dual-fiber ferrule with respect to the lens facet. In one case one fiber core extends beyond the other fiber core (Fig. 5.15(a)), and in the other case the fiber cores are flush (Fig. 5.15(b)). It is reported that GRIN collimators typically use the first orientation and shaped-lens collimators use the second orientation [5].

In light of the change in pointing direction with lateral fiber offset, studied in Fig. 5.14, the angle between output collimated beams is calculated using (5.2.38) on page 230. The divergence angle θ_{div} is

$$\theta_{\text{div}} = \frac{y_2 - y_1}{f} \quad (5.3.3)$$

where the position y_k is taken from the centerline of the lens. For small gap lengths, the approximate and generally used expression for the divergence angle is

$$\theta_{\text{div}} = \frac{s}{f} \quad (5.3.4)$$

where s is the separation between fiber cores. A typical divergence angle for commercially available collimators is 3° .

While small pointing errors of a single-fiber collimator are not significant since the collimator housing is simply tilted in compensation, the divergence angle of a dual-fiber collimator is inviolate as the two fibers in the ferrule are fixed in position. For a dual-fiber collimator that receives two output beams, the component architecture must account for the requirement that the beams must enter the collimator at the divergence angle.

5.4 Fiber-to-Fiber Coupling by a Lens Pair

The function of the collimator is to minimize the fiber-to-fiber insertion loss when the beam transits a component core. The elements within the core and the wavefront distortion imparted by the lenses both introduce loss. These losses must be accounted for on a case-by-case basis. Errors in imaging may also lead to losses, and these losses can be estimated by the coupling coefficients derived below.

Figure 5.16 illustrates the two optimal coupling configurations for a gap of length L_{gap} . The first optimal coupling is when the lenses produce a collimated beam. In this case the beam waist between the two lenses is $w = f \times \text{N.A.}$ and the beam suffers minimum diffraction. The second optimal coupling is when the lenses produce an intermediate focus positioned between the lenses. The waist on this focal plane is determined from the magnification of the lenses:

$$M = \frac{w_b}{w_a} = \frac{L_{\text{gap}}}{2f} - 1 \quad (5.4.1)$$

where the latter equation comes from (5.2.34). While not a necessary condition, it is worth noting that the depth of focus equals the gap length when

$$w_b = \sqrt{\frac{\lambda L_{\text{gap}}}{\pi}} \quad (5.4.2)$$

For example, with ideal lenses, a gap length of 100 mm, and $\lambda = 1.545 \mu\text{m}$, a beam diameter of $\sim 450 \mu\text{m}$ puts the depth of focus at the gap length. In turn this corresponds to a magnification factor of $M \sim 43$.

That the aforementioned coupling conditions are optimal is shown using the ABCD matrix formalism. The matrix concatenation for two like collimators is

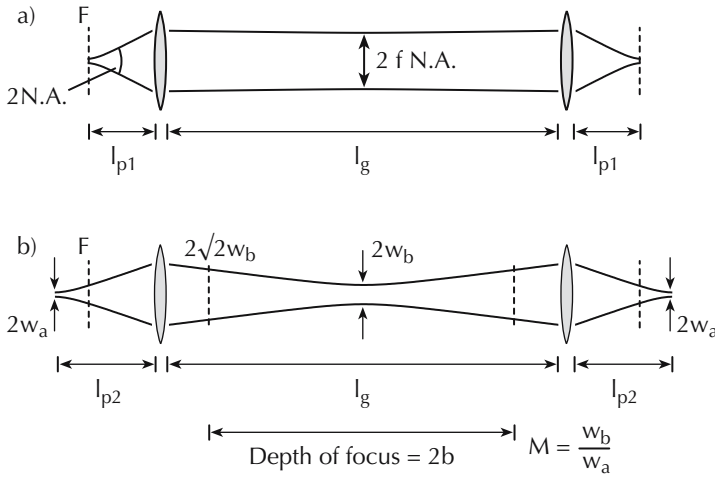


Fig. 5.16. Gaussian beam profile from one fiber to another with two similar lenses. There are two optimal coupling conditions for fixed gap L_{gap} : collimation (a) and focusing (b). a) To collimate the fibers are located on the front focal plane and the beam between lenses nominally does not expand. b) To focus the fibers are located behind the front focal plane and the intermediate beam achieves focus between the lenses. The image on the back focal plane is an image of the fiber face magnified by M .

$$\bar{A} = \begin{pmatrix} 1 & l_p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & L_{\text{gap}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix} \begin{pmatrix} 1 & l_p \\ 0 & 1 \end{pmatrix}$$

The resulting product is

$$\bar{A} = \frac{1}{f^2} \begin{pmatrix} L_{\text{gap}}l_p - 2fl_p - fL_{\text{gap}} + f^2 & (l_p - f)(L_{\text{gap}}l_p - 2fl_p - fL_{\text{gap}}) \\ L_{\text{gap}} - 2f & L_{\text{gap}}l_p - 2fl_p - fL_{\text{gap}} + f^2 \end{pmatrix}$$

To achieve optimal coupling, the B entry must be zero. It is in this circumstance that a point source is imaged to another point source with no regard to the inclination of the rays that emanate from the original source. One finds two possible solutions:

$$(l_p - f) \left(\frac{1}{f} - \frac{1}{L_{\text{gap}}/2} - \frac{1}{l_p} \right) = 0$$

The solutions are

$$\text{Collimate:} \quad l_p = f$$

$$\text{Focus:} \quad \frac{1}{f} = \frac{1}{L_{\text{gap}}/2} + \frac{1}{l_p}$$

To collimate, the fiber facet is positioned on the front focal planes of the lens. The ABCD matrix is

$$\bar{\bar{A}}_{\text{coll}} = \begin{pmatrix} -1 & 0 \\ \frac{L_{\text{gap}} - 2f}{f^2} & -1 \end{pmatrix} \quad (5.4.3)$$

Alternatively, to focus, the fiber facets are pulled behind the front focal planes of the lenses. A focus is reached midway between the lenses with the magnification factor (5.4.1). The ABCD matrix in this case is

$$\bar{\bar{A}}_{\text{focus}} = \begin{pmatrix} 1 & 0 \\ \frac{L_{\text{gap}} - 2f}{f^2} & 1 \end{pmatrix} \quad (5.4.4)$$

Consider first the case of collimation in the ray-optic framework. Denote the excess gap between the two focal lengths of the lens to lens gap d such that $d = L_{\text{gap}} - 2f$. The output point and slope is related to the inputs as

$$\begin{pmatrix} y_1 \\ y'_1 \end{pmatrix} = \begin{pmatrix} -y_o \\ (d/f^2)y_o - y'_o \end{pmatrix}$$

All points y_o in the object plane are reconstructed on the image plane with up-down inversion and unity magnification. For the point on the object plane that is also on axis ($y_o = 0$) the angle of an image ray is the negative of the angle of an associated object ray: $y'_1 = -y'_o$. For points on the object plane off axis, the angle is adjusted to ensure that at the image plane the points reconstruct the object with up-down inversion. In the gaussian framework, the output q parameter is related to the input q parameter as

$$\frac{1}{q_1} = -\frac{f^2}{d} + \frac{1}{q_o} \quad (5.4.5)$$

When the object is in focus, $1/q_o = -j2/kw_o^2$. Clearly the waist of q_1 is the same as q_o , representing unity magnification, and field curvature is added when the excess gap is non-zero.

The same observations regarding the collimating lens pair apply to the focusing lens pair with the exception that the image is not up-down inverted but is instead recovered upright. The notion of excess gap is not applicable to the focusing system because the focal position is designed to lie between the two lenses. Field curvature is imparted at the object plane via the same mechanism that applies to the collimating lens pair. Only when the object and image are at infinity do they both have zero field curvature.

5.4.1 Coupling Coefficients

Calculation of the coupling from fiber to fiber is essential to the design of robust components. Gaussian optics provides a framework from which to calculate the coupling coefficients. However, it is once again emphasized that robust tolerancing must be done with commercially available ray-tracing software, but reasonable loss estimates can be made with gaussian optics.

One method to calculate the coupling loss is that of overlap integrals. A first gaussian mode profile e_1 from one beam is multiplied by a second profile of another beam e_2 and integrated over the cross-section. The general mode overlap integral is

$$\mathcal{I} = \frac{\iint_{-\infty}^{\infty} e_1^* e_2 \, dx dy}{\sqrt{\iint_{-\infty}^{\infty} e_1^* e_1 \, dx dy \iint_{-\infty}^{\infty} e_2^* e_2 \, dx dy}} \quad (5.4.6)$$

The overlap integral $\mathcal{I}$ is bound by zero and unity. The overlap integrals derived in the following account for magnification errors, offset, tilt, and defocusing. The method of overlap integrals allows the coupling coefficient to be calculated at any convenient plane along the optical path, not just, for instance, at the output. The midpoint along a path is often a convenient location to calculate the mode overlap. Note, however, that care must be taken to account properly for propagation direction when the overlap integral is taken away from the source or target; a mode reverse propagated to the overlap plane must be reversed on the plane to account for the complex conjugate found in the integral.

Magnification and Offset Errors

Consider a first gaussian mode with circular beam radius w_1 and centered along the axis of propagation. The normalized mode profile is

$$e_1(x, y) = \frac{1}{\sqrt{\pi w_1^2}} \exp\left(-\frac{x^2 + y^2}{2w_1^2}\right) \quad (5.4.7)$$

This mode might be the approximate eigenmode of a fiber. Now, consider a second gaussian mode that is imaged onto the first mode. The second mode has magnification error, where the beam radius w_2 is not w_1 , and offset error where the offset is taken, without loss of generality, along the x axis. The second mode profile is written as

$$e_2(x, y) = \frac{1}{\sqrt{\pi w_2^2}} \exp\left(-\frac{(x - \Delta x)^2 + y^2}{2w_2^2}\right)$$

The coupling coefficient of these two modes is calculated from (5.4.6):

$$\mathcal{I}_a = 2 \frac{w_1 w_2}{w_1^2 + w_2^2} \exp\left(-\frac{\Delta x^2}{2(w_1^2 + w_2^2)}\right) \quad (5.4.8)$$

When both mode radii are the same, the overlap integral is unity for zero offset and decreases to zero as the offset is increased. In the asymptotic limit that one mode is much larger than the other, the overlap integral is dominated by the smaller mode size.

Tilt Error

Another error is the tilt error. This can come from misalignment of the collimators or may be built into the architecture of the device, e.g. a wedge-type polarization independent isolator. The tilt is accounted for by a phase front rotation by the angle between the two beams. For a tilt angle of γ , the phase rotation is $\exp(+jkx \tan \gamma)$, where k is the wavenumber. For normalized modes and in the small angle limit, the phase tilt is added to (5.4.6) via

$$\mathcal{I} = \iint_{-\infty}^{\infty} e_1^* e^{-j\gamma kx} e_2 dx dy \quad (5.4.9)$$

Two beams may be tilted and offset along different directions. The perpendicular and parallel cases are considered here. When the tilt and offset are perpendicular, the resultant coupling coefficient is

$$\mathcal{I}_{b\perp} = 2 \frac{w_1 w_2}{w_1^2 + w_2^2} \exp \left(-\frac{\gamma^2 k^2 w_1^2 w_2^2 + \Delta x^2}{2(w_1^2 + w_2^2)} \right) \quad (5.4.10)$$

When the beams are aligned but for the tilt and the magnification is unity, the tilt penalty is

$$\mathcal{I}'_{b\perp} = \exp \left(-\frac{\gamma^2 k^2 w^2}{4} \right) \quad (5.4.11)$$

When the tilt and offset are parallel, the resultant coupling coefficient is

$$\mathcal{I}_{b\parallel} = 2 \frac{w_1 w_2}{w_1^2 + w_2^2} \exp \left(-\frac{\gamma^2 k^2 w_1^2 w_2^2 + 2j\gamma k \Delta x w_1^2 + \Delta x^2}{2(w_1^2 + w_2^2)} \right) \quad (5.4.12)$$

This coupling coefficient is a complex number. The phase of the coefficient does not matter when one beam overlaps with another. However, accounting for the phase is critical when two or more co-polarized beams are coupled into the same fiber or waveguide; in this case coherent interference will occur and more or less coupling can be achieved with offset of tilt of the beam. In the present case, only one beam couples to another, in which case the magnitude of (5.4.12) is what matters. That magnitude is the same as written in (5.4.10), that is

$$|\mathcal{I}_{b\parallel}| = \mathcal{I}_{b\perp} \quad (5.4.13)$$

Focus Error

Focus error can be treated in the same matter as tilt error where a phase term is added to the overlap integral. For tilt the phase front was modified by a linear increase in phase as a function of lateral coordinate. For focus error, within the gaussian approximation where all field curvatures are spherical, a quadratic phase increase (or decrease) as a function of lateral coordinate is inserted. The overlap integral takes the form

$$\mathcal{I} = \iint_{-\infty}^{\infty} e_1^* e^{\left(-j \frac{x^2+y^2}{2r^2}\right)} e_2 \, dx dy \quad (5.4.14)$$

where r is the field curvature. Eliminating tilt and offset errors but retaining magnification error, the coupling coefficient with focus error is

$$\mathcal{I}_c = \frac{1}{w_1 w_2} \left(\frac{1}{2w_1^2} + \frac{1}{2w_2^2} + \frac{j}{2r^2} \right)^{-1} \quad (5.4.15)$$

As with the tilt error, the integral (5.4.15) is a complex number. The imaginary part is associated with the field curvature error. When more than one co-polarized beam is coupled to the same output port, interference due to the field curvature results. However, in the present case no such interference is considered and the magnitude of the coupling coefficient generates the loss penalty. The magnitude of (5.4.15) is

$$|\mathcal{I}_c| = \frac{1}{w_1 w_2} \left(\left(\frac{1}{2w_1^2} + \frac{1}{2w_2^2} \right)^2 + \left(\frac{1}{2r^2} \right)^2 \right)^{-1/2} \quad (5.4.16)$$

and finally when the magnification is unity the defocus penalty reduces to

$$|\mathcal{I}'_c| = \frac{1}{\sqrt{\left(\frac{w^2}{2r^2} \right)^2 + 1}} \quad (5.4.17)$$

Taylor expansion of (5.4.17) shows that the initial penalty accrues as quickly as that for offset (5.4.8) and tilt (5.4.11) errors.

References

1. K. Asano and H. Hosoya, "Collimator lens, fiber collimator and optical parts," U.S. Patent Application 2002/0 168 140 A1, Nov. 14, 2002.
2. P. Bernard, M. A. Fitch, P. Fournier, M. F. Harris, and W. P. Walters, "Fabrication of collimators employing optical fibers fusion-spliced to optical elements of substantially larger cross-sectional areas," U.S. Patent 6,360,039 B1, Mar. 19, 2002, same spec as US 2002/041742 A1 and US 2002/0054735 A1.
3. —, "Fabrication of collimators employing optical fibers fusion-spliced to optical elements of substantially larger cross-sectional areas," U.S. Patent Application 2002/0 041 742 A1, Apr. 11, 2002, same spec as US 6,360,039 B1 and US 2002/0054735 A1.
4. —, "Fabrication of collimators employing optical fibers fusion-spliced to optical elements of substantially larger cross-sectional areas," U.S. Patent Application 2002/0 054 735 A1, May 9, 2002, same spec as US 6,360,039 B1 and US 2002/0041742 A1.
5. C. Brophy and A. K. Thompson, "Dual fiber collimator," U.S. Patent Application 2003/0 021 531, Jan. 30, 2003.
6. Casix Quality Assurance Department, "Reliability test report on c-collimator," Casix, Inc., Fuzhou, Fujian, P.R. China, Tech. Rep. TR1201020 Issue 01, Oct. 1999. [Online]. Available: <http://www.casix.com>
7. —, "Reliability test report on collimator," Casix, Inc., Fuzhou, Fujian, P.R. China, Tech. Rep. TR1201001 Issue 01, Dec. 1999. [Online]. Available: <http://www.casix.com>
8. "Mode-field diameter measurement method," Corning Incorporated, Corning, NY, Aug. 2001, MM16.
9. "Corning SMF-28 optical fiber product information," Corning Incorporated, Corning, NY, Aug. 2002, PI1036.
10. H. A. Haus, *Waves and Fields in Optoelectronics*. Englewood Cliffs, New Jersey: Prentice-Hall, 1984.
11. "Micro optics for telecom catalog 2002," Kocent Communications, Fuzhou, Fujian, P.R. China, 2002. [Online]. Available: <http://www.koncent.com/>
12. J. A. Kong, *Electromagnetic Wave Theory*. New York: John Wiley & Sons, 1989.
13. "Lightpath technologies product catalog," LightPath Technologies, Orlando, FL, 2003. [Online]. Available: <http://www.lightpath.com/literature.html>
14. Z. Liu, "Optical collimator with long working distance," U.S. Patent 6,469,835 B1, Oct. 22, 2002.
15. NSG America, Inc., Somerset, NJ, object at 'Dispersion Equations and Paraxial Optics Formulae'. [Online]. Available: <http://www.nsgamerica.com/technical.shtml>
16. "Selffoc microlens table," NSG America, Inc., Somerset, NJ. [Online]. Available: <http://www.nsgamerica.com/technology/microlens.cfm>
17. I. Ooyama, T. Fukuzawa, and S. Kai, "Optical fiber collimator," U.S. Patent Application 2002/0 094 163 A1, July 18, 2002.
18. J.-J. Pan, M. Shih, and J. Xu, "Integrable fiberoptic coupler and resulting devices and system," U.S. Patent 5,889,904, Mar. 30, 1999.
19. C. Qian and Y. Qin, "Optical fiber collimator with long working distance and low insertion loss," U.S. Patent Application 2002/0 197 020 A1, Dec. 26, 2002.

20. M. Shirasaki, "Optical apparatus which uses a virtually imaged phased array to produce chromatic dispersion," U.S. Patent 5,930,045, July 27, 1999.
21. *Generic Reliability Assurance Requirements for Passive Optical Components*, Telcordia Technologies Std. GR-1221-CORE, 1999. [Online]. Available: <http://telecom-info.telcordia.com/site-cgi/ido/index.html>
22. L. Ukrainczyk, "Optical fiber collimators and their manufacture," U.S. Patent Application 2003/0026535 A1, Feb. 6, 2003.
23. A. Yariv and P. Yeh, *Optical Waves in Crystals*. Hoboken, New Jersey: Wiley-Interscience, John Wiley & Sons, Inc., 2003.
24. Y. Zheng, "Dual fiber optical collimator," U.S. Patent 6,148,126, Nov. 14, 2000.
25. —, "Reliable low-cost dual fiber optical collimator," U.S. Patent 6,246,813 B1, June 12, 2001.

Isolators

An isolator is a one-input, one-output component that transmits light in a forward direction and blocks light in a reverse direction. All isolators incorporate at least one Faraday rotator (FR) as the nonreciprocal element. There are two classes of isolators: polarizing isolators that isolate by rejecting the unwanted polarization state, and polarization-independent isolators that isolate by spatial filtering. Polarizing isolators have a single optical path that transits a first polarizer, the FR, and a second polarizer. In the isolation direction the unwanted polarization is either absorbed or deflected. Polarization-independent isolators use polarization diversity to form two co-directional optical paths, one for each polarization. Along the forward direction the two paths recombine at the collimator while along the reverse direction the two paths fall outside of the fiber aperture.

Analysis of the polarizing isolator highlights the issues regarding transmission and isolation as a function of temperature, wavelength, and manufacturing error. The materials-related shortcomings of iron garnets are apparent in this analysis. The polarization-diversity schemes to realize polarization-independent isolators guide the best choice lensing system and open up the issue of path-length balancing. A path-length imbalance imparts polarization-mode dispersion (PMD). PMD-compensated isolators are realized by select changes in the optical elements used to build the isolator or by explicit addition of a compensating element.

6.1 Polarizing Isolator

Polarizing isolators were the first isolator type to demonstrate the importance of the nonreciprocal Faraday rotator. Early literature shows a Verdet-based isolator [28] and a terbium-aluminum garnet-based isolator [16]. A garnet-based polarizing isolator is illustrated in Fig. 6.1. The polarizing axes of the first and second polarizers are rotated 45° with respect to one another. A Faraday rotator with a fixed magnetization direction is placed between the

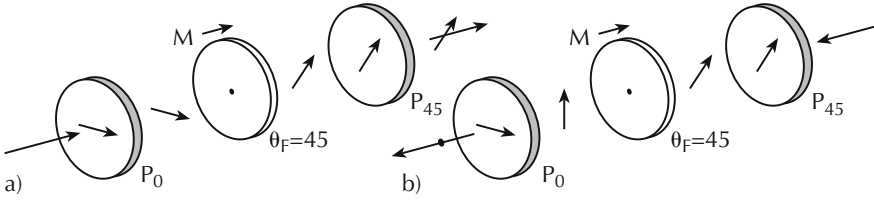


Fig. 6.1. Faraday rotation of 45° between two polarizers. Polarizer and analyzer are rotated 45° with respect to one another to maximize transmission and isolation. a) Forward path allows transmission. Horizontal linear polarization is rotated $+45^\circ$ by FR and transits analyzer $\mathcal{P}_{45}$ without loss. b) Reverse, isolation path. $+45^\circ$ linear polarization is rotated $+45^\circ$ by FR and is extinguished by analyzer $\mathcal{P}_0$.

two polarizers. In practice the FR is a saturated Bi:RIG iron garnet (cf. §4.2.3) where the magnetization is fixed by a permanent magnet, such as Sm-Co, or where the iron garnet is latching and pre-poled. Multi-magnet schemes have been proposed to concentrate the magnetic field around the FR [5, 6, 18], but in practice a single magnet is used. The FR is designed to rotate a linear polarization state by $+45^\circ$ (or -45°) irrespective of transit direction.

In the forward, or transmission, direction, the lead polarizer $\mathcal{P}_0$ polarizes the light along the horizontal (the absolute direction being, of course, immaterial). The FR subsequently rotates the polarization by $+45^\circ$, which aligns it for complete transmission through the second polarizer $\mathcal{P}_{45}$. In the reverse, or isolation, direction, the lead polarizer $\mathcal{P}_{45}$ polarizes the light at $+45^\circ$. The FR rotates the polarization by $+45^\circ$, at which point the polarization is clipped by the second polarizer $\mathcal{P}_0$. The second polarizer absorbs the light.

Problems with this system arise from the wavelength and temperature dependence of the FR plate, manufacturing error of the plate length, residual linear birefringence in the garnet, and multiple reflections due to imperfect antireflection coatings. Residual linear birefringence is reduced by removing the garnet from its substrate and annealing the film, although linear birefringence remains the ultimate limiting factor. Imperfect antireflection coatings cause multiple reflections inside the material; each full pass rotates the polarization by approximately 90° , so every other round-trip reflection creates a polarization component that reduces isolation.

The specific rotation of an iron garnet has temperature and wavelength dependencies:

$$\theta'_F(\lambda, T) = \frac{2\pi}{\lambda} \Delta n(\lambda, T) \quad (6.1.1)$$

The Faraday rotation angle θ_F for a plate of length L is

$$\theta_F = \theta'_F L \quad (6.1.2)$$

At a nominal temperature, wavelength, and thickness the target rotation is θ_{Fo} . The actual rotation for small deviations is $\theta_F = \theta_{Fo} + \Delta\theta_F$, where the total deviation from the target is

$$\Delta\theta_F = \frac{d\theta_F}{dL} \Delta L + \frac{d\theta_F}{dT} \Delta T + \frac{d\theta_F}{d\lambda} \Delta\lambda \quad (6.1.3)$$

The first term comes from manufacturing error, the second from temperature dependence of the garnet, and the third from the total wavelength dependence. The wavelength dependence has two components, one from the fixed waveplate thickness and the other from wavelength-dependent Δn of the garnet:

$$\frac{d\theta_F}{d\lambda} \Delta\lambda = \frac{\Delta\omega}{\omega_o} \theta_{Fo} + \frac{1}{\Delta n} \left(\frac{d\Delta n}{d\lambda} \right) \theta_{Fo} \Delta\lambda \quad (6.1.4)$$

The first term is simply the frequency dependence of the waveplate and comes from (4.6.13) on page 182. The second term highlights the material dependence on wavelength. This term is not zero and generally increases the overall wavelength dependence of the plate.

Recall that the eigen-axis of a Faraday rotator is $\pm\hat{s}_3$. The associated Jones operator U_F is

$$U_F = \cos\theta_F I \mp j\sigma_3 \sin\theta_F \quad (6.1.5)$$

The operator U_F must be treated carefully: due to the nonreciprocal nature of the FR, the signs of θ_F and σ_3 are invariant to transit direction. The $\mp$ sign encompasses only the magnetization direction $\mathbf{M}$ and the particular Bi:RIG material. Once these are fixed, the sign is fixed. The point of including the $\mp$ sign is to represent the possibility that the FR or permanent magnet can be flipped around. Without loss of generality, the $(-)$ sign will be used in the following.

Also, recall that a polarizer is represented by the projection matrix (2.5.2) on page 52. The two polarizers for this nominal isolator are

$$\mathcal{P}_o = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ and } \mathcal{P}_{45} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (6.1.6)$$

Equipped with these polarizer and FR matrices, the forward and reverse isolator paths can be analyzed.

In the forward direction, output state $|t\rangle$ is generated from input state $|s\rangle$ via

$$|t\rangle = \mathcal{P}_{45} U_F \mathcal{P}_o |s\rangle$$

Recalling that $\mathcal{P}^2 = \mathcal{P}$, the output intensity is

$$\langle t|t\rangle = \langle s|\mathcal{P}_o^\dagger U_F^\dagger \mathcal{P}_{45} U_F \mathcal{P}_o |s\rangle \quad (6.1.7)$$

Combining (6.1.5-6.1.6) into (6.1.7) gives the forward transmission matrix

$$T_{\text{iso}}^+ = \frac{1}{2} (1 + \sin 2\theta_F) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

At the nominal Faraday rotation angle $\theta_{Fo} = +45^\circ$ the transmission is unity.

In the reverse direction, the output state $|s'\rangle$ is generated from input state $|t'\rangle$ via

$$|s'\rangle = \mathcal{P}_o U_F \mathcal{P}_{45} |t'\rangle$$

The output intensity is

$$\langle s' | s' \rangle = \langle t' | \mathcal{P}_{45}^\dagger U_F^\dagger \mathcal{P}_o U_F \mathcal{P}_{45} | t' \rangle$$

which yields the reverse transmission matrix

$$T_{\text{iso}}^- = \frac{1}{4} (1 - \sin 2\theta_F) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Stripping away the polarization orientation, the norm of T_{iso}^- gives the total reverse intensity

$$|T_{\text{iso}}^-| = \frac{1}{2} (1 - \sin 2\theta_F) \quad (6.1.8)$$

At the nominal Faraday rotation angle $\theta_{Fo} = +45^\circ$ the transmission is zero.

The nominal forward and reverse transmissions are unity and zero, respectively. However, rotation deviations included in (6.1.3) degrade the performance. Accounting for deviations, the forward and reverse transmissions are

$$T_{\text{iso}}^+ = \cos^2 \Delta\theta_F, \quad \text{and} \quad T_{\text{iso}}^- = \sin^2 \Delta\theta_F$$

Isolation is defined as

$$I_{\text{iso}} = -10 \times \log_{10} (T_{\text{iso}}^-) \quad (\text{dB}) \quad (6.1.9)$$

The forward transmission, or insertion loss (IL), changes to second order with change in $\Delta\theta_F$ and the isolation changes to first order.

Consider the frequency dependence of the rotation alone:

$$\Delta\theta_F = \frac{\omega - \omega_o}{\omega_o} \theta_{Fo} \quad (6.1.10)$$

where θ_{Fo} coincides with frequency ω_o . Figure 6.2(a) plots the isolation as a function of frequency over the C-band. This idealized FR imparts $+45^\circ$ rotation at 194.1 THz. Recalling the spectral coverage for the C-band from (4.1.5) on page 146, the frequency deviation from center is $\Delta f/f_o = \pm 1.03\%$, which translates to $\Delta\theta_F \simeq \pm 0.45^\circ$. In turn, the minimum isolation on either side of the C-band is $I_{\text{iso}} \simeq 42$ dB. This is a high level of isolation that under ideal conditions would be suitable for many applications.

The remaining three factors further degrade the practical performance of an isolator. One factor is that the FR plate is generally toleranced to $\theta_F = 45 \pm 1^\circ$ [17]. This translates into a frequency shift of the maximum isolation point. Figure 6.2(b) shows such a frequency shift for $\Delta\theta_F = -0.25^\circ$. The isolation for a 1° rotation error, combined with the 0.45° frequency-dependent

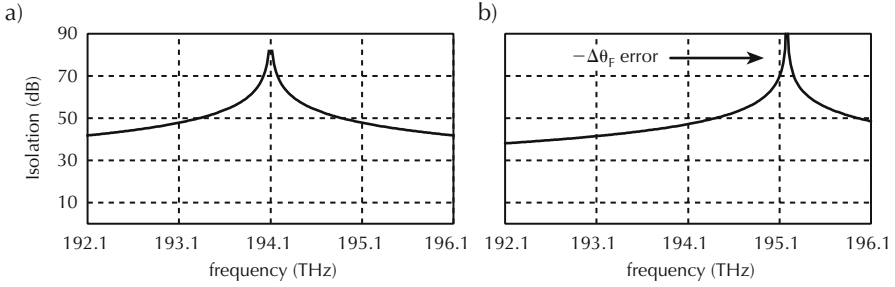


Fig. 6.2. The frequency dependence of the Faraday rotation plate, ignoring material wavelength dependence, results in frequency-dependent isolation. a) Isolation over the C-band. b) Any change in the FR angle changes the frequency of peak isolation.

rotation, gives $\Delta\theta_F = 1.45^\circ$, or $I_{\text{iso}} \simeq 32$ dB. This level of isolation is commonly found in component specification sheets of polarization-independent isolators at room temperature [9].

The remaining two factors are the change in specific rotation as a function of temperature and wavelength. Examples of practical temperature and wavelength ranges are 0°C to $+70^\circ\text{C}$ and the short to long wavelength sides of the C-band. Taking room temperature as $\text{RT} = 25^\circ\text{C}$, the maximum temperature excursion is $\Delta T = 45^\circ\text{C}$. The C-band is covered by $\Delta\lambda = \pm 16$ nm.

The total of all four deviation factors should not cause worst-case isolation to fall below a specified value. For example, $I_{\text{iso}} = 20$ dB requires $\Delta\theta_F|_{\text{max}} = \pm 5.75^\circ$. Subtracting 1.0° for manufacturing tolerance, there remains 4.75° for temperature and total wavelength dependencies. Setting the coefficients equal gives

$$\frac{d\theta_F}{d\lambda} \simeq 0.08^\circ/\text{nm}, \text{ and } \frac{d\theta_F}{dT} \simeq 0.08^\circ/^\circ\text{C} \quad (6.1.11)$$

These coefficients translate to a 1.3° allowance for total wavelength dependence and a 3.6° allowance for temperature dependence. Moreover, the wavelength-dependent component of the waveplate is 0.45° , so the material component must be less than 0.85° . As discussed in §4.2.3, these coefficients are challenging from a materials standpoint.

That the worst-case isolation can fall to 20 dB even with perfect polarizing elements is quite a remarkable fact. The relatively poor performance of the FR over the full operating range necessitates two-stage isolators for high-performance applications.

One simple improvement ubiquitous in the industry is to change the angle between polarizers to account for manufacturing error of the FR plate. A 1° manufacturing error in rotation is 18% of the total error budget. In the isolation direction, a manufacturing error of $\Delta\theta_F$ is made up with a one-to-one change in the rotation of the analyzing polarizer. This, however, changes the insertion loss of the forward direction. If the analyzer is rotated by angle 45°

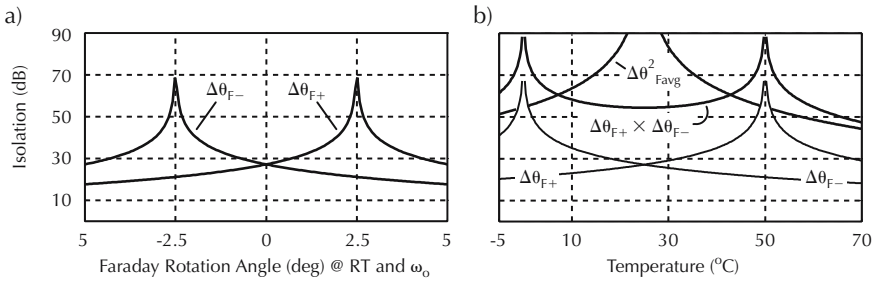


Fig. 6.3. Two-stage, complementary FR configuration [19]. a) One stage has FR biased to an angle below 45° while the other stage is biased above 45° . b) Two-stage isolation across temperature. The product of positive- and negative-biased FR angles provides better overall isolation than two stages with the same FR centered at 25°C .

$+\rho$, the insertion loss varies with analyzer error ρ as

$$|T_{\text{iso}}^+(\rho)| = \cos^2 2\rho \quad (6.1.12)$$

Note that the IL goes as 2ρ , but still imparts only a second-order change.

To improve the overall isolation, a two-stage isolator is necessary. Shiraishi proposes a two-stage isolator where the two FRs are detuned in frequency about a center frequency [19, 20]. The detuning is easily achieved by polishing the iron garnets to different thicknesses. For detuning rotations $\Delta\theta_{F1}$ and $\Delta\theta_{F2}$, the forward and reverse transmissions in cascade are

$$T_{\text{iso}}^+ = \cos^2 \Delta\theta_{F1} \cos^2 \Delta\theta_{F2} \quad (6.1.13a)$$

$$T_{\text{iso}}^- = \sin^2 \Delta\theta_{F1} \sin^2 \Delta\theta_{F2} \quad (6.1.13b)$$

As the peak isolation frequencies shift with wavelength and temperature, the overall isolation remains higher than a two-stage isolator using the same FR. Figure 6.3 illustrates the principle. Figure 6.3(a) plots the frequency-dependent isolation from two individual isolators where the FRs are different. The combined isolation over temperature is shown in Fig. 6.3(b). The combined isolation does not fall below 48 dB while that for two stages of a single FR type centered at 25°C is slightly worst at both temperature extremes.

6.2 Comparison of Lens Systems

The next logical step in isolator development is the polarization-independent (PI) isolator. There are two different types of PI isolators: deflection-type and displacement-type. These types are detailed in the following sections. Here, a look at the different optimal lensing systems serves to emphasize the contrast between the two isolator types.

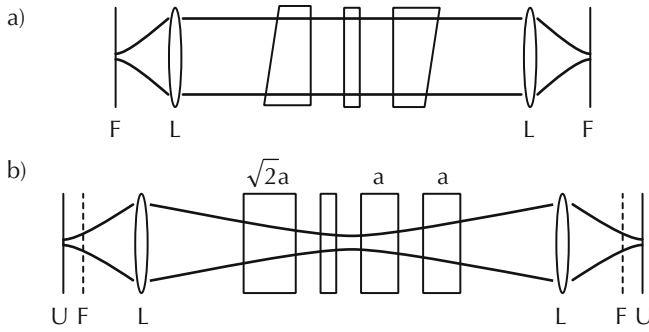


Fig. 6.4. Lensing systems for deflection and displacement isolators. a) Deflection isolator uses a collimating design for shortest length. A collimating lens system transforms angle in image space to position on the focal plane. b) Displacement isolator uses a focusing design to minimize the necessary walkoff, in turn minimizing the length of the birefringent crystals.

Figure 6.4 illustrates a deflection-type isolator (top) and displacement-type isolator (bottom) and their respective collimating and focusing lens systems (cf. §5.4). The core of the deflection isolator changes the angle of the beam as a function of propagation direction. The collimating system transforms angle into lateral offset at the focal plane, which leads to spatial filtering in the isolation direction. The core of the displacement isolator operates on the principle of walkoff, where beams are laterally displaced but remain collinear. The lensing system transforms offset in the image plane to offset in the object plane, the magnification of the system being the scale factor between the two offsets. These offsets also lead to spatial filtering in the isolation direction.

Figure 6.5 illustrates the different configurations relevant to isolator designs. Figure 6.5(a) shows an idealized alignment between the axis of a collimated beam and a lens. All rays converge at the focal plane F . Figure 6.5(b) shows the same collimating lens but where the beam is offset from the lens axis by y_o , the angle remaining aligned to that axis. Since angle transforms to offset for a collimating system, all rays converge at the original focal point of Fig. 6.5(a), albeit along the inclination θ_c . Figure 6.5(c) shows a collimated beam deflected by angle θ_d ; the beam is focused to point y_i , which is offset from the on-axis focal point. The focusing lens shown in Fig. 6.5(d) shows a diverging beam offset from and parallel to the lens axis. On the object plane U the rays focus to a point offset by y_i . If the magnification of the system is M , the spot size w_o and offset y_o on the object plane are scaled to $w_i = w_o/M$ and $y_i = y_o/M$ on the image plane.

With these two configurations in mind, the action of the deflection-type and displacement-type isolator cores and their relationship to the lensing systems will be more apparent.

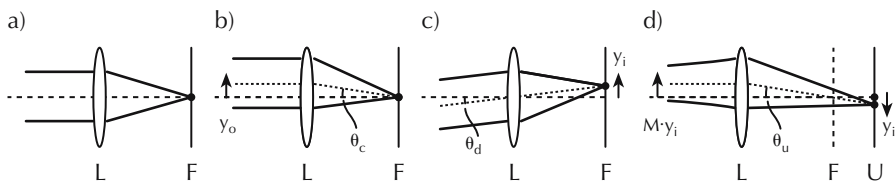


Fig. 6.5. Isolator-related beam transformations through a lens. Collimated systems are in (a–c) and a focusing system is in (d). a) Collimated beam travels through center of lens. b) Offset collimated beam focuses on same spot as (a). c) Deflected beam translates to focal-point offset. d) Offset diverging beam is imaged to a point offset from center.

6.3 Deflection-Type Isolators

The first polarization-independent isolators are reported in [12, 13] and [26]. These isolators used iron garnet Faraday rotators and polarization diversity schemes. However, the number of parts and required size of the isolators, much less the deleterious inclusion of a bandwidth-limiting half-wave waveplate in the design by [12], makes these PI isolators impractical.

The first practical polarization-independent isolator was invented by Shirasaki [21, 22]. The Shirasaki isolator is a deflection-type isolator that uses birefringent wedges to create polarization diversity and to polarize and analyze the light. The core of the deflection isolator is illustrated in Fig. 6.6. Two birefringent wedge prisms (cf. §4.7) are oriented as shown and a Faraday rotator is placed in between. The wedge angle of the prism deflects the incoming light, and the birefringence of the prism creates a polarization-dependent deflection. The extraordinary axis of the birefringent crystal is cut to lie in the face of the prism; in this way there is no walkoff but there is polarization-dependent refraction. The only required relation between the extraordinary axes of the two wedges is that they are $\pm 45^\circ$ apart in the plane perpendicular to propagation. As illustrated in Fig. 6.6(a), when the wedge is cut such that the e -axis is at $+22.5^\circ$ from an edge of the rectangular aperture, the same wedge can be used at both ends. The Faraday rotator is a Bi:RIG iron garnet that has its saturated magnetization maintained by a permanent magnet, Fig. 6.6(b). Originally, a YIG garnet was used.

The forward and reverse ray-trace diagrams are shown in Fig. 6.7. (A detailed ray-trace that includes chief and marginal rays is available in [1]). In the forward direction (Fig. 6.7(a)), the input light is refracted by the first wedge into u - and v -paths, the paths being orthogonally polarized. On plane (a) the point and polarization of the u - and v -paths are indicated, where the polarization of the v -path is extraordinary in relation to the first wedge while the u -path is ordinary. Transit of the two beams through the FR rotates their linear polarization states by 45° in a clockwise manner. The polarizations of these beams are now aligned to the extraordinary axis of the second wedge. The v -path refracts based on the extraordinary index of the second wedge

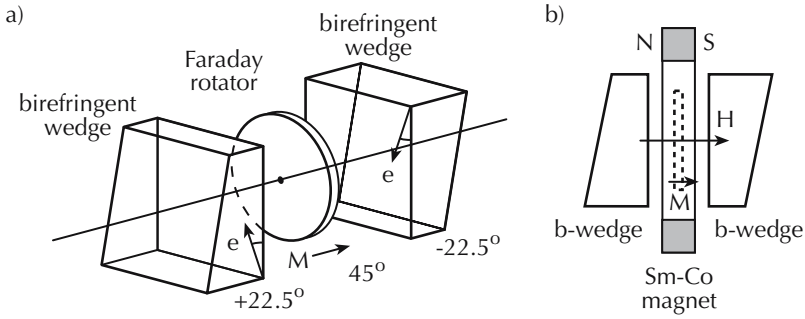


Fig. 6.6. Core of single-stage deflection-type isolator. a) Two birefringent wedges having the same wedge angle are configured as shown. A Faraday rotator plate is located between the wedges and the magnetization vector points along the optic axis. b) The magnetization of the FR is typically held in saturation by a permanent magnet such as Sm-Co.

while the u -path refracts based on the ordinary index. Provided wedge angles and materials are the same, the second wedge deflection cancels the first. The u - and v -paths run collinear before the second lens, the lens brings the two beams to focus at the same point.

In the reverse direction the two wedges conspire to deflect the beams out of the aperture of the return fiber. Accounting for the polarization rotation from the Faraday rotator, the two wedges form a Wollaston prism as viewed from the isolation direction. As shown in Fig. 6.7(b), the wedge w_2 imparts double refraction based on the polarization state of the input light (the same as wedge w_1 for the forward path). The linear polarization states of paths u' and v' are shown at plane (a'). Transit through the FR again rotates the linear polarization states by 45° in a clockwise manner. At the leading face of wedge w_1 the polarizations on the two paths are not aligned to the wedge. The v' -path which was refracted by the extraordinary index in wedge w_2 now refracts by the ordinary index, retaining a residual deflection that is calculated below. The opposite alignment occurs for the u' -path with the effect of a residual deflection in the opposite direction. Together, the two wedges split and deflect the incoming light.

There are four calculations required for the deflection-type isolator: in the forward direction, the loss due to the beam offset before lens L_2 ; in the reverse direction, the loss due to the beam deflection by the wedge pair; the angle of the wedge; and the path-length imbalance, or PMD, in the forward direction.

In the forward direction, offset y_o on the left side of lens L_2 is mapped by the lens into tilt angle θ_c . For simplicity consider that the overall magnification of the lens pair is unity. The beam waists w_o of the fiber and focused-beam modes are then the same. Using (5.4.11) on page 243, the mode-overlap due to tilt θ_c is

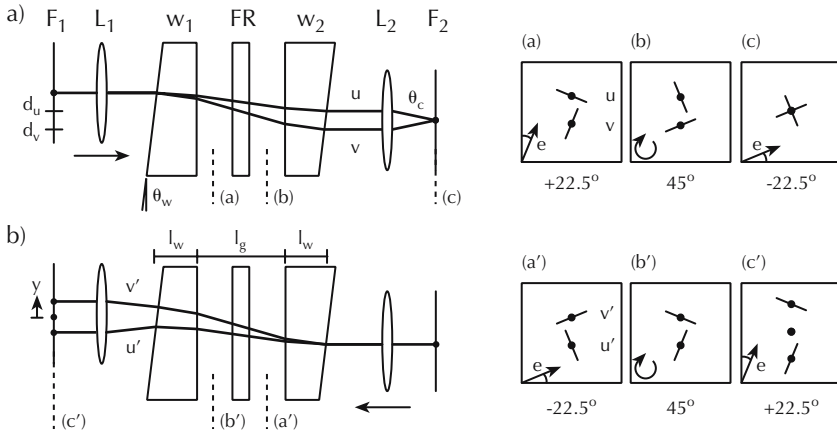


Fig. 6.7. Ray-trace diagrams for forward and isolation directions in a deflection-type isolator. a) Forward-path ray trace: beams of cross polarizations converge at the output fiber. Right: spot diagram through core. b) Isolation-path ray trace: beams of cross polarizations are deflected by the wedges, falling outside of the aperture of the return fiber. The frames around the spot diagrams are a guide for the eye only.

$$\mathcal{I}_b = \exp \left(-\frac{\theta_c^2 k^2 w_o^2}{4} \right) \quad (6.3.1)$$

where k is the wavenumber and the tilt angle θ_c is related to offset y_o and lens focal length f via

$$y_o = \theta_c f \quad (6.3.2)$$

The beam offset is therefore related to the lens parameters and mode overlap as

$$y_o = \frac{2f}{kw_o} \sqrt{-\ln \mathcal{I}_b} \quad (6.3.3)$$

The offset in turn is determined by the difference in displacement between the u and v paths. The displacement difference is

$$d_v - d_u = (n_e - n_o) \theta_w \left(\frac{2l_w}{n_e n_o} + l_g \right) \quad (6.3.4)$$

where θ_w is the angle of the wedge, l_w and l_g are the wedge and gap lengths, and n_e and n_o are the extraordinary and ordinary refractive indices. (Note that if the wedge prisms were exchanged such that the flat facets face the lenses, only the gap length l_g would contribute to the displacement.) Provided that the axis of lens L_2 bisects the displacements d_u and d_v , the offset is related to the displacement difference as $y_o = (d_v - d_u)/2$.

To appreciate the order of magnitude for tolerable offset y_o , consider $f = 1$ mm, $\omega = 2\pi \times 194.1$ THz, and $w_o = 5$ μ m. For an insertion loss of $\mathcal{I}_b = -0.05$ dB attributable to the beam displacement (and not imperfect AR coatings or material losses), the required offset is $y_o \simeq 7$ μ m.

In the reverse direction, deflection θ_d imparted by the wedge pair is mapped by lens L_1 to an offset position y_i on focal plane F_1 . Displacement of the reverse beams from lens center also creates a tilt angle when mapped through the lens, but given the small displacement calculated for the forward direction this angle is ignored here. Provided unity magnification, the mode-overlap due to displacement y_i between the focused beams and the fiber facet, given by (5.4.8) on page 242, is

$$\mathcal{I}_a = \exp\left(-\frac{y_i^2}{4w_o^2}\right) \quad (6.3.5)$$

where the displacement y_i is related to the deflection angle via

$$y_i = \theta_d f \quad (6.3.6)$$

The required deflection angle for a given isolation $\mathcal{I}_a$ is therefore

$$\theta_d = \frac{2w_o}{f} \sqrt{-\ln \mathcal{I}_a} \quad (6.3.7)$$

Continuing with the same example, an isolation of $I_{\text{iso}} = -45$ dB requires a deflection angle of $\theta_d \simeq 1.8^\circ$.

The deflection angle θ_d is directly related to the wedge angle θ_w and the birefringence of the crystal. Recall the deflection angle β for a small-angle prism, given by (4.7.2) on page 199. Consider first the forward u -path. The deflections due to wedges w_1 and w_2 are

$$\beta_{u1} = -(n_o - 1)\theta_w \quad (6.3.8a)$$

$$\beta_{u2} = (n_o - 1)\theta_w \quad (6.3.8b)$$

Both deflections are based on the ordinary refractive index seen by the u path. One sign is the opposite of the other because the based of the wedge prisms are inverted. The total deflection is $\beta_{u1} + \beta_{u2} = 0$. This is consistent with the previous analysis of the forward path.

In the reverse direction the wedge deflections do not cancel. Following the same analysis, the deflections of the u' -path are

$$\beta_{u'2} = (n_o - 1)\theta_w \quad (6.3.9a)$$

$$\beta_{u'1} = -(n_e - 1)\theta_w \quad (6.3.9b)$$

The total deflection of the u' -path is therefore

$$\beta_{u'} = \theta_d = -(n_e - n_o)\theta_w \quad (6.3.10)$$

The deflection along the v' -path is the negative of (6.3.10).

The deflection angle θ_d can now be associated with the isolation requirement. Substitution of (6.3.10) into (6.3.7) gives

$$\theta_w = \frac{2w_o}{\Delta n f} \sqrt{-\ln \mathcal{I}_a} \quad (6.3.11)$$

Using the exemplar values from above and using YVO_4 as the wedge material, the wedge angle to provide $I_{\text{iso}} = -45$ dB is $\theta_w \simeq 9.0^\circ$. This angle is consistent with YVO_4 wedges that are used in the industry.

Together, equations (6.3.3), (6.3.4), and (6.3.11) define the specification of a wedge-type isolator. For these three equations there are three free variables: focal length f , wedge thickness l_w , and gap length l_g . The remaining parameters are the fiber, the material, and the transmission and isolation specifications.

Table 6.1 presents deflection-type isolator specifications for a one-stage isolator reported by a manufacturer. The high return loss is achieved by collimator selection as well as orienting the angled facet of the wedges toward the collimators to minimize back reflection. Power handling is limited by the collimator technology (cf. §5.1), and as reported here, the collimators are likely air-gap type.

The remaining calculation is the path imbalance, commonly referred to as the PMD of the device. Here, the use of the term PMD is not precise, and while the industry will not likely change its terminology based on a small discrepancy, the discerning reader should know the difference. In the forward direction the u - and v -paths experience different refractive indices. Accordingly, one path is “fast” while the other is “slow.” Since the paths are split according to polarization, one polarization state is delayed with respect to the other. This is, precisely, differential-group delay (DGD). Polarization-mode dispersion results from the concatenation of multiple, non-aligned DGD elements and is characterized in most cases by a PMD vector that changes its pointing direction with frequency. This is not the case for a single isolator. In this work, the PMD of a component will be used when relating to industry usage; but otherwise DGD, or τ , is used. For the deflection-type isolator, note that in the forward direction the u experiences the ordinary index, while the v experiences the extraordinary. This is true through both crystals. The FR does not impart any significant differential delay. Since the deflection angle differences are small for a practical isolator, the wedge length l_w approximates the actual path. The differential-group delay τ between the two paths is

$$\tau \simeq \frac{2(n_{g,e} - n_{g,o})l_w}{c} \quad (6.3.12)$$

where $n_{g,e}$ and $n_{g,o}$ are the group indices of the e - and o -axes. A YVO_4 wedge 0.5 mm long at the base with a 9° wedge has a path length of approximately 0.35 mm on the thin side. The differential-group delay for an isolator made from this deflection-type component is $\tau \simeq 0.45$ ps (cf. §4.2.2). Substitution of LiNbO_3 for YVO_4 can further reduce the differential-group delay at the expense of an increase in gap length and wedge angle. For example, LiNbO_3 wedges of the same length yields a differential-group delay of $\tau \simeq 0.16$ ps.

Table 6.1. Deflection-Type Isolator Technology Comparison

Specification ^(a)	1-Stage	2-Stage	PMD-Comp	Units
Isolation ($\lambda_c \pm 15$ nm, 23°C, all SOP)	32	60	32	dB
Isolation ($\lambda_c \pm 15$ nm, 0-70°C, all SOP)	22	42	22	dB
Insertion Loss (λ_c , 23°C, all SOP)	0.2	0.4	0.3	dB
Insertion Loss (λ_c , 0-70°C, all SOP)	0.3	0.6	0.5	dB
PMD	0.20	0.05	0.02	ps
PDL (λ_c , 23°C)		0.05		dB
Return loss (λ_c , 23°C)		60		dB
Power handling		1,000		mW
Operating temperature		0-70		°C

^(a)Specification values reported by Koncent for C-band operation [9].

As a concluding remark, the deflection-type isolator can be generalized to a 2×2 port component to reduce size while increasing functionality [23].

6.4 Displacement-Type Isolators

Building on the spatial-walkoff method of polarization diversity first proposed in [12], Chang and Sorin developed a practical alternative to the Shirasaki isolator [2]. The Chang and Sorin isolator, herein called a displacement-type isolator, uses three birefringent walkoff crystals cut as parallelepiped blocks to achieve polarization independence. One of several variations of their isolator core is illustrated in Fig. 6.8 and its relation to the associated lensing system is found in Fig. 6.4(b).

The core is characterized by two walkoff blocks of length a and the remaining block of length $\sqrt{2}a$ placed in the sequence $\sqrt{2} : 1 : 1$. The Faraday rotator is located between the longer block and the first shorter block. All blocks are cut so as to impart beam walkoff between orthogonal linear polarization states (cf. §3.6.1). Accordingly, the extraordinary axis has vector components that lie both transverse and longitudinal to the optical path. To maximize that walkoff angle, the extraordinary axis is inclined into the crystal by about 45° ; the precise optimal angle, being material dependent, is calculated from (3.6.24) on page 117. For example, the optimal angle of inclination from the optical axis for YVO₄ and rutile are both $\alpha_{\max} \simeq 47.8^\circ$. The three walkoff crystals are cut so that the walkoff directions are different from one another. Following Fig. 6.8, the walkoff directions are $+90^\circ$, -45° , and $+45^\circ$ from front to back, respectively.

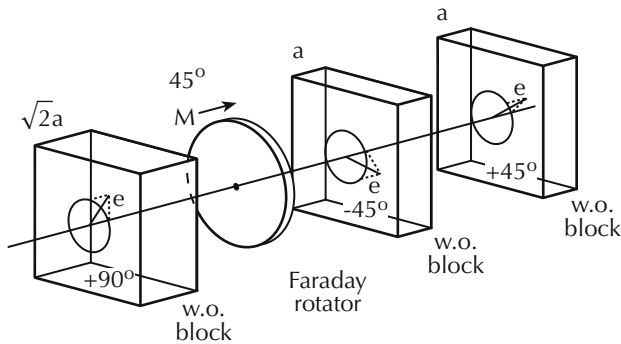


Fig. 6.8. Core of single-stage displacement-type isolator. a) Three birefringent walkoff blocks are configured as shown. The first block is $\sqrt{2}$ times longer than the other blocks. A Faraday rotator plate is located between first and second blocks. The cut of the extraordinary axes all maximize the walkoff and are oriented relative to one another as 90° , -45° , and -45° . The FR magnetization is typically held in saturation by a permanent magnet.

The principle of the displacement-type isolator differs from the deflection type. Within the displacement-type core, all beam paths run collinear between crystals. Within the crystals, polarizations are split and laterally shifted from one position to another. The interaction of the core with the lenses also differs because both offset and tilt are the principal means of spatial filtering. In order to minimize the crystal lengths, and corresponding spot-offset distances, a focusing lens system is used. The spatial distribution of spots in the transverse ray-trace diagram in the image plane (located within the core) is imaged onto the object plane (at the fiber face) by the receiving lens with a magnification factor M^{-1} . Only spots that fall within the fiber aperture are coupled, the remaining spots and associated beam paths are lost.

The forward and reverse ray-trace and spot diagrams are illustrated in Fig. 6.9. In the forward direction, walkoff block w_1 splits the two polarizations and translates the extraordinary polarized light along the v -path to the top line. The FR rotates both u - and v -path polarization states by $+45^\circ$. Walkoff block w_2 translates the v -spot to the point indicated in frame (d) of Fig. 6.9(a). Lastly, walkoff block w_3 translates the u -spot to overlap with the v -spot. Running collinear and along the L_2 lens axis, the combined forward beam is coupled into the output fiber.

In the reverse direction, the ray-trace through blocks w_2 and w_3 is the same as along the forward path. After the FR, however, the beams diverge. The Faraday rotator imparts a $+45^\circ$ rotation to the u' - and v' -path polarization states, creating a misalignment to the extraordinary axis of block w_1 . The misalignment causes the v' -path to trace a straight line through last block and the u' -path to walkoff in a downward direction. The resultant spot dia-

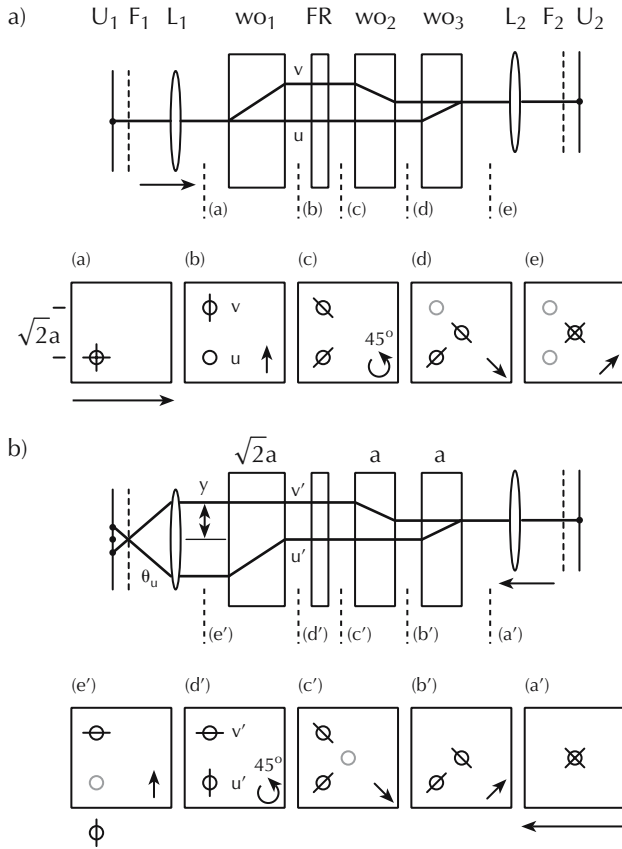


Fig. 6.9. Ray-trace and spot-trace diagrams for forward and isolation directions. a) Forward path splits polarizations along paths u - and v -paths. These beams, always collinear between blocks, converge at the output lens. b) Isolation path prevents beam convergence at the return lens and instead displaces both beams out of the aperture of the return fiber.

gram before lens L_1 is shown in frame (e') of Fig. 6.9(b). With proper design the u' - and v' -spots fall outside of the fiber face and are lost.

There are several interlocking calculations required for the displacement-type isolator. In the forward direction the fiber-to-fiber coupling must have maximum transmission. While the fiber-to-fiber magnification is unity the single-lens magnification sets the Rayleigh length, which in turn should be on the same order as the lens-to-lens gap. In the reverse direction the requisite isolation determines, in conjunction with the lens magnification, the necessary spot offset. The spot offset in turn determines the unit crystal length a . Finally, the path-length imbalance, or DGD, is calculated for the forward direction.

In the forward direction, the depth-of-focus $2b$ should be on the same order as the lens-to-lens gap L_{gap} to ensure good coupling. In view of the focusing diagram in Fig. 5.16(b), the gaussian beam waist w_b between the lenses is related to the lens magnification and fiber beam waist as $w_b = Mw_o$. Using the expression for the confocal parameter (5.2.10) on page 222, the magnification is

$$M = \frac{1}{w_o} \sqrt{\frac{2b}{k_o}} \quad (6.4.1)$$

Note that the wavenumber k_o here is for free space.

In the reverse direction the requisite isolation along with the lens attributes determine the necessary spot offset. Referring to Fig. 6.9(b), the offset in frame (e') maps to beam offset and tilt at the object plane U_1 . The tilt angle θ_u is related to the offset via the focal length:

$$\theta_u = \frac{y_o}{f} \quad (6.4.2)$$

Moreover, the offset y_o to the right of the lens relates to the offset y_i on the U_1 plane through the magnification

$$y_o = My_i \quad (6.4.3)$$

Folding these relations into the overlap integral (5.4.10) on page 243, the offset is related to the isolation, lens, and fiber parameters via

$$y_o = \frac{2w_o M}{\sqrt{1 + \left(\frac{k_o n w_o^2 M}{f}\right)^2}} \sqrt{-\ln \mathcal{I}_b} \quad (6.4.4)$$

As the mode overlap occurs within the fiber glass, the wavenumber k_o in (6.4.4) is scaled by the fiber index n . Finally, the u' - and v' -spot offsets from the axis of lens L_1 are both

$$y_o = \sqrt{2}a\gamma \quad (6.4.5)$$

where the walkoff angle γ is determined in general from (3.6.15) on page 110, or from (3.6.25) on page 117 for maximum walkoff. For example, the maximum walkoff in a YVO_4 crystal is $\gamma_{\text{max}} \simeq 5.70^\circ$.

Together, equations (6.4.1), (6.4.4), and (6.4.5) determine the crystal thickness a . For example, consider an approximate solution using YVO_4 where $I_{\text{iso}} = -45$ dB, $2b = 5$ mm, $f = 1$ mm, $n = 1.44$ (of the fiber), and $\omega = 2\pi \times 194.1$ THz. The requisite magnification is $M \simeq 7.0$ mm. This in turn sets the necessary spot offset to $y_o \simeq 156$ μm . For this magnification, the minimum beam waist between the lenses is $2w_b \simeq 70$ μm and the displacement by the walkoff crystals is $y_o \simeq 2.2 \times (2w_b)$. Finally, the unit block length given the above walkoff angle is $a \simeq 1.1$ mm. As a check, the total length of the blocks is $(2 + \sqrt{2})a \simeq 3.8$ mm, which is a bit less than the depth-of-focus.

The remaining calculation is the path imbalance in the forward direction. It should be clear that the differential-group delay through walkoff blocks w_{o2} and w_{o3} cancel. The imbalance is solely due to walkoff block w_{o1} . Within this block, the u -path sees the ordinary group index $n_{g,o}$, while the v -path sees the effective group index. The effective refractive index is a mixture of the ordinary and extraordinary refractive indices, the mixture set by the inclination angle α of the extraordinary from the optic axis, see (3.6.26) on page 117. The differential-group delay requires the effective group index, which generally is approximately the same as the refractive index for high birefringent crystals in the near infrared. Nonetheless, keeping track of the proper indices, the differential-group delay τ is

$$\tau = \frac{\sqrt{2}a (n_{g-\text{eff},e} - n_{g,o})}{c} \quad (6.4.6)$$

Continuing with the example and using an effective refractive index, use of YVO₄ walkoff crystals imparts a differential-delay of $\tau \simeq 1.1$ ps. There is, in fact, a simple alternation that eliminates τ for a two-stage block-type isolator. This is detailed in §6.6.

As a concluding remark, folded architectures of the one-stage block-type isolator have been proposed wherein a single relay lens with a mirror backing is placed at a midpoint along the polarization evolution [4, 8, 15]. These isolators can be very compact, but none of the reports demonstrates an epoxy-free path.

6.5 Two-Stage Isolators

To increase the isolation over a wide temperature and wavelength range, two or more isolation stages are necessary. Two-stage isolators were already considered in §6.1 with respect to the garnet design. However, that discussion did not include polarization diversity. Nonetheless, the lessons from that earlier section apply here just as well.

Both the deflection-type and displacement-type isolators are expandable to two stages. Rather than blindly cascade two like stages, however, economies can be found to improve the performance or reduce the size of a two-stage component.

An example of a two-stage deflection-type isolator is illustrated in Fig. 6.10. Here the extraordinary axes of second wedge pair are rotated by 90° with respect to the first wedge pair. As indicated in the ray-trace diagram, this rotation leads to a convergence of the u - and v -paths in the forward direction. This is an improvement in PDL and alignment sensitivity over the single-stage deflection-type isolator. Moreover, the path imbalance along the forward path is cancelled. That is, the expected differential-group delay through the component is zero, although in practice differences in where the light paths intersect the the wedges leaves a residual imbalance. Table 6.1 shows the reported reduction in PMD value for a two-stage deflection-type isolator.

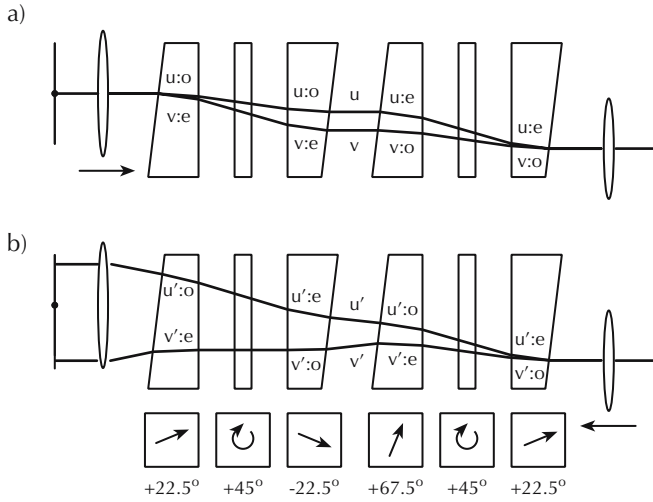


Fig. 6.10. A two-stage deflection-type isolator, one particular realization. To recombine the forward beams and to balance the u - and v -path lengths, the extraordinary axes of the second stage wedges are cut 90° in rotation from the first stage. a) Forward direction ray-trace. b) Isolation direction ray-trace. The deflection angles are twice that of the single-stage counterpart.

In the reverse direction, the u' - and v' -beams are deflected twice. The total deflection for either beam is therefore

$$\theta_d = \pm 2(n_e - n_o)\theta_w \quad (6.5.1)$$

Since the mode coupling goes exponentially with tilt angle, some redesign from a one-stage deflection may be possible to reduce the wedge angle.

A two-stage displacement-type isolator, in the configuration proposed by Chang and Sorin [3], is illustrated in Fig. 6.11. The $\sqrt{2}a$ block is placed between the two a blocks instead of in front. All three extraordinary axes are cut to maximize the walkoff, as before. Two FRs having the same rotary direction are added as indicated, and the walkoff direction in the plane perpendicular to the optic axis changes by 45° from one block to the next. Unlike the two-stage deflection-type isolator, this isolator does not increase the spatial filtering of the principal rays. However, the isolation is increased because light is scattered into more locations. Also, the differential-group delay is reduced, but not eliminated. Figure 6.11(a,b) details the principal and “error” paths along the forward and reverse directions, while (c) shows a detailed spot-evolution diagram. The error-paths originate from incorrect Faraday rotation. The differential-group delay for this two-stage isolator is

$$\tau = \frac{(\sqrt{2} - 1)a\Delta n_g}{c} \quad (6.5.2)$$

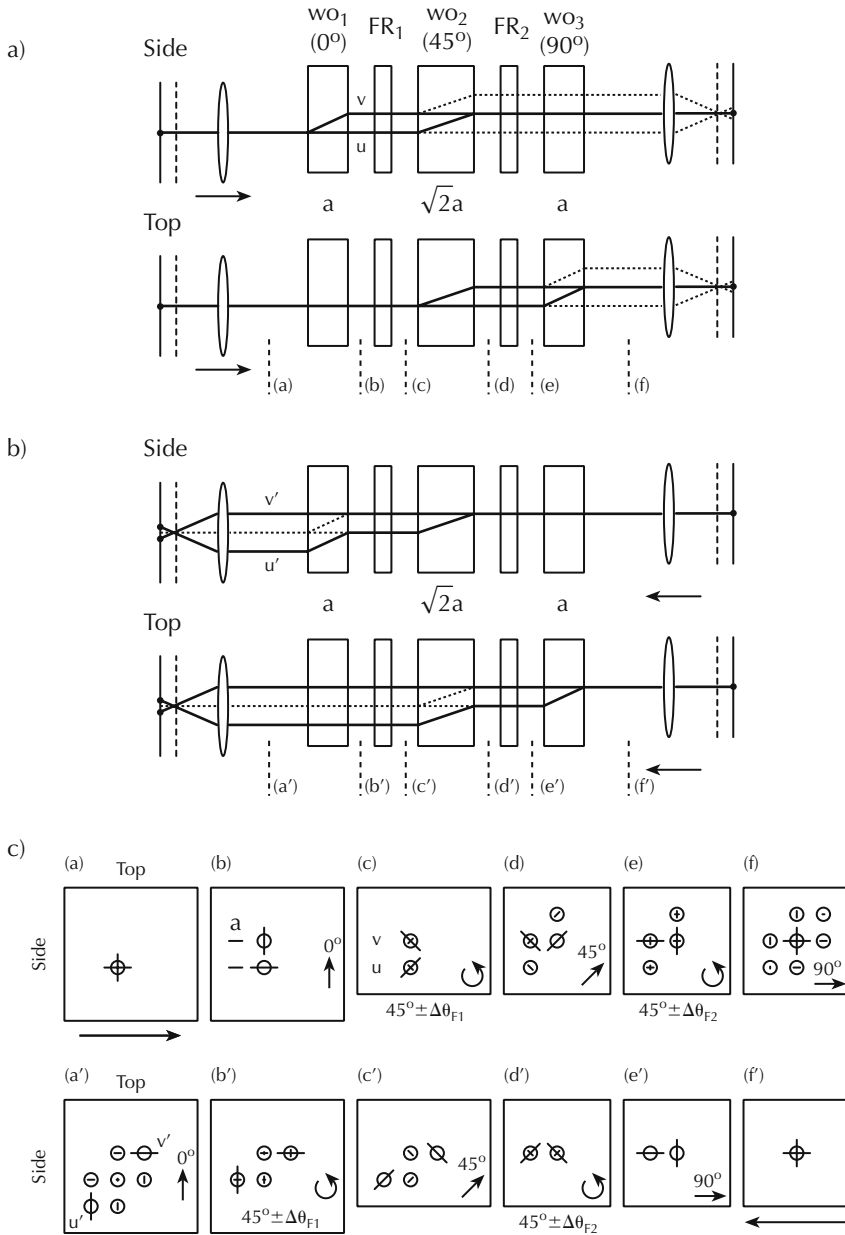


Fig. 6.11. Ray-trace and spot-trace diagrams for forward and isolation directions. This two-stage displacement isolator places walkoff blocks in a $1 : \sqrt{2} : 1$ sequence, with FRs located between blocks. Solid lines are nominal beam paths; dashed lines are error paths that occur when the Faraday rotation is not precisely 45° . a) Side and top views of forward paths. b) Side and top views of isolation paths. c) Spot-trace diagrams that include residual light from imperfect Faraday rotation.

where, as shown in the figure, the u -path somewhat cancels the differential-group delay from the v -path.

6.6 PMD-Compensated Isolators

The two-stage deflection-type isolator (Fig. 6.10) illustrated one way to compensate for the differential-group delay, colloquially known as PMD, in that configuration. There are other significant PMD-compensation techniques for one-stage isolators, both for deflection and displacement types.

For a one-stage deflection-type isolator, two methods have been invented to compensate for differential-group delay and differential beam displacement. Swan proposes the addition of a rhombohedral crystal with its extraordinary axis cut in the plane perpendicular to the optical path [24, 25] (Fig. 6.12(a)). The rhombohedral angle is set to combine the u - and v -paths through differential refraction. Alternatively, Xie proposes the addition of a parallelepiped crystal with its extraordinary axis cut as a walkoff block [7], Fig. 6.12(b). The parallelepiped length and extraordinary axis compensation angle are set concurrently to eliminate the DGD and differential beam displacement. There is some advantage of the Swan method over the Xie method because the former scheme lets the two wedge prisms be the same part where the latter method is best suited for two different prism parts.

Figure 6.12(a) shows the orientation of the three birefringent crystals in the Swan scheme. As in §6.3 the extraordinary axis of each wedge is cut at 22.5° . The orientation of the second wedge with respect to the first, as shown in the figure, automatically positions the second e -axis at 45° with respect to the first. In the forward path, the v -path refracts more than the u , assuming the wedges are made of positive uniaxial crystal. Concomitant with the greater refraction is an arrival-time delay of the v -path with respect to the u -path. In the forward direction the FR ensures that the v -path sees the e -axis of both wedges. Therefore, the ignoring the second-order correction for the refraction angle, DGD is accrued over length $2l_w$.

Accordingly, the rhombohedral compensation crystal is cut as a waveplate with its e -axis rotated 90° to the e -axis of the second wedge. That is, the extraordinary path through the wedge pair becomes the ordinary path through the compensation crystal. The waveplate cut ensures that the e -axis lies in the plane perpendicular to the optical path. With length $l_c = 2l_w$ (under that assumption that the wedge and compensation parts are the same material) the DGD is eliminated to first order. Second-order corrections can be made to account for the refracted paths in the crystals.

The differential beam displacement $d_v - d_u$ can also be corrected by cutting the compensation crystal to rhombohedral angle θ_r . In general, the refraction angle into the crystal depends on the input polarization. The e - and o -ray refraction angles in the small-angle approximation are

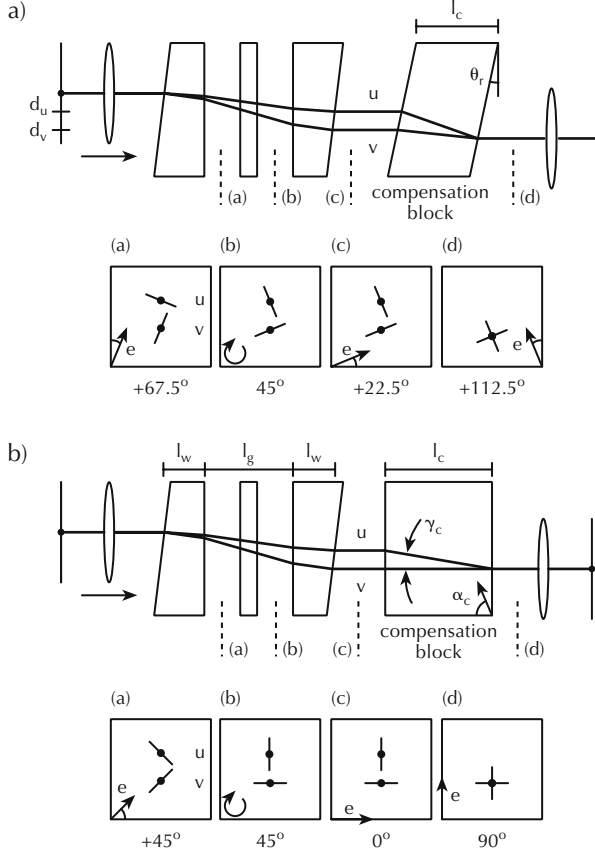


Fig. 6.12. Forward ray-trace diagrams of Swan and Xie PMD-compensated deflection-type isolators. a) Swan method uses compensation crystal cut as a waveplate and rhombohedral angle θ_r to compensate the DGD from the wedge pair and to refract differentially the u - and v -paths. b) Xie method uses compensation crystal cut as a partial walkoff and partial waveplate parallelepiped to compensate the DGD from the wedge pair and to translate the u -path onto the v -path.

$$\theta_{e,o} = \left(\frac{1}{n_{e,o}} - 1 \right) \theta_r \quad (6.6.1)$$

The difference in displacement through the compensation crystal is therefore

$$d_e - d_o = l_c \theta_r \left(\frac{1}{n_e} - \frac{1}{n_o} \right) \quad (6.6.2)$$

The differential beam displacement due to the wedge pair, (6.3.4) on page 256, is cancelled by the compensation crystal when the two displacements are equal. The required rhombohedral angle is therefore

$$\theta_r = \frac{d_v - d_u}{l_c \left(\frac{1}{n_e} - \frac{1}{n_o} \right)} \quad (6.6.3)$$

Precisely speaking, a non-zero rhombohedral angle requires the plane in which the extraordinary axis of the compensation crystal lies to be tilted so that the refracted paths remain perpendicular to that plane. This prevents walkoff.

Figure 6.12(b) shows the orientation of the three birefringent crystals in the Xie scheme. There are two mechanisms that simultaneously affect the light through the compensation crystal. One is the walkoff experienced by one path because the e -axis of the crystal is cut with a vector component along the optical path. The other is differential-group delay experienced because one path sees the ordinary index and the other sees an effective index that depends on the inclination angle of the e -axis. The compensation crystal length l_c and e -axis inclination angle α_c are tailored to compensate the DGD and differential beam displacement.

Two equations with two free parameters together determine the length and e -axis angle of the compensation crystal. The DGD of the crystal is

$$\tau_c = \frac{l_c (n_g(\alpha_c) - n_{g,o})}{c} \quad (6.6.4)$$

and the differential beam displacement is

$$d_v - d_u = l_c \tan \gamma_c(\alpha_c) \quad (6.6.5)$$

The effective index $n_{\text{eff}}(\alpha_c)$ is given by (3.6.26) on page 117, or for normal incidence,

$$n_{\text{eff}} = \frac{n_e n_o}{\sqrt{n_e^2 \cos^2(\alpha_c) + n_o^2 \sin^2(\alpha_c)}} \quad (6.6.6)$$

The walkoff angle, (3.6.15) on page 110, is also governed by the inclination angle:

$$\tan \gamma_c = \frac{(n_e^2 - n_o^2) \sin \alpha_c \cos \alpha_c}{n_e^2 \cos^2 \alpha_c + n_o^2 \sin^2 \alpha_c} \quad (6.6.7)$$

As a technical point, the group indices are used in (6.6.6) while the refractive indices are used in (6.6.7). The two free parameters in these two equations are l_c and α_c . Given the DGD from the wedge pair and the displacement $d_v - d_u$, a unique solution can be found.

Note that a shortcoming of the Xie design is that the e -axis orientation in the two wedges is different than before (Figure 6.12(b)). To recombine the u - and v -paths perfectly, the linear polarization state orientation should be parallel and perpendicular to the walkoff direction.

Table 6.1 shows the improvement of a single-stage deflection-type PMD-compensated isolator over single-stage deflection-type isolator. It is not known which of the two compensation schemes is used for this product.

Displacement-type isolators can also have PMD-compensation. The best overall performance with the fewest parts is a two-stage configuration similar to the original Chang and Sorin proposal [3] but with a variation proposed by Konno [10] and further improved here. Using the spot-vector diagrams so clearly presented in [11], four displacement-type isolators and their behavior are illustrated in Fig. 6.13. Another variation of the spot diagram is found in [14], and an alternative PMD-compensation scheme for displacement-type isolators is reported in [27].

The original Chang and Sorin device [2] is shown in Fig. 6.13(a). The associated spot-vector diagrams trace the beam locations through the device in a plane perpendicular to the optical path. In the forward direction, both input polarizations begin at point P_o . The first and second walkoff blocks translate the v -path along vectors p_{v1} and p_{v2} , respectively. The third block translates the u -path along vector p_u to combine the two beams. In the reverse direction, the input polarizations begin at point P'_o and trace the indicated paths. Both beams are displaced out of the P_o aperture, leading to isolation.

The path lengths p_u and p_v are analogues for the DGD accrued through the isolator core. The optical phase of either path through a crystal is $\phi = \omega n L / c$. The phase difference between a walkoff path and the associated straight-through path is $\Delta\phi = \omega(n_{\text{eff}} - n_o)L/c$, or, accounting for the walkoff angle γ ,

$$\Delta\phi = \frac{\omega}{c}(n_{\text{eff}} - n_o)p_{u,v} \cot \gamma \quad (6.6.8)$$

Provided that the e -axis inclination angle is the same for each crystal, ensuring that n_{eff} and γ are constant, the path-length difference $p_v - p_u$ is proportional to the DGD. The first Chang and Sorin isolator has a DGD that goes as $\tau \propto \sqrt{2}a$, where a is the unit crystal length.

Clearly, then, the two-stage Chang and Sorin component [3] (Fig. 6.13(b)) has a DGD that scales as $\tau \propto (2 - \sqrt{2})a$. Recognizing the importance of path balancing, Kuzuta proposes the modified two-stage block isolator shown in Fig. 6.13(c). Addition of the second $\sqrt{2}a$ -length block translates both u - and v -paths equally. As shown in the diagram, $p_{u1} + p_{u2} = p_{v1} + p_{v2}$.

A more economic and elegant approach is shown in Fig. 6.13(d), which is a variation of the Konno proposal [10]. The path lengths are balanced by increasing the effective index of the p_u in relation to the v -paths. Using (6.6.8) and $\tau = d\Delta\phi/d\omega$, the path lengths are balanced when

$$(n_{\text{g-eff}}(\alpha_c) - n_{\text{g,o}}) \cot \gamma_c = \sqrt{2} (n_{\text{g-eff}} - n_{\text{g,o}}) \cot \gamma \quad (6.6.9)$$

The center crystal has to be lengthened accordingly. This scheme is a small change to the original Chang and Sorin two-stage proposal but with the significant improvement in DGD mitigation.

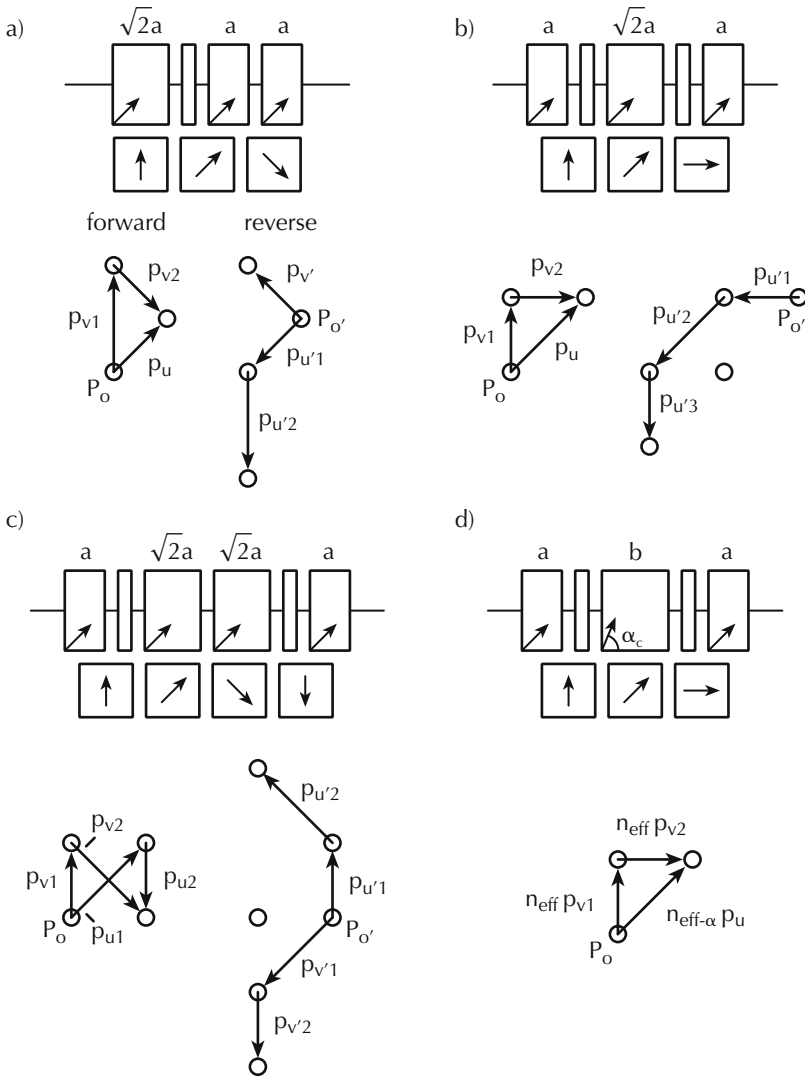


Fig. 6.13. Displacement isolators, non-PMD-compensated (a,b) and PMD-compensated (c,d), with corresponding spot vector diagrams. a) Single-stage displacement isolator as before. Forward path lengths are such that $p_{v1} + p_{v2} \neq p_u$, generating PMD. b) Dual-stage displacement isolator as before. c) PMD-compensated dual-stage displacement isolator. Forward path lengths are equalized between the two polarizations. d) Elegant PMD-compensated dual-stage displacement isolator. Orientation of extraordinary axis in second block increases the u -path delay to equalize to the v -path.

References

1. D. W. Anthon and D. L. Sipes, "Multi-function optical isolator," U.S. Patent 6,088,153, July 11, 2000.
2. K. W. Chang and W. V. Sorin, "Polarization independent isolator using spatial walkoff polarizers," *IEEE Photonics Technology Letters*, vol. 1, no. 3, pp. 68–80, 1989.
3. —, "High-performance single-mode fiber polarization-independent isolators," *Optics Letters*, vol. 15, no. 8, pp. 449–451, 1990.
4. Y. Cheng and G. S. Duck, "Multi-stage optical isolator," U.S. Patent 5,768,005, June 16, 1998.
5. D. J. Gauthier, P. Narum, and R. W. Boyd, "Simple, compact, high-performance permanent-magnet faraday isolator," *Optics Letters*, vol. 11, no. 10, pp. 623–625, 1986.
6. A. J. Heiney and D. K. Wilson, "Optical isolators employing oppositely signed faraday rotating materials," U.S. Patent 5,087,984, Feb. 11, 1992.
7. Y. Huang, P. Xie, X. Luo, and L. Du, "Optical isolator with reduced insertion loss and minimized polarization mode dispersion," U.S. Patent 2002/0 060 843, May 23, 2002.
8. R. S. Jameson, "Polarization independent optical isolator," U.S. Patent 5,033,830, July 23, 1991.
9. "Micro optics for telecom catalog 2002," Kocent Communications, Fuzhou, Fujian, P.R. China, 2002. [Online]. Available: <http://www.koncent.com/>
10. Y. Konno, S. Aoki, and K. Ikegai, "Polarization independent optical isolator," U.S. Patent 5,774,264, June 30, 1998.
11. N. Kuzuta, "Optical isolator," U.S. Patent 5,237,445, Aug. 17, 1993.
12. T. Matsumoto, "Polarization-independent isolators for fiber optics," *Electronics and Communicatinos in Japan*, vol. 62-C, no. 7, pp. 113–119, 1979.
13. —, "Optical nonreciprocal device," U.S. Patent 4,239,329, Dec. 16, 1980.
14. H. Ohta and N. Nakamura, "Optical isolator," U.S. Patent 5,151,955, Sept. 29, 1992.
15. J.-J. Pan, "Highly miniaturized, folded reflection optical isolator," U.S. Patent 6,212,305, Apr. 3, 2001.
16. F. J. Sansalone, "Compact optical isolator," *Applied Optics*, vol. 10, no. 10, pp. 2329–2331, 1971.
17. K. Shirai, M. Sumitani, N. Takeda, and M. Ariei, "Optical isolator," U.S. Patent 5,278,853, Jan. 11, 1994.
18. K. Shiraishi, F. Tajima, and S. Kawakami, "Compact faraday rotator for an optical isolator using magnets arranged with alternating polarities," *Optics Letters*, vol. 11, no. 2, pp. 82–84, 1986.
19. K. Shiraishi and S. Kawakami, "Cascaded optical isolater configuration having high-isolation characteristics over a wide temperature and wavelength range," *Optics Letters*, vol. 12, no. 7, pp. 462–464, 1987.
20. K. Shiraishi, S. Sugaya, and S. Kawakami, "Fiber faraday rotator," *Applied Optics*, vol. 23, no. 7, pp. 1103–1105, 1984.
21. M. Shirasaki, "Optical device," U.S. Patent 4,548,478, Oct. 22, 1985.
22. M. Shirasaki and K. Asama, "Compact optical islator for fibers using birefringent wedges," *Applied Optics*, vol. 21, no. 23, pp. 4296–4299, 1982.
23. J. Y. Song, "Optical isolator," U.S. Patent 6,061,167, May 9, 2000.

24. C. B. Swan, "Optical isolator with polarization dispersion and differential transverse deflection correction," U.S. Patent 5,631,771, May 20, 1997.
25. —, "Optical isolator with polarization dispersion and differential transverse deflection correction," U.S. Patent 5,930,038, July 27, 1999.
26. T. Uchida and A. Ueki, "Optical isolator," U.S. Patent 4,178,073, Dec. 11, 1979.
27. T. Watanabe, S. Sugiama, and T. Ryuo, "Multiple-stage optical isolator," U.S. Patent 6,049,425, Apr. 11, 2000.
28. C. G. Young, "Multiple wavelength optical isolator," U.S. Patent 3,602,575, Aug. 31, 1971.

Circulators

An optical circulator is a generalized isolator having three or more ports. While an isolator causes loss in the isolation direction, a circulator collects the light and directs it to a nonreciprocal output port. Figure 7.1 illustrates several possible circulator configurations. Figure 7.1(a) illustrates the port mapping for a four-port circulator. The ports cyclically map $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$. This is called a strict-sense circulator because every input port has a specific nonreciprocal output port. Construction of a strict-sense circulator with more ports becomes inelegant but ones with three ports can be simple [22]. Figure 7.1(b) illustrates a non-strict-sense circulator having any number of ports greater than two. In this case each input port has a specific nonreciprocal output port except for the last port; the light input to the last port is lost. The ladder diagram reflects the optical path within the component and indicates the disconnect between the first and last ports. Figure 7.1(c) illustrates a three-port non-strict-sense circulator. This circulator has significance in telecommunications applications because return of light from port 3 to port 1 is often not necessary. For instance, the reflected light from a fiber Bragg grating need only be separated from the input light without loss, but as optical links are not typically operated in reverse there is no need for strict-sense behavior.

Other than architecture, most of the considerations for the optical circulator have already been addressed in preceding chapters. As a nonreciprocal device, a circulator has at least one Faraday rotator (FR) in the optical path. The wavelength and temperature performance of FRs was treated in §6.1. In that same section, the transmission and isolation as a function of polarizer alignment was treated. The modern deflection-type circulators use birefringent prism combinations such as the Wollaston and Rochon prisms, detailed in §4.7, along with dual-fiber collimators, detailed in §5.3. Likewise, the differential-group delay due to path-length imbalance was treated in §6.6. Accordingly, all the tools necessary to appreciate and design optical circulators have been earlier developed, allowing the current chapter to focus exclusively on architectures and performance.

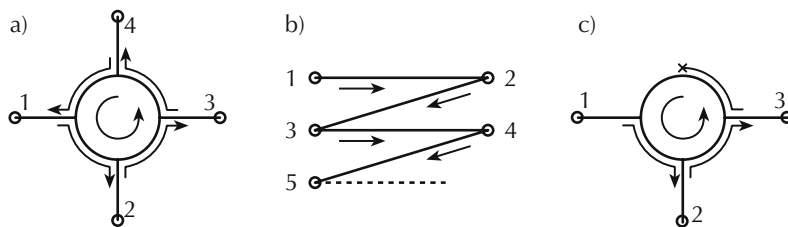


Fig. 7.1. Three types of circulator port connections. a) Strict-sense circulator with four ports. Each input port has a specific nonreciprocal output port. b) Non-strict-sense circulator in ladder topology. Any number of ports greater than two is possible; however, light input to the last port is lost. c) Non-strict-sense three-port circulator. This topology has significant applications in unidirectional telecommunications links and has good economies compared with (a) or (b).

As with isolators, circulators can be polarization dependent or polarization independent. The polarization-dependent circulator is an important starting point because the minimum requirements for circulatory behavior are clear. Polarization-independent circulators are further categorized as displacement-type and deflection-type. Except for the earliest designs, displacement-type circulators use birefringent walkoff crystals to achieve polarization diversity. Deflection-type circulators use birefringent prisms and dual-fiber collimators for the same purpose. The most compact and least expensive circulators use deflection-type designs.

7.1 Polarizing Circulator

Figure 7.2 illustrates a YIG-based polarizing circulator first demonstrated by Shibukawa [32, 33]. This configuration is the minimum necessary to achieve circulatory behavior and should be compared to the polarizing isolator in Fig. 6.1 on page 248. Circulatory behavior exists only for prescribed linear input polarization states. The main issue addressed by Shibukawa was the addition of input and output ports on either side of the FR, as compared to a polarizing isolator, to realize a strict-sense circulator. To create two ports on either side of the FR, Glan-Taylor prisms were used (Fig. 7.2(c)). The Glan-Taylor prisms were used for two reasons. The inventors realized the importance of high polarization contrast yet thin-film polarization beam-splitting cubes were not readily available nor did they exhibit good performance. Moreover, as compared with the Glan-Thompson, Wollaston, and Rochon prisms, the Glan-Taylor prism has a wide deflection angle, e.g. 110° for calcite. This enabled them to build a relatively compact components.

A Glan-Taylor prism is a birefringent prism pair cut so that along the hypotenuse one polarization state, e.g. the extraordinary ray for a positive uniaxial crystal, experiences total-internal reflection at the crystal/air bound-

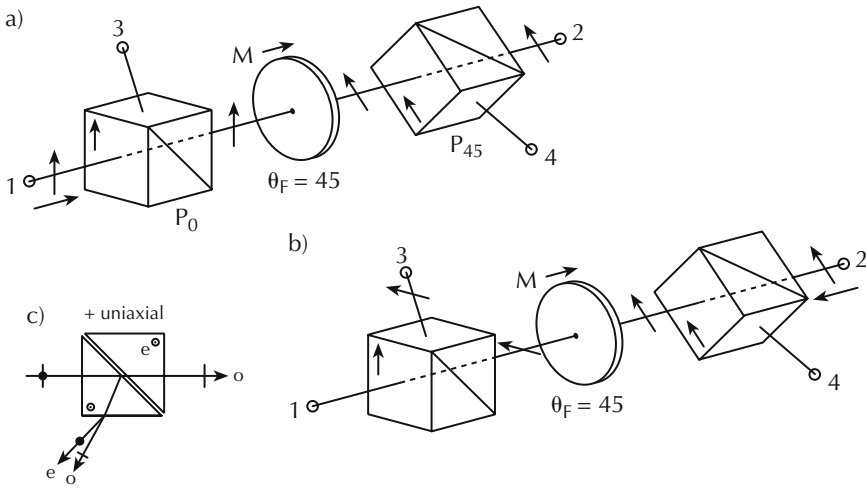


Fig. 7.2. A polarizing circulator in its simplest embodiment. Two polarization-dividing prisms and a Faraday rotator are used. The prisms are rotated by 45° with respect to one another along the longitudinal axis. a) Transmission of $1 \rightarrow 2$ for linear vertical polarization input. Linear horizontal input is deflected by prism P_0 . Light after the FR not aligned to prism P_{45} is output on port 4, reducing path isolation. b) Transmission of $2 \rightarrow 3$ for linear polarization input. Light after the FR not aligned to prism P_0 is output on port 1, reducing path isolation. c) A Glan-Taylor birefringent prism, where e -rays are totally internally reflected and o -rays are partially reflected at the crystal / air interface along the hypotenuse.

ary while the orthogonally polarized ray exits interface. The birefringent prism can be cut at Brewster's angle to maximize the transmission, but such was not the case in the work of Shibukawa. Accordingly, reflected ordinary light co-propagates with the TIR extraordinary light but is refracted at a different angle upon exiting the crystal. It should be noted that in most early demonstrations none of the optical interfaces were anti-reflection coated.

Referring to Fig. 7.2, all that is required for minimal circulatory action is a Faraday rotator with $\theta_F = 45^\circ$, an input polarization splitter, and an output polarization splitter rotated by 45° . To pass from port 1 to port 2, a vertically aligned linear polarization state is input. This state transits the first polarization splitter, is rotated by the FR, and transits the second polarization splitter. In the reverse direction, the same linear polarization state now input to port 2 is again rotated by the FR and is diverted by the first polarization splitter to port 3. Further path tracing shows that this is a strict-sense polarizing circulator.

With high-quality thin-film polarization beam-splitting cubes and high-performance iron garnet materials now available, the polarizing circulator Fig. 7.2 can exhibit good performance other than PDL. However, at the time, losses were incurred through lack of AR coatings, poor extinction ratio of the

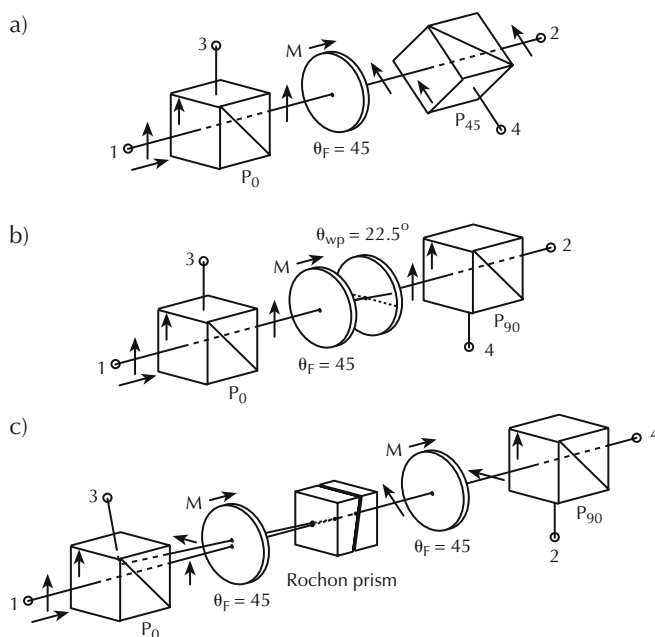


Fig. 7.3. Conceptual development of polarizing circulators. a) Simple polarizing circulator like that of Shibukawa; PBS cubes illustratively replace the Glan-Taylor prisms. The four ports do not lie in a plane. b) Addition of reciprocal element, here a half-wave waveplate, allows all four ports to lie in a plane. The polarization plane can be rotated by 90° . c) A two-stage polarizing circulator. The reciprocal element of (b) is replaced with a second FR and a polarizing beam splitter, such as a Rochon prism, located between FR plates. The polarization plane is rotated 90° along the forward path. The polarizing-cube, FR, prism, FR, polarizing-cube sequence forms two-stages of isolation.

prisms, losses in the YIG garnet, and poor fiber coupling. The performance reported at the time is an insertion loss of 2 dB and an isolation variation from 13 dB to 28 dB, depending on port combination. The authors were cognizant of wavelength dependence but made no reports on temperature dependence.

The polarizing circulator leads to a conceptual framework that encompasses essential aspects of any circulator design. The circulator of Shibukawa, or one similar as in Fig. 7.3(a), where the Glan-Taylor prisms are illustratively replaced with polarization beam splitting (PBS) cubes, has four ports that do not lie in a plane. This makes the component form-factor less convenient. To rectify this shortcoming, the FR can be preceded or followed by a reciprocal element such as a half-wave waveplate, having its birefringent axis at $\pm 22.5^\circ$ with respect to the cube axis, or an optically active crystal that rotates the linear polarization state by 45° (Fig. 7.3(b)). In either case, from the FR

looking through the reciprocal plate, the PBS cube that is in view appears rotated by 45° . Optically the cube is rotated but physically it is not, allowing all ports to lie in the same plane.

The reciprocal and nonreciprocal rotators together can rotate the plane of linear polarization by 90° . The 90° rotation is characteristic of most circulators. However, as a general rule a waveplate reduces the bandwidth of a component. Whether such limitation is tolerable or not depends on many factors. Yet a better method is to use two Faraday rotators with an intermediate polarizing beam splitter (Fig. 7.3(c)). Note that other than material dispersion, the second FR has the same wavelength dependence as an equivalent reciprocal rotator. However, the dual FR design accomplishes two goals: the plane of polarization is rotated by 90° in the forward direction and 0° in reverse, and the isolation of the circulator is squared because this is a two-stage circulator. A two-stage circulator is a natural consequence of placing all ports on the same plane.

7.2 Historical Development

Circulator architectures are characterized by rapid development toward an “optimal” design – given available materials, sub-assemblies, and recognized application-specific requirements – and plateaus during which little changed. Very early work in optical circulators circa 1960s derived motivation from radio-frequency circulators which, at the time, required only single-polarization performance [28, 31]. The late 1970s are characterized by the application-specific realization that single-mode optical fiber does not preserve polarization and therefore circulators have to be polarization-independent [17, 29, 36]. Indeed early descriptions claimed a 3 dB loss for a component with infinite polarization-dependent loss (PDL) [17]. Optical circulators thus transformed from polarization-dependent (PD) to polarization-independent (PI) via polarization-diversity schemes. The goal at the time was to achieve high isolation and low loss, which in turn required the development of high-contrast polarization splitters. Specifications of wavelength-dependence and PDL were also recognized at the time, but performance was often poor by today’s standards. Temperature dependence was almost never referred to (the exception being [36]) as well as differential-group delay. Lastly, the early PI circulators were strict-sense four-port designs.

Figure 7.4 illustrates the architectural development between 1978 and 1981 from PD circulators to PI circulators. Indeed the submission dates of the articles are so close it is difficult to reconstruct precisely who invented what first. Figure 7.4(a) illustrates the Shibukawa PD circulator detailed in the preceding section. Figures 7.4(b–d) are all PI circulators with an additional important distinction. In all three topologies the four ports lie in the same plane. To do this a reciprocal polarization rotator was added: a half-wave waveplate was used by Matsumoto and Shirasaki, while an optically active

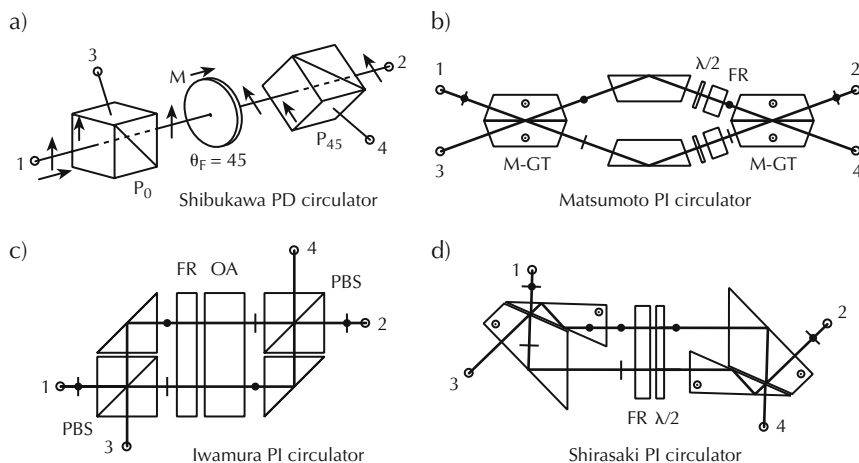


Fig. 7.4. Evolution of early four-port strict-sense circulators. Goals were to provide polarization-independent circulatory behavior with high isolation. a) The polarizing circulator circa 1978. b) First polarization-independent (PI) proposal circa 1979. Modified Glan-Thompson (M-GT) prisms and dual half-wave waveplate plus FR pairs were used. c) Alternative PI proposal circa 1979. Thin-film polarization beam splitters were used, along with optically active quartz for the reciprocal 45° rotation. d) Shirasaki circulator circa 1980 using high-extinction-ratio Shirasaki polarization splitters, an FR and a half-wave waveplate.

(OA) rotator was used by Iwamura. At the time an OA rotator was considered by some as advantageous because of easy alignment [17], although the required crystal length for quartz is 15.8 mm at $1.55 \mu\text{m}$ [20]. In either case, the addition of the reciprocal rotator reduces the bandwidth of the component.

The Matsumoto PI circulator [28, 29] in Fig. 7.4(b) uses modified Glan-Thompson (M-GT) birefringent prisms to separate the polarization. Similar to the Glan-Taylor prism, the Glan-Thompson prism extracts one polarization component through TIR. The latter prism has the gap between the two prism sections filled with a bonding agent such as epoxy to reduce the angular deflection of the TIR light. Using calcite, Matsumoto reported a 36.4° full-angle deflection. Unlike the Shibukawa PD circulator, the present circulator captures both transmitted and deflected light and directs the two paths through separate half-wave and FR pairs. The waveplate was a true zero-order half-wave waveplate and the FR was YIG with an Sm-Co permanent magnet for saturation. The polarizations output from the rotators are combined by a second Glan-Thompson prism. The reported insertion loss was 3.7 dB.

One principle drawback of the Matsumoto scheme is that the modified Glan-Thompson prisms reflected -12 dB of the non-TIR light in the direction of the TIR light. This made for a very low isolation floor. The other drawback is the duplication of the reciprocal and nonreciprocal rotators.

The Iwamura PI circulator [17] in Fig. 7.4(c) is a far more suitable architecture, but the inventors were limited by the low-quality thin-film polarization beam splitters. The significant improvement is a single reciprocal/nonreciprocal rotator pair through which both optical paths transit. While none of the optical surfaces were anti-reflection coated, the inventors reported an insertion loss of $1.2 - -1.6$ dB and an isolation of $16 - -19$ dB.

The Shirasaki circulator [22, 36] in Fig. 7.4(d) employs the Shirasaki birefringent prism (cf. §4.7) to act as a high-extinction ratio polarization splitter. Taken as a pair, the rutile prisms exhibited over 40 dB contrast with a loss less than 0.5 dB. Like the Iwamura isolator, the two optical paths run parallel and transit a single rotator pair. The FR was YIG and the half-wave waveplate was a true zero-order quartz plate. It should be noted that the birefringent axis of the waveplate was inclined by 22.5° in the plane perpendicular to the optical path. The resultant circulator had $0.4 - -0.6$ dB insertion loss and $25 - -32$ dB isolation. Moreover, the inventors characterized their circulator over a $5 - -45^\circ$ temperature range and demonstrated a 0.3 dB insertion loss shift and ± 1 dB isolation variation.

Emkey [7, 9, 10] plays a pivotal role in the development of circulators for two reasons. First, he recognized that birefringent walkoff crystals have a higher extinction ratio than did the polarization-splitting prisms that preceded him. Second, he recognized that for optical communication links a circulator need not be a strict-sense four-port component, but rather a non-strict-sense three-port circulator was satisfactory and certainly more economical. While his component design is awkward and not repeated here, Emkey set the stage for Koga, Fujii, Xie, and others to develop displacement-type circulators in the early 1990s. Displacement-type circulators also incorporated superior iron garnet materials, specifically, the Bi:RIG garnets being developed at the time, and superior single-fiber collimators. A large advance in performance was thus recorded.

The final substantial improvement come about in the late 1990s with the development of the dual-fiber collimator. The dual-fiber collimator simplified and miniaturized the housing size of the lenses. Equally importantly, due to its convenient interaction with Wollaston, Rochon, and Kaifa compound prisms, the use of dual-fiber collimators ushered in a family of deflection-type schemes that substantially reduced the necessary volume of birefringent material, which in turn further reduced the size and cost.

7.3 Displacement Circulators

A displacement-type circulator is one where the polarization components are spatially separated and combined by birefringent crystals cut as walkoff blocks. Birefringent walkoff can yield an extinction ratio between the two polarization components in excess of 50 dB. Such an extinction ratio is better than most commercially available thin-film polarization beam splitting cubes.

Accordingly, the displacement-type circulators developed in the early 1990s exhibited superior performance in comparison with the earlier developed Matsumoto, Iwamura, and Shirasaki circulators.

The series of displacement-type circulators developed by Koga and Matsumoto highlight conceptual developments that have bearing on deflection-type circulators. Their first-reported design was a strict-sense four-port single-stage circulator. The complexity and limited performance of this design was overcome by their second design, which was a non-strict-sense ladder-type two-stage circulator. Even this design, however, was limited in part by the incorporation of reciprocal polarization rotators. These rotators add to the part count and reduce the isolation bandwidth. Their last-reported design was a non-strict sense ladder-type two-stage circulator using only nonreciprocal polarization rotation. This is the most compact and best performing of their devices.

Other inventors have proposed displacement-type circulators as well. In particular, Fujii proposed several designs [11–13] using both walkoff crystals and PBS cubes. Separately, Cheng developed a variety of reflection-based displacement circulators that were more compact than previous architectures [2, 4, 5]. He also addressed issues of low PMD and alignment improvements for manufacturing [3, 6]. More recently, Xie and Huang proposed a compact two-stage design that incorporates thermally expanded-core (TEC) fibers [42]. Use of TEC fibers reduces the necessary beam displacement and in turn the size of the component. TEC fibers will be seen again in the deflection circulators to follow. Finally, Liu *et al.* have proposed a strict-sense four-port two-stage circulator that includes a polarization beam-splitter cube to route the last port back to the first [27].

Figure 7.5 illustrates the first Koga and Matsumoto PI circulator [18, 21]. The core of their circulator is a variation of the Shibukawa PD circulator where the Glan-Taylor prisms are replaced with walkoff blocks (Fig. 7.5(a)). The center reciprocal rotator allows all four ports to lie in the same plane. This PD circulator was embedded between two polarization-conditioning sections that, all together, form a strict-sense circulator (Fig. 7.5(b)). The spot-trace diagrams at the bottom of the figure detail the connection between all port pairs.

As a critique, one can say that the polarization beam splitters on either side of the center FR each require four reciprocal parts. The extinction ratio of the splitters and the insertion loss of the device critically depend on the alignment of these elements. As the circulator is single-stage and the reciprocal parts will be misaligned in a real product, one cannot expect too much from this architecture.

A better design is illustrated in Fig. 7.6. This is a two-stage circulator having fewer components than the first design [18, 20, 21]. The simpler architecture was achieved by switching to a ladder-type device from a strict-sense device. Reciprocal rotators are still included, but the two-stage design helps with the isolation. As with the first device, all paths into and out of the cir-

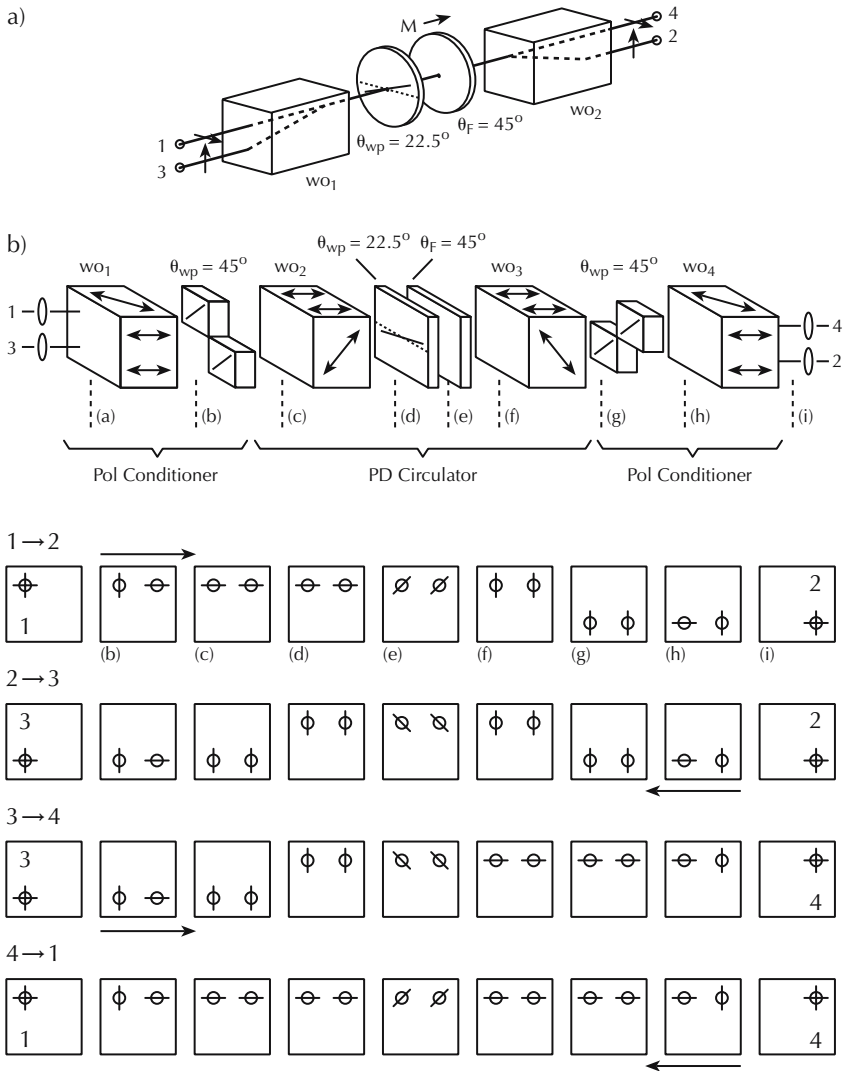


Fig. 7.5. Koga and Matsumoto strict-sense single-stage displacement-type circulator [18, 21], altered by the author to include half-wave waveplates rather than optically active plates and to simplify the quarter-size elements. a) A variation of the Shibukawa PD circulator, where walkoff blocks substitute for Glan-Taylor prisms and a half-wave waveplate allows all four ports to lie in the same plane. b) The PI circulator with embedded PD circulator. The polarization-conditioning stages generate the strict-sense PI circulatory behavior. All paths into and out of the circulator run parallel, and four separate collimators couple the light to and from fiber. At the bottom, spot-trace diagrams detail the connection between port pairs.

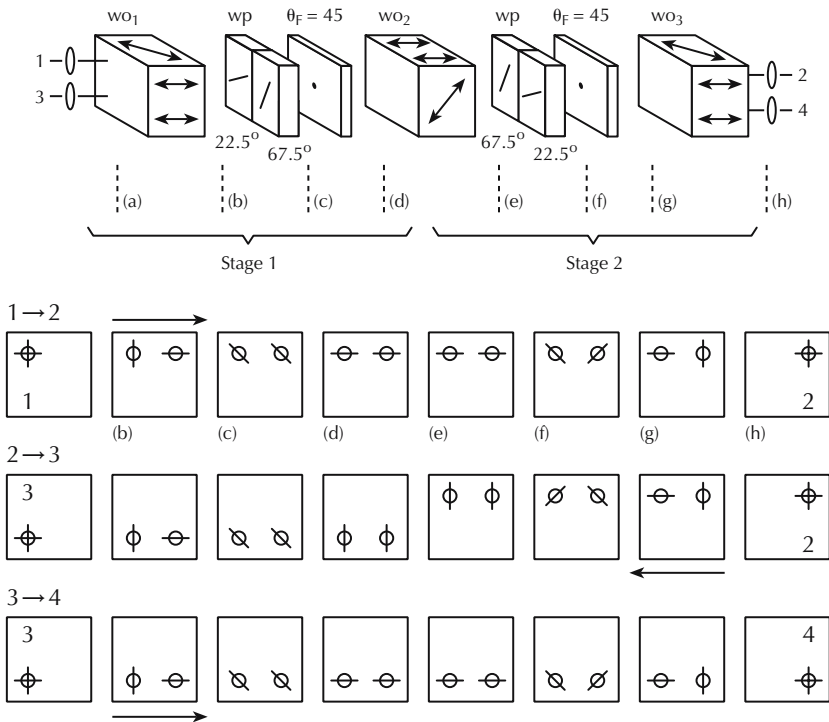


Fig. 7.6. Koga and Matsumoto ladder-type two-stage displacement circulator [18, 20, 21], altered by the author to include half-wave waveplates rather than optically active plates. The center walkoff block is the polarizing element between the first and second Faraday rotators, resulting in a two-stage circulator. All paths into and out of the circulator run parallel, and four separate collimators couple light to and from fiber. At the bottom, spot-trace diagrams detail the connection between port pairs. The connection 4 $\rightarrow$ 1 is not available.

culator are parallel, and separate collimating lenses were used to couple to and from the fiber. Given a minimum spacing of adjacent collimators based on the form factor, the displacement crystals must be long enough to couple to either lens. Koga and Matsumoto somewhat overcame this limitation by using turning prisms to deflect the light from a small core to a more widely spaced lens pair. However, use of turning prisms in production is not often attractive. Nonetheless, the inventors reported an insertion loss of < 1.5 dB, a PDL of 0.25 dB, and isolation at room temperature and center wavelength of over 67 dB. They reported a 70 nm wavelength range centered at 1550 nm where the isolation was at least 60 dB.

The final architecture in this series eliminates the reciprocal rotators all together [19], developed by Koga. The stated purposed by the inventor was to reduce the component length by removing the optically active rotators.

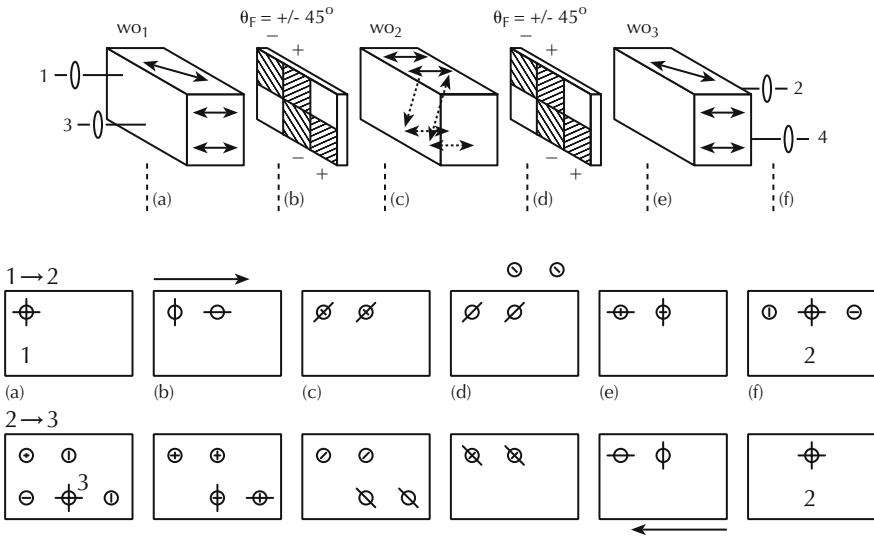


Fig. 7.7. A ladder-type two-stage displacement circulator with no reciprocal rotators, proposed by Koga [19]. Tiled FR elements are located between walkoff blocks. The center walkoff block walks the extraordinary rays along a 45° line in the plane perpendicular to the long component axis. All paths into and out of the circulator run parallel, and four separate collimators couple light to and from fiber. At the bottom, spot-trace diagrams detail the connection between port pairs as well as the error paths.

However, substantial length reduction could well have been achieved through substitution of half-wave waveplates. Nonetheless, the bandwidth of the resultant component is increased by the removal of the reciprocal rotators.

The quartz-free two-stage ladder-type circulator demonstrated by Koga is illustrated in Fig. 7.7. Here a checkerboard of FR elements is used to intersect the internal optical paths in the appropriate way to create a PI circulator. The center walkoff block is also changed to impart walkoff along a 45° direction in the plane perpendicular to the light path. Since the spacing in the FR checkerboard was small, only one external magnet could be used. To achieve both clockwise and counterclockwise polarization rotation the inventor used two different Faraday materials, YIG and a Bi:RIG derivative. The drawback was that the YIG garnet was 2.1 mm thick and the Bi:RIG garnet was 0.48 mm thick. This leads to a relatively high PDL (1.1 dB) and a differential group delay of 21 ps (although no more than about 8 ps can be accounted for via index and path-length difference alone). Also, while not reported, the temperature coefficients of these two materials differ, reducing the isolation over a normal operating range. Nonetheless, it was reported that the insertion loss was below 1.75 dB and the isolation better than 65 dB.

It should be noted that the YIG and Bi:RIG garnets can be replaced today with matched latching garnets. The $\pm 45^\circ$ rotations are realized by reversing the orientation of one part with respect to the other part. Also, the permanent magnet is removed. Using latching garnets, one expects the PDL, temperature dependence, and differential-group delay to improve substantially.

As a final note, like the earlier displacement circulators, the input and output ports are parallel to one another and individual collimators couple the light to fiber. The component size cannot be reduced beyond the displacement necessary for light to couple to either of two collimators on the same side of the component. As a consequence, there is a minimum volume of required crystal material which sets a floor on the price. These limitations are overcome by deflection circulators.

7.4 Deflection Circulators

The size of a circulator can be reduced by using a dual-fiber collimator. A dual-fiber collimator uses a single lens to collimate or focus the light from two closely-spaced fibers threaded through the same ferrule. As derived in §5.3, the consequence of using a single lens for two fibers is an angular divergence between collimated light paths. A typical full-angle divergence is 3° , although this varies by product.

Combination of dual-fiber collimators (DFC) with deflection prisms is natural. The simplest example of coupling a single-fiber collimator to a dual-fiber collimator is the polarization-beam splitter. Figure 7.8 illustrates four polarization-beam splitters, all of which use a birefringent prism for deflection. Figure 7.8(a) uses a Wollaston prism to convert parallel paths to paths that match the angular aperture of the DFC, and uses a walkoff crystal to laterally translate the beams into position [14]. Figure 7.8(b) combines the first two crystals into one, and when combined with a second prism, as shown, the combination is called the Kaifa prism (cf. §4.7.2). As with the first example, one beam is displaced before the two beams are deflected into the angular aperture of the DFC. Figure 7.8(c) achieves displacement by concatenating two Wollaston prisms with a “complete gap” in between [16]. The second Wollaston prism imparts stronger deflection than the first and orients the light into the angular aperture of the DFC. These three polarization-beam splitters all require translation of one or both beams because there is a presumption that large optical elements will be inserted between the collimators. However, very small-size designs may place the effective deflection plane of a prism right at the crossing point of the DFC [15, 25]. Specially designed DFCs can have crossing distances of ~ 2.4 mm. Figure 7.8(d) illustrates a Wollaston prism located as described to form a compact polarization-beam splitter.

As a critique of the first three designs, the third polarization-beam splitter uses less crystal volume than the first two. The deflection from the first prism makes the two beams diverge, creating displacement without the need for

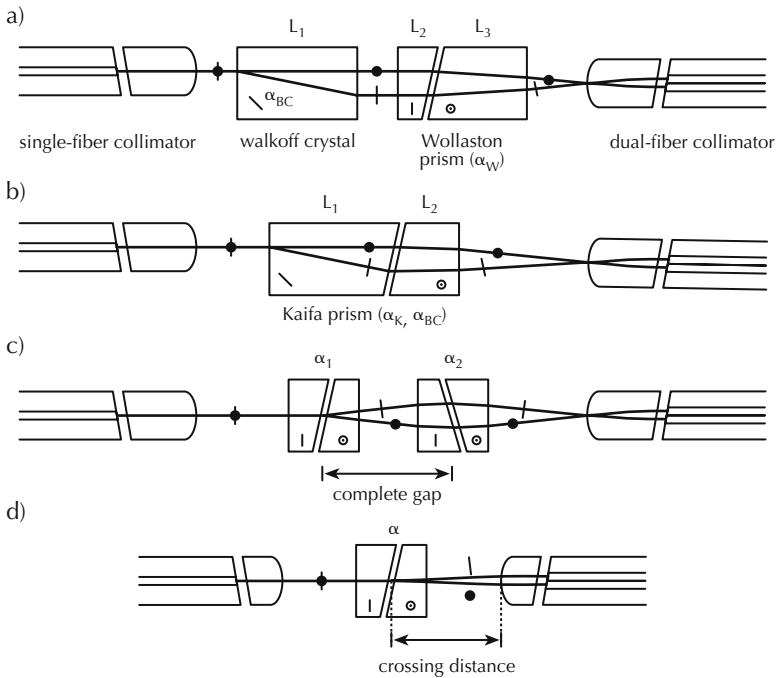


Fig. 7.8. Four polarization-beam splitters using deflection from compound birefringent prisms to match the angular aperture of an output dual-fiber collimator. a) Combined walkoff crystal and Wollaston prism. The prism imparts deflection into the angular aperture of the DFC, while the walkoff crystal provides the necessary translation. b) The Kaifa compound prism combines walkoff and deflection in one compound prism. c) A pair of Wollaston prisms with an intermediate complete gap. The second prism imparts more deflection than the first. The complete gap is adjusted to fine-tune the spatial translation for optimal coupling. d) Single Wollaston prism placed at the crossing point of a DFC. Such a system typically has a small gap between lenses.

crystal. Also, the complete gap is a convenient alignment degree-of-freedom unavailable in the first two designs. The forth splitter uses the least crystal volume of all but requires that all necessary parts fit between the ferrule and the crossing point.

Based on this primer, five significant circulator architectures are detailed in the following. These circulators are all deflection-type, non-strict-sense three or four port components that were independently developed between 1997 and 1999. The first two circulators are the Kaifa designs, invented by Li, Au-Yeung, Guo, and Wang. The third and last circulators are New Focus designs invented by Xie and Huang. The fourth circulator is a Fujitsu-Avanex design invented by Shirasaki and Cao. These designs highlight the aforementioned techniques to integrate dual-fiber collimators into two-stage circulators.

Kaifa Circulators

Figure 7.9 illustrates a Kaifa non-strict-sense two-stage three-port circulator that uses a deflection scheme [23, 24]. The chain of optical elements is shown in Fig. 7.9(a). Polarization diversity is carried out in the vertical plane via walkoff crystals wo_1 and wo_3 , while deflection happens in the horizontal plane via the Wollaston prism and walkoff crystal wo_2 . The separation of polarization diversity and deflection into orthogonal planes is important because fine placement adjustments can be made independently of one another. The circulator is two-stage because each FR is located between a walkoff-crystal polarizer or compound prism polarizer. The nonreciprocal FR plates are each preceded by a reciprocal half-wave waveplate pair. The pair is split along the horizontal and oriented so that, at the center wavelength of the waveplates, one plate rotates the polarization state by $+45^\circ$ while the other imparts a rotation of -45° . Figures 7.9(b,c) show the spot-trace diagrams, and the top- and side-view of the component for connections $1 \rightarrow 2$ and $2 \rightarrow 3$.

As a critique, the reciprocal elements add to the part count, may be difficult to precisely align, and reduce the isolation bandwidth. Moreover, the second walkoff-block is an unnecessary addition to the part count.

These two problems are remedied in the second Kaifa design, shown in Fig. 7.10 [1]. Like the first, this circulator is a non-strict-sense two-stage three-port circulator, but here the Kaifa compound prism is incorporated to impart displacement and deflection. Moreover, the reciprocal half-wave waveplates are eliminated. To accommodate the elimination, the Faraday rotator plates are each split in two, the upper having a rotary direction clockwise and the lower having an opposite rotary direction; and the first and third walkoff crystals are re-cut to displace light along a 45° angle (rather than vertical) in the plane perpendicular to the longitudinal axis of the component. This change in crystal cut rotates by 45° the output polarizations of the displaced and straight-through beams. The polarization states impinging on the Kaifa prism in either direction remain that same as those in Fig. 7.9.

The result of the re-cut of the first and third walkoff blocks is a coupling of the horizontal and vertical axes, which makes alignment more difficult. However, the component has three fewer parts and no reciprocal rotators.

A First Xie-Huang Circulator

A first Xie-Huang deflection-type circulator is illustrated in Fig. 7.11 [40, 41, 43]. This circulator is in the same spirit as the Kaifa circulators, and is a non-strict-sense two-stage three-port design. While the first Kaifa circulator uses waveplates and the second circulator eliminates them but couples the horizontal and vertical axes, the present Xie-Huang circulator combines the best features of the two. At the core of the Xie-Huang circulator is a pair of Wollaston prisms separated by a complete gap. As with the aforementioned polarization-beam splitters, one prism deflects more strongly than the other.

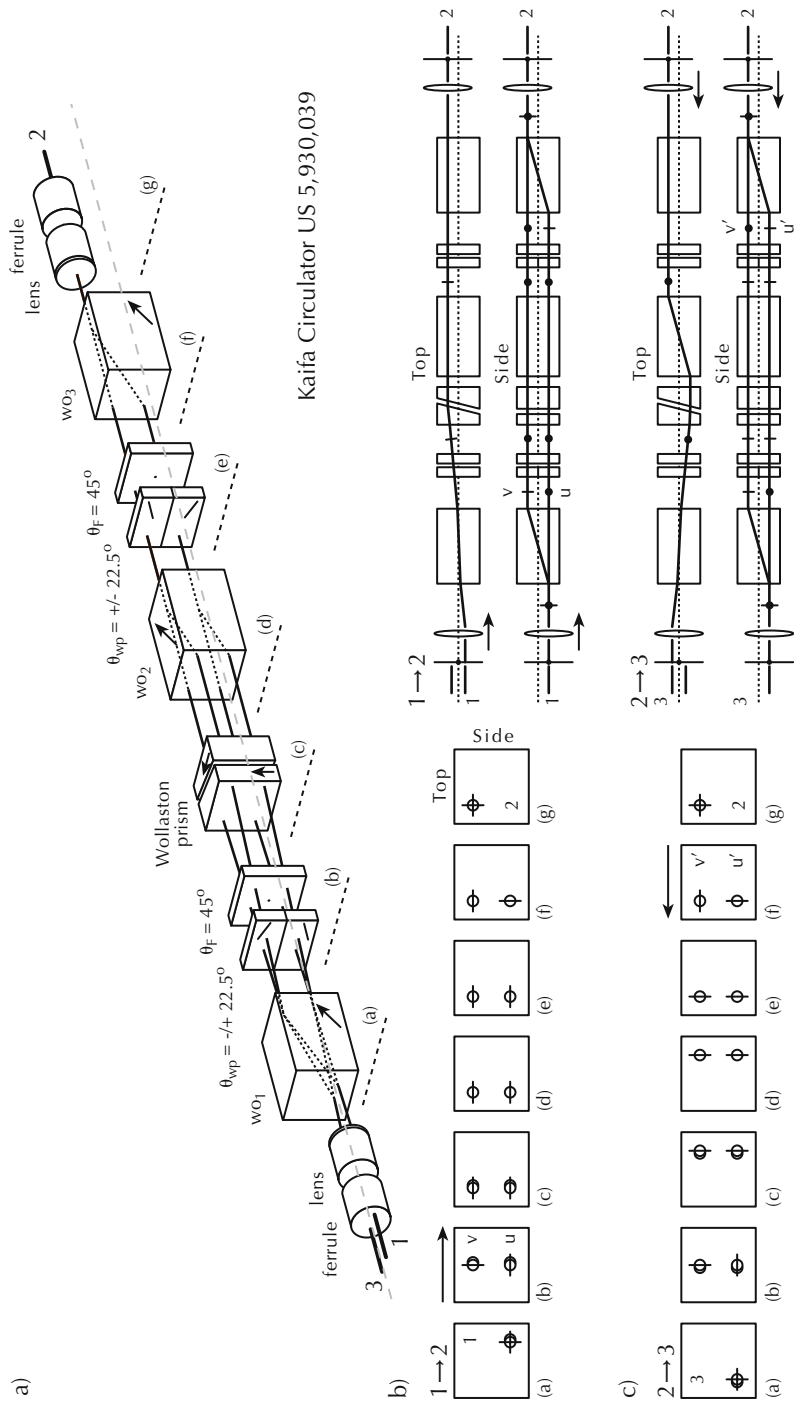


Fig. 7.9. Kaifa non-strict-sense two-stage three-port circulator. Deflection via Wollaston prism and walkoff crystal.

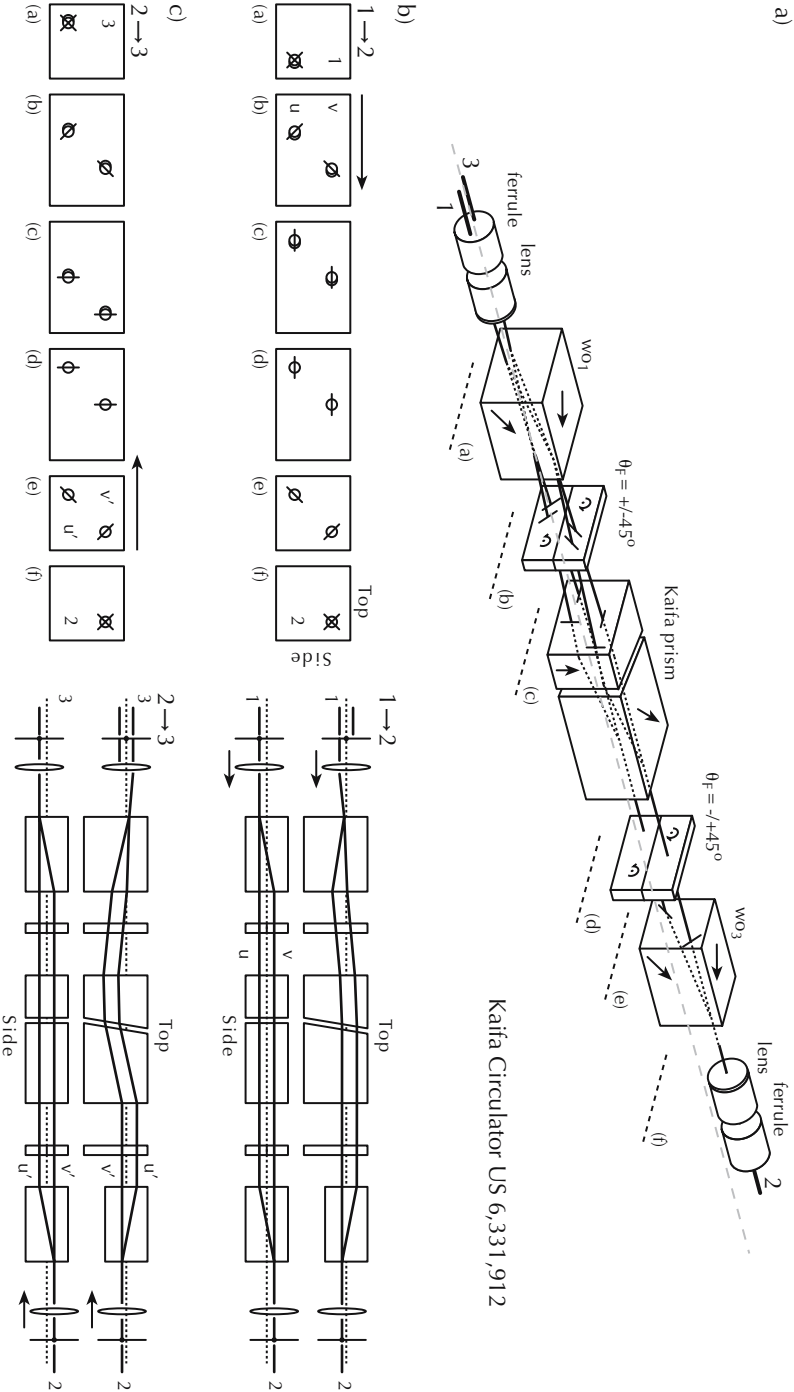


Fig. 7.10. Kaifa non-strict-sense two-stage three-port circulator. Deflection via Kaifa prism.

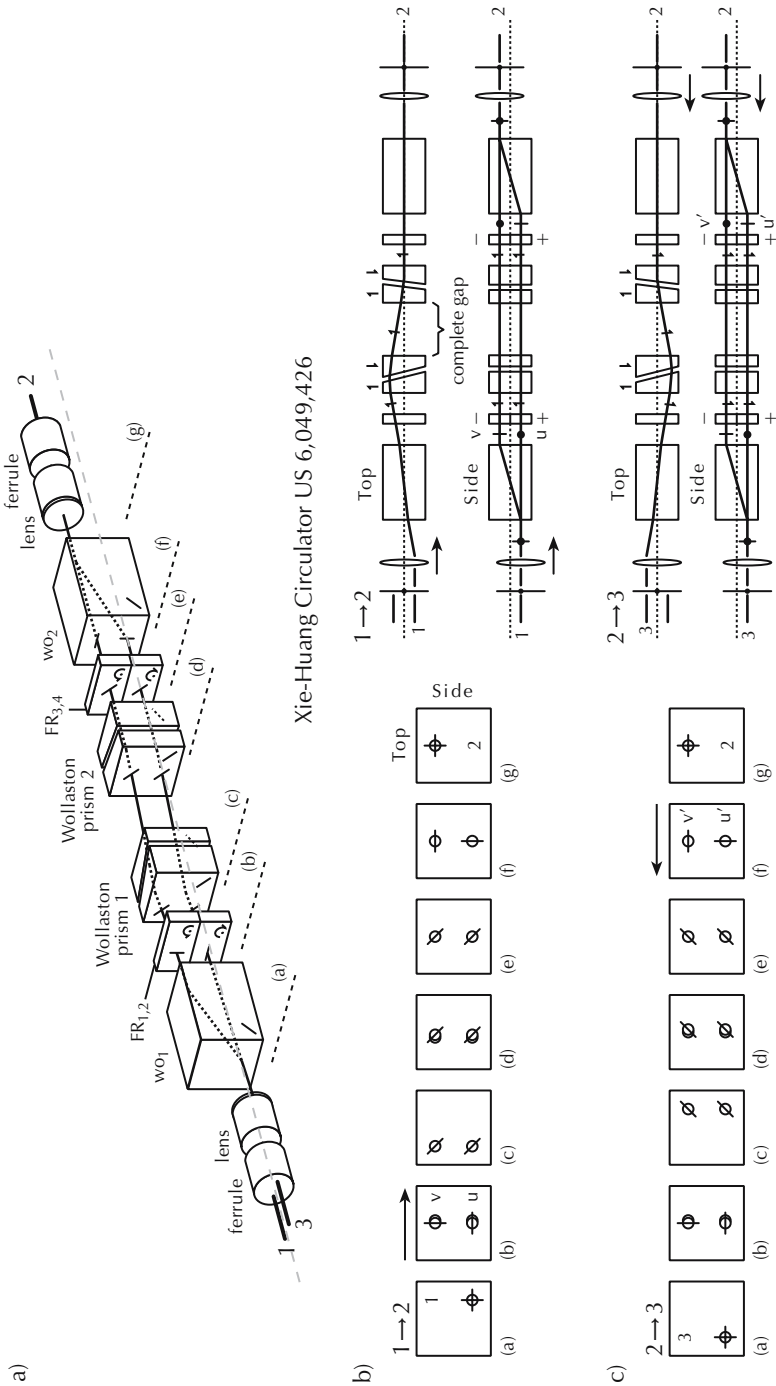


Fig. 7.11. Xie-Huang non-strict-sense two-stage three-port circulator. Deflection via Wollaston prism pair and complete gap.

The combination of the weaker prism and the complete gap together generate the displacement embodied by the Kaifa designs. However, since the orientation of the birefringent axes in the Wollaston prisms is arbitrary, the Xie-Huang design uses a modified Wollaston prism pair with birefringent axis orientations of $\pm 45^\circ$. These axes are directly aligned to the polarization states resolved by the walkoff crystals and the Faraday rotator pairs. Accordingly, the polarization diversity generated by walkoff crystals w_{o1} and w_{o2} may be in the plane perpendicular to the deflection of the Wollaston prism pair. Moreover, the Wollaston prism pair allows for simultaneous elimination of differential-group delay and balancing of diffraction along the two paths. These properties lead to low PMD and low PDL, respectively.

The complete gap is an ingenious feature that allows fine-tuned alignment between collimators. The convergence angle may be just a few degrees. For instance, a 3° convergence angle gives a 20 : 1 ratio between longitudinal adjustment and lateral displacement.

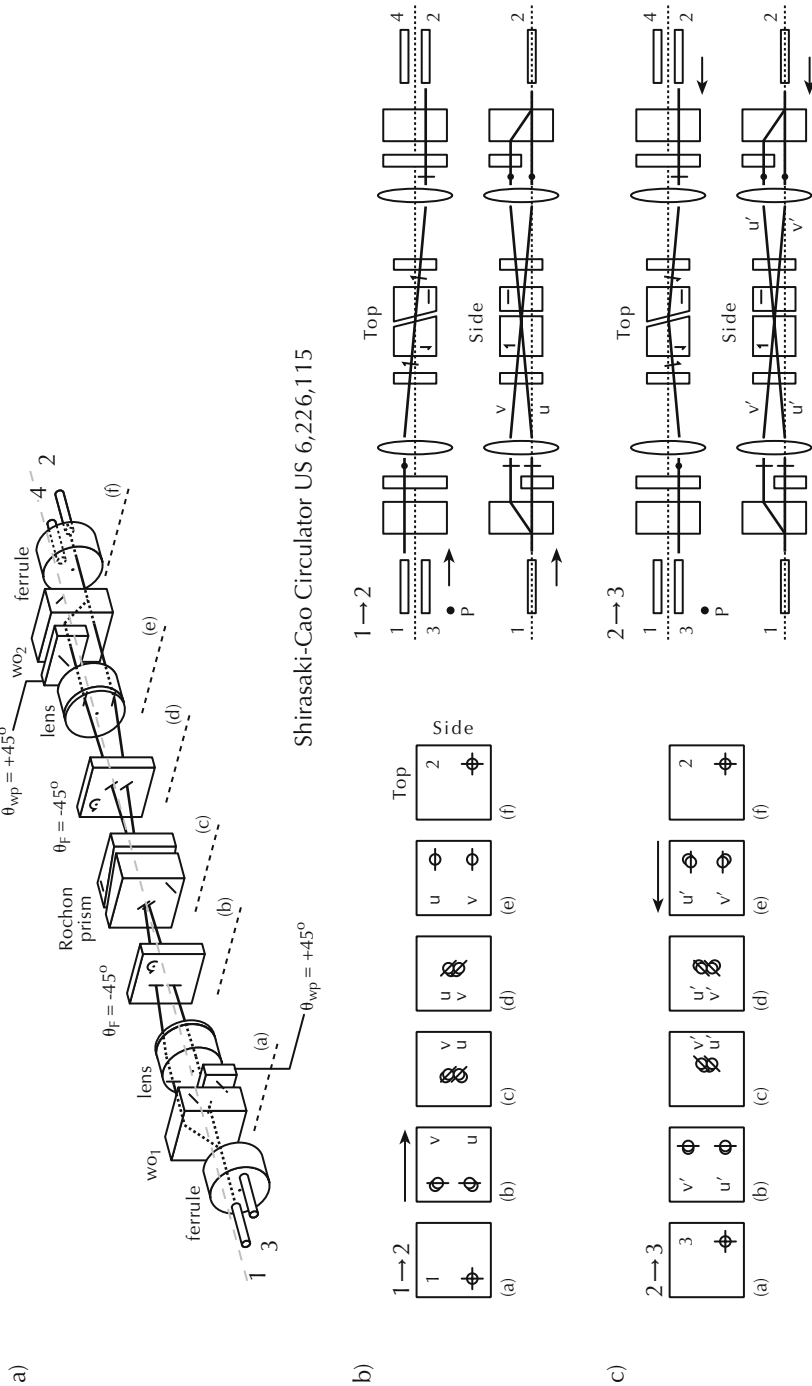
The inventors use latching iron garnet Faraday rotators to eliminate the permanent magnets and to allow the same material to be used for both rotary directions in the pair. This is an important innovation that reduces size and balances paths, and is possible because the two-stage isolation accommodates the increased temperature sensitivity of the latching garnet.

Shirasaki-Cao Circulator

To make a circulator smaller yet, the displacement for the polarization diversity must be reduced. The displacement, of course, must be sufficient to separate light into two distinct beams. Shirasaki and Cao impart displacement at the ferrule and before the lens, where the beam waist is on the order of $10\ \mu\text{m}$ (separately noted in [26]). The displacement crystal need only be $\sim 100\ \mu\text{m}$ thick, a factor of $25\times$ reduction from a typical crystal length where the walkoff follows the lens. With the removal of walkoff crystals from the optical path between the lens pair, there is room to place the deflection prism at the crossing point of the dual-fiber collimator. Architectures of this type are the most compact of all circulator.

Figure 7.12 illustrates the Shirasaki-Cao non-strict-sense two-stage four-port circulator [35], an improvement on an earlier design by Shirasaki [34]. As a four-port scheme, dual-fiber collimators are located on either end of the component. While any deflection prism may be used, the illustration shows a Rochon prism. As shown in Fig. 7.12(a), the walkoff and deflection directions are orthogonal, which decouples the adjustment of polarization diversity from isolation. Also note that in contrast to the preceding deflection-type circulators the present design uses a cross-over design in the plane of polarization-diversity. The lens axis is located between the axes of the walkoff light and straight-through light so as to deflect both paths equally.

In this architecture, a walkoff crystal and half-wave waveplate are located between the ferrule and lens (Fig. 7.12(a)). Accordingly, the polarizations on



Shirasaki-Cao Circulator US 6,226,115

Fig. 7.12. Shirasaki-Cao non-strict-sense two-stage four-port circulator. Deflection via single Wollaston prism.

the two paths at the lens are nominally the same. After the lens, the polarization state is rotated by 90° by transit of the first and second Faraday rotators. The deflecting prism located between the two FRs is cut so that the birefringent axes are at $\pm 45^\circ$. Using a Rochon prism, one polarization state (i.e. $+45^\circ$) is transmitted straight through while the orthogonal state is deflected. As shown by the ray-trace diagrams in Fig. 7.12(b,c), the component can be designed so that no deflection makes the connections $1 \rightarrow 2$ and $3 \rightarrow 4$, while deflection makes the connection $2 \rightarrow 3$. Light input to port 4 is deflected and lost toward point P. As discussed by the inventors, the diffraction of the u -path is less than the v -path because the former path transits the two waveplates. Moreover, the differential diffraction occurs in the high N.A. region between ferrule and lens. In principle this leads to either higher insertion loss or higher PDL. If a particular design cannot be made satisfactorily, a glass plate can be inserted beside the waveplate to balance the diffraction.

As a critique, the presence of the waveplates limits the bandwidth of the device. Moreover, the Rochon prism should be as thin as possible to minimize the imparted differential-group delay. Otherwise, the Shirasaki-Cao design is very compact and overall rather simple to make.

A Second Xie-Huang Circulator

Many variations are possible with the Shirasaki-Cao architecture, some of which depend on the particular technologies that are employed. An independently invented Xie-Huang circulator that falls into this class is illustrated in Fig. 7.13 [44]. Here the Faraday rotator plates are brought between the ferrule and lens. The plates are split in two along a horizontal line and the plates impart opposite rotations. This allows the removal of the half-wave waveplates present in the former scheme. This step is technically difficult for the following reasons. The FRs must be latching garnets to enable opposite rotation with the same material. A latching garnet thickness for $1.55 \mu\text{m}$ applications is $\sim 500 \mu\text{m}$ thick instead of $\sim 350 \mu\text{m}$ thick for non-latching Bi:RIG garnets. The latching garnet thickness is to be compared to the half-wave waveplate thickness of $92 \mu\text{m}$.

The five-fold thickness increase, especially between ferrule and lens, requires more displacement and a larger lens. Alternatively, the inventors use TEC fiber to reduce the numerical aperture of the light emergent from the ferrule. It is in this way the split-FR highly compact design can be realized.

Figure 7.13(a) illustrates their non-strict-sense two-stage three-port circulator. Since only three ports are used (and accordingly only one DFC), the deflecting prism is naturally a Wollaston prism. When viewed from port 2, one polarization is deflected toward port 1 and the other towards port 3. The ray-trace diagrams in Fig. 7.13(b,c) show the $1 \rightarrow 2$ and $2 \rightarrow 3$ connections. Light input to port 3 is deflected to the region above port 2 and is lost. Since this circulator is diffraction, path-length, and temperature balanced between

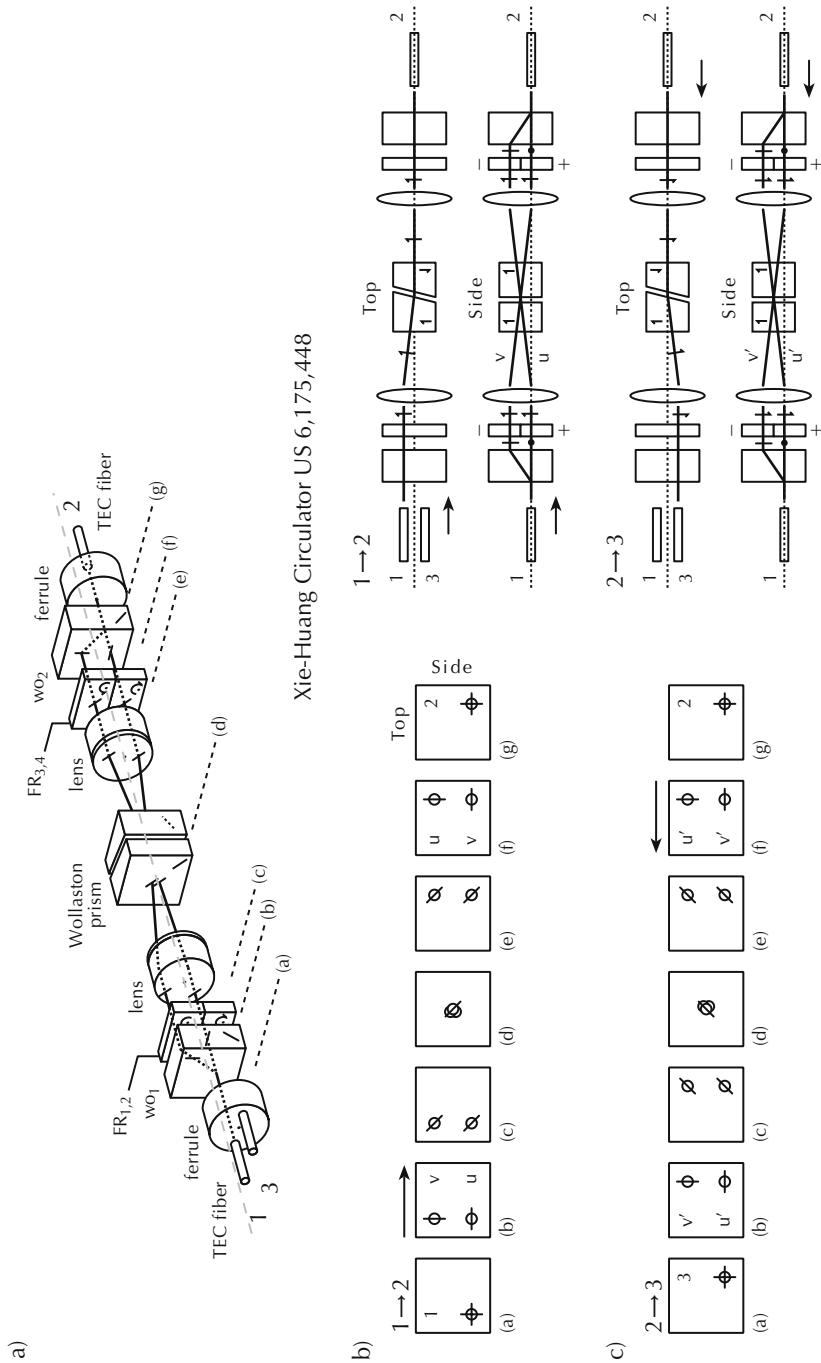


Fig. 7.13. Xie-Huang non-strict-sense two-stage three-port circulator. Deflection via single Wollaston prism.

Table 7.1. Two-Stage Deflection-Type Circulator Performance

Specification ^(a)	Nominal	Units
Center wavelength	1550	nm
Bandwidth	50	nm
Isolation (λ_c , 23°C, all SOP)	50	dB
Isolation ($\lambda_c \pm 25$ nm, 0-65°C, all SOP)	40	dB
Insertion Loss (λ_c , 23°C, all SOP)	0.5	dB
Insertion Loss ($\lambda_c \pm 25$ nm, 0-70°C, all SOP)	0.7	dB
PMD	0.05	ps
PDL	< 0.15	dB
Return loss	50	dB
Maximum crosstalk	−50	dB
Power handling	500	mW
Operating temperature	0–65	°C

^(a)Specification values reported by New Focus [30].

the two paths, one can expect very good performance. Table 7.1 lists the specifications for a high-performance compact circulator such as this one.

7.5 Summary

Circulator architectures have experienced periods of rapid and competitive development followed by relatively stable design concepts. The two driving factors behind periods of development are the advent of new materials and sub-assemblies, such as iron garnets and dual-fiber collimators, as well as application-specific demands, such as compact form factor and low polarization-dependent loss. This chapter has selected from the broader patent and technical literature those circulators that highlight the conceptual evolution. Within the framework developed here other circulators may be categorized and then one can make a prediction on its relative performance.

At the time of this writing initiatives to create new functionality are being developed. The bi-isolator [39] and the bi-circulator [8, 37, 38] are two leading examples. Both components are designed to work on a uniformly spaced wavelength-division multiplexed grid in order to facilitate bi-directional communication. Consider the separation of DWDM wavelengths into even and odd channels, and run the even channels “east” and the odd channels “west.” On even channels, the bi-isolator prevents light from travelling west; on odd

channels the bi-isolator prevents light from travelling east. The bi-circulator operates as a similar principle in that the component circulates “clockwise” for even channels and “counterclockwise” for odd channels. The bi-circulator can act as a gateway between unidirectional and bidirectional communication systems. The element common to both the bi-isolator and bi-circulator is an interleaving filter. The filter designs are discussed in the references.

References

1. V. Au-Yeung, Q.-D. Gao, and X. L. Wang, “Optical circulator,” U.S. Patent 6,331,912, Dec. 18, 2001.
2. Y. Cheng, “Reflective optical non-reciprocal devices,” U.S. Patent 5,471,340, Nov. 28, 1995.
3. —, “Optical circulator,” U.S. Patent 5,574,596, Nov. 12, 1996.
4. —, “Optical circulator,” U.S. Patent 5,878,176, Mar. 2, 1999.
5. —, “Optical circulator,” U.S. Patent 5,930,422, June 27, 1999.
6. —, “Optical circulator,” U.S. Patent 5,991,076, Nov. 23, 1999.
7. J. S. V. Delden, “Optical circulator having a simplified construction,” U.S. Patent 5,212,586, May 18, 1993.
8. T. Ducellier, K. Tai, K.-W. Chang, J. Chen, and Y. Cheng, “Bi-directional circulator,” U.S. Patent 2002/00 224 730, Feb. 28, 2002.
9. W. L. Emkey, “A polarization-independent optical circulator for 1.3 microns,” *Journal of Lightwave Technology*, vol. LT-1, no. 3, pp. 466–469, Sept. 1983.
10. —, “Optical circulator,” U.S. Patent 4,464,022, Aug. 7, 1984.
11. Y. Fujii, “High-isolation polarization-insensitive optical circulator,” *Journal of Lightwave Technology*, vol. 9, no. 10, pp. 1238–1243, Oct. 1991.
12. —, “High-isolation polarization-insensitive optical circulator coupled with single-mode fiber,” *Journal of Lightwave Technology*, vol. 9, no. 4, pp. 456–460, Apr. 1991.
13. —, “High-isolation polarization-insensitive quasi-optical circulator,” *Journal of Lightwave Technology*, vol. 10, no. 9, pp. 1226–1229, Sept. 1992.
14. Y. Huang and P. Xie, “Optical polarization beam combiner/splitter,” U.S. Patent 6,331,913, Dec. 18, 2001.
15. —, “Optical polarization beam combiner/splitter,” U.S. Patent 6,282,025, Aug. 28, 2001.
16. —, “Optical polarization beam combiner/splitter,” U.S. Patent 6,373,631, Apr. 16, 2002.
17. H. Iwamura, H. Iwasaki, K. Kubodera, Y. Torii, and J. Noda, “Compact optical circulator for near-infrared region,” *Electronics Letters*, vol. 15, no. 25, pp. 830–831, Dec. 1979.
18. M. Koga, “Optical circulator,” U.S. Patent 5,204,771, Apr. 20, 1993.
19. —, “Compact quartzless optical quasi-circulator,” *Electronics Letters*, vol. 30, no. 17, pp. 1438–1440, Aug. 1994.
20. M. Koga and T. Matsumoto, “Polarisation-insensitive high-isolation nonreciprocal device for optical circulator application,” *Electronics Letters*, vol. 27, no. 11, pp. 903–905, May 1991.

21. —, “High-isolation polarization-insensitive optical circulator for advanced optical communication systems,” *Journal of Lightwave Technology*, vol. 10, no. 9, pp. 1210–1217, Sept. 1992.
22. H. Kuwahara, “Optical circulator,” U.S. Patent 4,650,289, Mar. 17, 1987.
23. W.-Z. Li, V. Au-Yeung, and Q.-D. Gao, “Optical circulator,” U.S. Patent 5,909,310, June 1, 1999.
24. —, “Optical circulator,” U.S. Patent 5,930,039, July 27, 1999.
25. W. Z. Li and Y. Yang, “Method and system for splitting or combining optical signal,” U.S. Patent 6,353,691, Mar. 5, 2002.
26. W. Z. Li, Y. Yang, F. Liu, and W. Luo, “Polarization splitter and combiner and optical devices using the same,” U.S. Patent 6,493,140, Dec. 10, 2002.
27. Z. Liu, M. S. Wang, and J. Xu, “Loop optical circulator,” U.S. Patent 2003/0007244 A1, Jan. 9, 2003.
28. T. Matsumoto, “Optical circulator,” U.S. Patent 4,272,159, June 9, 1981.
29. T. Matsumoto and K. Sato, “Polarization-independent optical circulator: An experiment,” *Applied Optics*, vol. 19, no. 1, pp. 108–112, Jan. 1980.
30. “New focus optical circulators (c-band),” New Focus Corporation, San Jose, California, 2003. [Online]. Available: http://www.newfocus.com/Online_Catalog/literature/Cband.pdf
31. W. B. Ribbens, “An optical circulator,” *Applied Optics*, vol. 4, pp. 1037–1038, 1965.
32. A. Shibukawa and M. Kobayashi, “Compact optical circulator for near-infrared region,” *Electronics Letters*, vol. 14, no. 25, pp. 816–817, Dec. 1978.
33. —, “Compact optical circulator for optical fiber transmission,” *Applied Optics*, vol. 18, no. 21, pp. 3700–3703, Nov. 1979.
34. M. Shirasaki, “Optical device,” U.S. Patent 5,982,539, Nov. 9, 1999.
35. M. Shirasaki and S. Cao, “Optical circulator or switch having a birefringent wedge positioned between faraday rotators,” U.S. Patent 6,226,115, May 1, 2001.
36. M. Shirasaki, H. Kuwahara, and T. Obokata, “Compact polarization-independent optical circulator,” *Applied Optics*, vol. 20, no. 15, pp. 2683–2687, Aug. 1981.
37. K. Tai, Q. Guo, K. W. Chang, J. Chen, and M. Xu, “4-port interleavers and fully circulating bi-directional circulators,” in *Tech. Dig., Optical Fiber Communications Conference (OFC’01)*, Anaheim, CA, Mar. 2001, paper MK5.
38. K. Tai, K.-W. Chang, J. Chen, T. Ducellier, and Y. Cheng, “Wavelength-interleaving bidirectional circulators,” *IEEE Photonics Technology Letters*, vol. 13, no. 4, pp. 320–322, 2001.
39. —, “Bi-directional isolator,” U.S. Patent 6,587,266, July 1, 2003.
40. P. Xie and Y. Huang, “Compact polarization insensitive circulators with simplified structure and low polarization mode dispersion,” U.S. Patent 6,049,426, Apr. 11, 2000.
41. —, “Compact polarization insensitive circulators with simplified structure and low polarization mode dispersion,” U.S. Patent 6,052,228, Apr. 18, 2000.
42. —, “Compact polarization insensitive circulators with simplified structure and low polarization mode dispersion,” U.S. Patent 6,212,008, Apr. 3, 2001.
43. —, “Compact polarization insensitive circulators with simplified structure and low polarization mode dispersion,” U.S. Patent 6,285,499, Sept. 4, 2001.
44. —, “Optical circulators using beam angle tuners,” U.S. Patent 6,175,448, Jan. 16, 2001.

Properties of Polarization-Dependent Loss and Polarization-Mode Dispersion

Polarization-dependent loss (PDL) and polarization-mode dispersion (PMD) are two linear properties that are encountered when dealing with long spans of single-mode fiber. PDL and PMD are not confined to fiber, however, and may be present in components, simple optical elements such as birefringent crystals and polarizers, and instruments that generate these effects. Yet the study of PDL and PMD and their concatenation properties was motivated by research conducted over the last 30 years on fiber-optic communication links. Indeed, polarization-dependent effects in fiber optics have been studied since the 1970s [37, 55].

Polarization-dependent loss refers to energy loss that is preferential to one polarization state. In the Jones picture, one axis suffers more loss than the other. This differential loss changes the output polarization state and imparts a common loss to an unpolarized light beam. The complement of polarization-dependent loss is polarization-dependent gain (PDG). PDG is an effect that is related to preferential gain between signal and noise in optical amplifiers, which, if left unaddressed, can cause impairments of its own sort. The specifics of PDL detailed in this work are equally applicable to PDG, but impairment analysis of the latter must include the combined effects of noise and signal. This is beyond the scope of the present text and the Reader is referred to Desurvire [7].

Polarization-mode dispersion refers to the polarization effects of concatenated lossless birefringent segments. Each homogeneous segment produces differential-group delay. PMD is generated when two or more differential-group delay segments are placed in cascade. While PMD can have the properties of a single homogeneous birefringent element (when the birefringent axes are aligned along a multi-segment cascade), a single homogeneous birefringent element does not possess all of the properties of PMD. PMD is not defined according to an eigenvector analysis of the transformation matrix as one is accustomed to for a single birefringent element.

The combination of PDL and PMD creates effects which are quite complicated and which can impair a communication system more than either effect

alone. For example, PDL is generally wavelength independent, while PMD is generally wavelength dependent. Addition of some PMD to PDL will generally result in wavelength-dependent PDL. The addition of some PDL to PMD can in some cases result in differential-group delay that exceeds what one would expect from the PMD alone.

In terms of partial polarization, PDL and PMD have opposite effects. PDL tends to polarize a partially polarized signal since it is a partial polarizer; PMD, however, tends to depolarize a signal since PMD creates polarization-state dispersion in frequency. Lyot depolarizers exploit polarization-state dispersion to (pseudo-)depolarize a transmission signal (cf. §1.5.3), but since this is tantamount to adding PMD to the system such one- and two-stage depolarizers, studied extensively in the late 1980s, have fallen out of favor.

The Mueller matrix for PDL and PMD highlights some of the properties one can expect from these two effects. PDL, being the effect of a partial polarizer, is represented by a Jones matrix that is Hermitian. When converted to a Mueller matrix $\mathbf{M}_{\text{PDL}}$ in Stokes space, all 16 entries may be occupied:

$$H \mapsto \mathbf{M}_{\text{PDL}} = \begin{pmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{pmatrix}, \text{ and } \underline{H} \mapsto \mathbf{M}_{\text{PMD}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix}$$

Given this form, PDL is not reversible unless gain is added. PMD, being a physical quantity that is measurable, is also represented by a Hermitian Jones matrix. Unlike PDL, however, the PMD matrix is identically traceless. One can argue this simply based on the lossless cascade of retarders that generates PMD. Accordingly, its Mueller matrix $\mathbf{M}_{\text{PMD}}$ only contains entries in the lower sub-matrix. The effect of PMD, therefore, can in principle be reversed. Finally, the addition of PDL anywhere along a birefringent cascade scatters the PMD Mueller-matrix entries into all 16 sites. The combination cannot therefore be strictly inverted.

The following sections of this chapter define the PDL and PMD vectors, show their effects on input states of polarization, derive equations of motion for concatenated vectors, and illustrate how these effects impact a waveform. The following chapter details the statistical properties of these effects.

8.1 Polarization-Dependent Loss

There are many mechanisms that generate polarization-dependent loss along a fiber-optic link. Micro-optic components such as those studied in the preceding chapters generate PDL when there is preferential coupling through a focusing lens of one polarization-diversity path over another. Integrated optic filters generate PDL because the passband frequency locations are polarization dependent. Fused fiber couplers generate PDL due to polarization-dependent

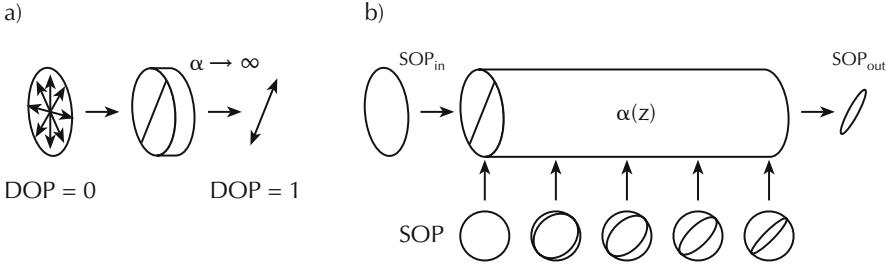


Fig. 8.1. Effects of PDL. a) A completely depolarized signal is completely polarized after transmission through a perfect polarizer. b) A partial PDL element continuously changes the polarization state of input light and diminishes the overall intensity.

coupling ratios. Optical amplifiers generate PDG due to preferential gain of the signal polarization over the orthogonal noise polarization [59]. Finally, micro-bends in fiber will generate PDL. Materials themselves can generate PDL because of dichroism of the molecules, such as in polymer waveguides. Regardless of the origin, however, a single formalism describes the effects.

8.1.1 Definitions

Table 8.1 gives a symbol list for the terms used in the following analysis. There is some inconsistency of notation in the literature, but the notation adopted here is intended to include the broadest overlap.

Figure 8.1 illustrates two effects of PDL. When PDL is complete, the element acts as a perfect polarizer (Fig. 8.1(a)). A perfect polarizer transmits only states that have a finite projection along the polarizer axis. After polarization, a completely depolarized signal becomes completely polarized. The PDL studied below is for partial polarization, not complete. Moreover, polarization occurs continuously through a differentially lossy medium. Figure 8.1(b) illustrates the evolution of light that is initially circularly polarized along a homogeneous PDL element. As the light travels the intensity along the loss axis diminishes while that along the neutral axis is unaltered. One can see how PDL transforms the polarization state along the element. Light launched along the PDL axis is unaltered and undiminished, and light launched along the orthogonal axis is unaltered in state but suffers loss.

Polarization-dependent loss ρ_{dB} is defined by international standards bodies such as the TIA and IEC as [34, 35, 60]

$$\rho_{\text{dB}} \equiv 10 \log_{10} \left(\frac{T_{\text{max}}}{T_{\text{min}}} \right) \quad (8.1.1)$$

where T_{max} and T_{min} are the maximum and minimum transmission intensities through the system. PDL is defined in decibels and is positive. Maximum

transmission intensity occurs when the polarization state of a completely polarized probe beam is aligned to the maximum transmission axis of the PDL element. That axis may be anywhere on the Poincaré sphere. Minimum transmission occurs for the orthogonal polarization (even in the presence of PMD).

Consider an intuitive example. The Jones matrix of a PDL element aligned to the horizontal (S_1 in Stokes space) may be written as

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & \\ & e^{-\alpha} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (8.1.2)$$

where α is the loss coefficient. When the input polarization is $(1, 0)^T$ the output intensity is $|v_1|^2 = 1$. Similarly, for $(0, 1)^T$ the output intensity is $|v_2|^2 = \exp(-2\alpha)$. Therefore the PDL is

$$\rho_{\text{dB}} = \alpha(20 \log_{10} e) \quad (8.1.3)$$

This relation holds true for any orientation of the PDL vector. Note that $20 \log_{10} e \simeq 8.686$. Moreover, for $\rho_{\text{dB}} = 3$ dB, $\alpha \simeq 0.345$.

In a fiber-optic link a PDL element is generally located somewhere between the terminations. As single-mode fiber does not preserve polarization in a practical link, the apparent orientation of the PDL vector at the fiber termination generally will not give a purely diagonal Jones matrix. Instead, the PDL vector can point in any direction. In particular the output matrix is $P' = UPV^\dagger$, where U, V are unitary operators.

The generalization of the PDL operator P , whose simple case is that in (8.1.2), comes in the matrix exponential form:

$$P = e^{-\alpha/2} \exp\left(\frac{\vec{\alpha} \cdot \vec{\sigma}}{2}\right) \quad (8.1.4)$$

where local PDL vector $\vec{\alpha} = \alpha \hat{\alpha}$ and $\hat{\alpha}$ is a unit vector in Stokes space that points in the direction of maximum transmission. This matrix exponential operator is expanded using (2.5.77) on page 62, yielding

$$P = e^{-\alpha/2} (I \cosh(\alpha/2) + (\hat{\alpha} \cdot \vec{\sigma}) \sinh(\alpha/2)) \quad (8.1.5)$$

Indeed (8.1.2) is recovered with $\hat{\alpha} = \hat{s}_1$.

An input state of polarization is altered by the PDL element according to

$$|t\rangle = P |s\rangle$$

Assuming that $\langle s | s \rangle = 1$, the transmission is

$$\langle t | t \rangle = \langle s | P^\dagger P | s \rangle$$

As the exponent of P is real, one has $P^\dagger P = P^2$. Therefore,

Table 8.1. Symbol Definitions for PDL

Symbol	Range	Definition
ρ_{dB} :	$0 \leq \rho_{\text{dB}} < \infty$	Polarization-dependent loss in decibels (dB) as defined by the TIA and IEC
$\vec{\alpha}, \hat{\alpha}$:		PDL vector and unit vector in Stokes space. $\vec{\alpha}$ points in the direction of maximum transmission
α :	$0 \leq \alpha < \infty$	Loss coefficient $ \vec{\alpha} $, related to ρ_{dB} via: $\rho_{\text{dB}} = \alpha(20 \log_{10} e)$
Γ :	$0 \leq \Gamma \leq 1$	Unit loss; normalization of α via: $\Gamma = \tanh \alpha$
$\vec{\Gamma}$:		Vector of cumulative PDL over a concatenation. $\vec{\Gamma}$ points in the direction of maximum transmission
$\vec{\alpha}(z)$:		PDL vector per unit length
$\vec{\Gamma}(z)$:		Cumulative PDL vector per unit length

$$\begin{aligned}
\langle t | t \rangle &= e^{-\alpha} \langle s | (I \cosh(\alpha) + (\hat{\alpha} \cdot \vec{s}) \sinh(\alpha)) | s \rangle \\
&= e^{-\alpha} (\cosh \alpha + (\hat{\alpha} \cdot \hat{s}) \sinh \alpha)
\end{aligned}$$

Denoting $T_p = \langle t | t \rangle$, the transmission intensity is

$$T_p = \frac{1}{1 + \tanh \alpha} (1 + \tanh \alpha (\hat{\alpha} \cdot \hat{s})) \quad (8.1.6)$$

The transmission depends not only on the loss coefficient α but the relative orientation of the PDL $\hat{\alpha}$ to the incoming state of polarization $\hat{s}$. Note that $\hat{s}$ is the state incident on the PDL element, not necessarily the state launched into a fiber far away from the element. The extrema of transmission are

$$T_p = \begin{cases} 1 & \hat{\alpha} \cdot \hat{s} = 1 \\ e^{-2\alpha} & \hat{\alpha} \cdot \hat{s} = -1 \end{cases}$$

Calculation of the PDL in dB recovers the definition (8.1.3). The relation between loss coefficient α PDL in decibels clearly holds in general.

One point to note is that transmission coefficients T_p do not multiply. That is, $T_{p21} \neq T_{p2} T_{p1}$. This is of course the result of change in the output polarization state from each PDL element, and a cascade exhibits multiple polarization alternations along its length. The correct calculation is $T_{p12} = \langle s | P_1^\dagger P_2^\dagger P_2 P_1 | s \rangle$.

The scalar transmission T_p can be mapped in Stokes space to show all possible polarization inputs. Figure 8.2 illustrates four such examples. The surfaces are the product of the transmission coefficient with input states that lie on the unit sphere. Figures 8.2(a,b) illustrate single PDL elements with 3 dB and 30 dB PDL, respectively. Note that the latter surface goes concave. The

transmission contour along the equator is projected to the bottom surface, where the orthogonal directions of no-loss and high-loss are apparent. Figures 8.2(c,d) illustrate the transmission through two PDL elements in cascade, both having 3 dB PDL. In the first case the PDL elements are aligned, resulting in a 6 dB overall PDL. However, in the second case the PDL elements are crossed. This results in a sphere with radius 0.5. There is no PDL at the output, but there is a uniform insertion loss of 3 dB. This shows a simple example of the more general fact that PDL can increase or decrease depending on the relative orientation of multiple elements, but the insertion loss always accumulates.

Thus far the input has been considered completely polarized. When the input is completely depolarized, the transmission is averaged over all polarization states. Since it is clear that $\langle \hat{\alpha} \cdot \hat{s} \rangle = 0$, the average transmission for unpolarized light is

$$\langle T_p \rangle = T_{\text{depol}}$$

where T_{depol} is the transmission coefficient for depolarized light. To connect with the literature, (8.1.6) is recast into

$$T_p = T_{\text{depol}} (1 + \Gamma(\hat{\alpha} \cdot \hat{s})) \quad (8.1.7)$$

where the loss coefficient is functionally normalized as

$$\Gamma \equiv \tanh \alpha \quad (8.1.8)$$

and the transmission for depolarized light is

$$T_{\text{depol}} = \frac{1}{1 + \Gamma} \quad (8.1.9)$$

The minimum and maximum transmissions are therefore

$$\begin{aligned} T_{\text{max}} &= T_{\text{depol}} (1 + \Gamma) \\ T_{\text{min}} &= T_{\text{depol}} (1 - \Gamma) \end{aligned}$$

The relationship between the normalized loss coefficient and the transmission extrema is

$$\Gamma = \frac{T_{\text{max}} - T_{\text{min}}}{T_{\text{max}} + T_{\text{min}}}, \quad \text{and} \quad \frac{T_{\text{max}}}{T_{\text{min}}} = \frac{1 + \Gamma}{1 - \Gamma} \quad (8.1.11)$$

The decibel expression for PDL can be written in these terms:

$$\rho_{\text{dB}} = 10 \log_{10} \left(\frac{1 + \Gamma}{1 - \Gamma} \right) \quad (8.1.12)$$

Moreover, these definitions are used to write the Jones and Mueller matrices for a PDL element oriented along the horizontal:

$$\mathbf{J}_{\hat{s}_1} = T_{\text{depol}}^{1/2} \begin{pmatrix} \sqrt{1 + \Gamma} & \\ & \sqrt{1 - \Gamma} \end{pmatrix} \quad (8.1.13)$$

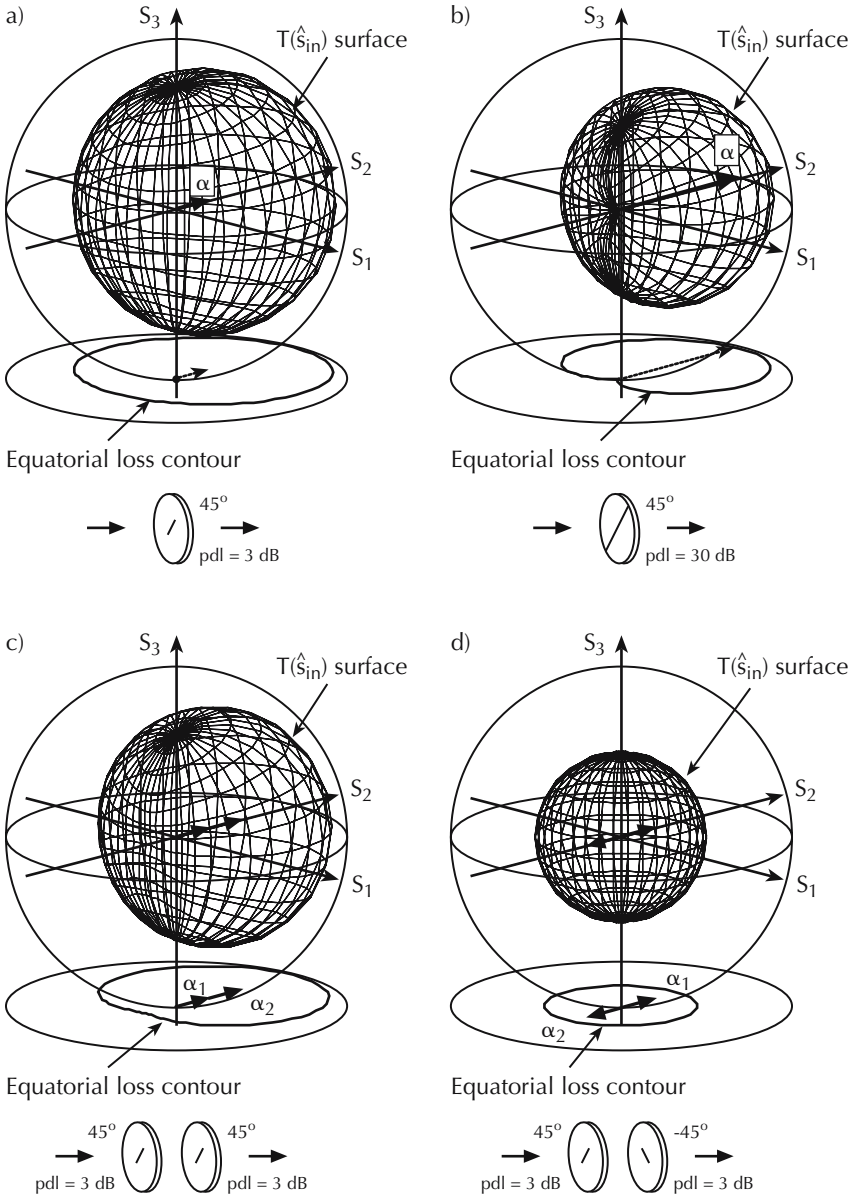


Fig. 8.2. These surfaces plot the transmission T_p as a function of input polarization state $\hat{s}_{in}$. The contours at the bottom are projections of T_p along the equator. a) Transmission surface for single 3 dB PDL vector at $+45^\circ$. b) Transmission surface for same PDL element but with 30 dB PDL. Note this surface is concave. c) Transmission surface after two aligned PDL elements, 3 dB PDL each. d) Transmission surface after two orthogonally aligned PDL elements, 3 dB PDL each. Note surface is spherical (no PDL) but has a smaller radius reflecting the insertion loss.

and

$$\mathbf{M}_{\hat{s}_1} = T_{\text{depol}} \begin{pmatrix} 1 & \Gamma & & \\ \Gamma & 1 & & \\ & & \sqrt{1-\Gamma^2} & \\ & & & \sqrt{1-\Gamma^2} \end{pmatrix} \quad (8.1.14)$$

where the Mueller matrix is related to the Jones matrix through the spin-vector expression (1.4.22) on page 18. Note that the Jones matrix can also be written as $\mathbf{J} = \text{diag}(\sqrt{T_{\text{max}}}, \sqrt{T_{\text{min}}})$.

8.1.2 Change of Polarization State

The state of polarization at the output from a PDL element generally differs from the input state. This is easily imaged: a right-hand circular state that is transmitted through a PDL element has one axis shortened. Accordingly, the state is altered from circular to elliptical.

The output polarization state is determined from the PDL operator P in the following way. The output unit Stokes vector is

$$\hat{t} = \frac{\langle t | \vec{\sigma} | t \rangle}{\langle t | t \rangle} \quad (8.1.15)$$

Substitution of the P expansion (8.1.5) into the numerator yields

$$\langle t | \vec{\sigma} | t \rangle = e^{-\alpha} \langle s | \left((I c_{\alpha/2} + s_{\alpha/2} (\hat{\alpha} \cdot \vec{\sigma})) \vec{\sigma} (I c_{\alpha/2} + s_{\alpha/2} (\hat{\alpha} \cdot \vec{\sigma})) \right) | s \rangle$$

where $s_{\alpha/2} = \sinh(\alpha/2)$ and $c_{\alpha/2} = \cosh(\alpha/2)$. Application of the identities

$$\begin{aligned} \langle s | \vec{\sigma} | s \rangle &= \hat{s} \\ \langle s | \vec{\sigma} (\hat{\alpha} \cdot \vec{\sigma}) + (\hat{\alpha} \cdot \vec{\sigma}) \vec{\sigma} | s \rangle &= 2\hat{\alpha} \\ \langle s | (\hat{\alpha} \cdot \vec{\sigma}) \vec{\sigma} (\hat{\alpha} \cdot \vec{\sigma}) | s \rangle &= 2\langle s | (\hat{\alpha} \hat{\alpha} \cdot) \vec{\sigma} | s \rangle - \langle s | \vec{\sigma} | s \rangle \\ &= 2\hat{\alpha} (\hat{\alpha} \cdot \hat{s}) - \hat{s} \end{aligned}$$

produces

$$\langle t | \vec{\sigma} | t \rangle = e^{-\alpha} \left[\hat{s} + s_{\alpha} (1 + t_{\alpha/2} (\hat{\alpha} \cdot \hat{s})) \hat{\alpha} \right]$$

Using the previously determined expression for $\langle t | t \rangle$, the unit Stokes vector $\hat{t}$ is governed by the relation [17]

$$\hat{t} = \frac{\sqrt{1-\Gamma^2}}{1 + (\Gamma \hat{\alpha} \cdot \hat{s})} \hat{s} + \frac{1 + \Gamma^{-2} (1 - \sqrt{1-\Gamma^2}) (\Gamma \hat{\alpha} \cdot \hat{s})}{1 + (\Gamma \hat{\alpha} \cdot \hat{s})} \Gamma \hat{\alpha} \quad (8.1.16)$$

where the following identification is made

$$\frac{\tanh(\alpha/2)}{\tanh \alpha} = \frac{1 - \sqrt{1-\Gamma^2}}{\Gamma^2}$$

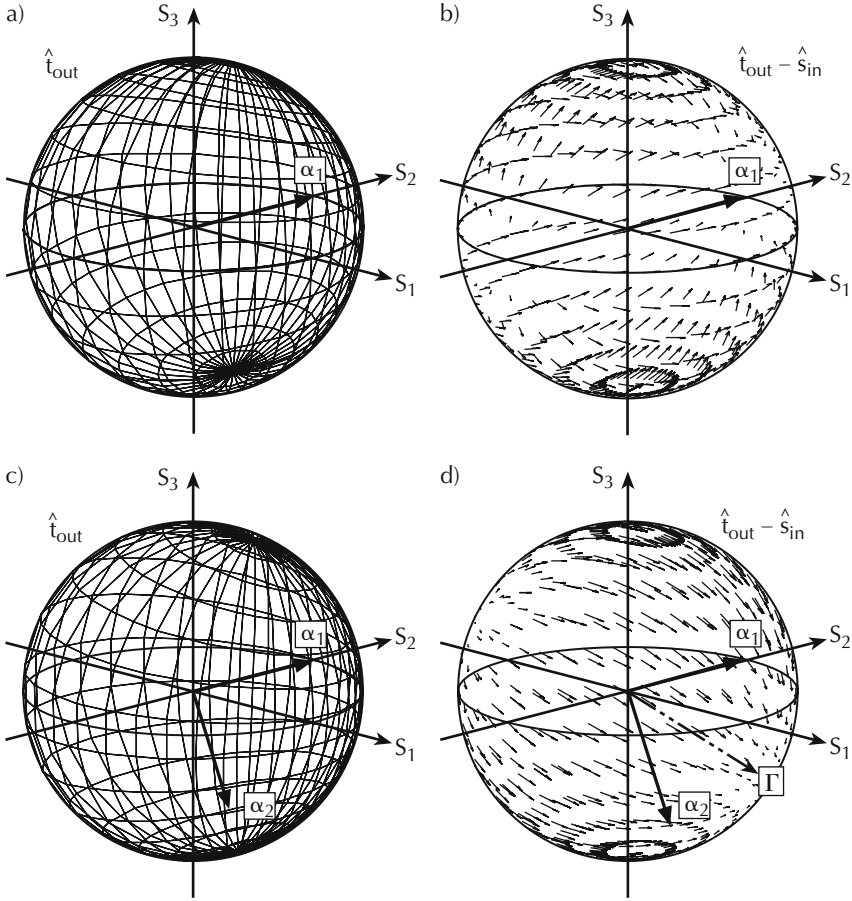


Fig. 8.3. Two examples of $\hat{t}_{\text{out}}$. Left figures show mesh of normalized $\hat{t}_{\text{out}}$ for one and two PDL elements. Right figures show vector difference plot to indicate the change $\hat{t}_{\text{out}} - \hat{s}_{\text{in}}$. a) and b) Single 3 dB PDL element aligned along S_2 . All states but $\pm S_2$ are pulled toward S_2 . c–d) Two 3 dB PDL elements cascaded, second element having elliptical PDL vector (or intermediate unitary rotation and linear PDL vector). Note in d) that cumulative PDL vector $\vec{\Gamma}$ points in a direction between $\vec{\alpha}_1$ and $\vec{\alpha}_2$. It is along $\vec{\Gamma}$ that the output states are pulled.

The transformation expression (8.1.16) has a characteristically different form than encountered previously in this text (Fig. 8.3). One is accustomed to $\hat{t} = R\hat{s}$, where R is a rotation operator in Stokes space. Instead, polarization-transformation through a PDL element has the form $\hat{t} = u\hat{s} + v\hat{a}$, where u, v are positive coefficients that are themselves a function of the relative orientation of $\hat{s}$ and $\hat{a}$. Physically, the input polarization state is “pulled” toward the PDL vector direction, where the pull strength is dictated by Γ and the relative orientation. This pulling effect is shown in Fig. 8.3(b,d).

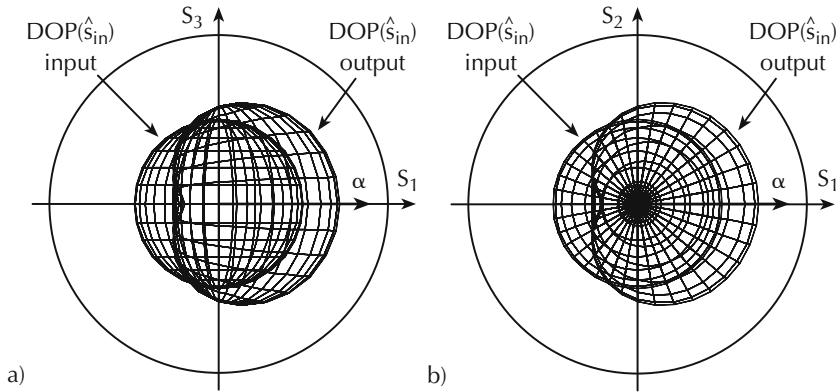


Fig. 8.4. Degree-of-polarization surfaces, before and after single PDL element with 3 dB loss and aligned along S_1 , illustrate repolarization. a) View of the (S_1, S_3) plane. b) View of the (S_1, S_2) plane. In both cases, spherical surface has $\mathcal{D}_{\text{in}} = 0.5$. The $\mathcal{D}_{\text{out}}$ surface is a map of $\mathcal{D}_{\text{out}}(\hat{s}_{\text{in}})$, or the DOP as a function of input polarization. The right hemisphere of both plots shows repolarization of the signal ($\mathcal{D}_{\text{out}} \geq \mathcal{D}_{\text{in}}$). The left hemisphere shows increased depolarization.

8.1.3 Repolarization

Intuitively one expects that depolarized light which passes through an ideal polarizer attains a unity degree of polarization. Repolarization [45] is the effect of generating partially polarized light from depolarized light by transiting through one or more PDL elements. However, transit of a PDL element with partially polarized light (cf. §1.5) can either increase or decrease the degree of polarization, depending on orientation [17].

The Mueller matrix for a single PDL element (8.1.14) governs re- and depolarization. Without loss of generality the matrix elements will remain fixed in the analysis below while the input polarization state varies. To summarize the findings:

$$\begin{aligned} \mathcal{D}_{\text{in}} = 1 &\longrightarrow \mathcal{D}_{\text{out}} = 1 \\ \mathcal{D}_{\text{in}} = 0 &\longrightarrow \mathcal{D}_{\text{out}} = \Gamma \\ \mathcal{D}_{\text{in}} = d &\longrightarrow \mathcal{D}_{\text{out}} = f(d, \Gamma, \hat{\alpha} \cdot \hat{s}) \end{aligned}$$

where f is the function (8.1.18). These cases are derived as follows. The output Stokes vector for an arbitrary input having $\mathcal{D}_{\text{in}} = d$ is

$$S_{\text{out}} = T_{\text{depol}} \begin{pmatrix} 1 & \Gamma \\ \Gamma & 1 \\ & \sqrt{1 - \Gamma^2} \\ & \sqrt{1 - \Gamma^2} \end{pmatrix} \begin{pmatrix} 1 \\ d \cos \phi \sin \theta \\ d \sin \phi \sin \theta \\ d \cos \theta \end{pmatrix} \quad (8.1.17)$$

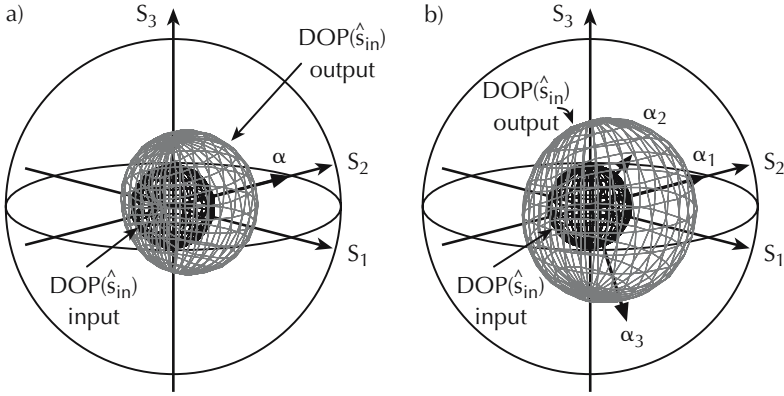


Fig. 8.5. Repolarization surfaces for $\mathcal{D}_{\text{in}} = 0.25$ after one (a) and three (b) PDL elements. All elements have 3 dB PDL. Note that for these large PDL values the repolarization is quickly established.

For $d = 1$ the output partial polarization in all cases is $\mathcal{D}_{\text{out}} = 1$. Conversely, for $d = 0$, only the S_0 and S_1 terms survive so $\mathcal{D}_{\text{out}} = \Gamma$. For an arbitrary d such that $0 \leq d \leq 1$, the output DOP is

$$\mathcal{D}_{\text{out}} = \frac{\sqrt{(1 + d\Gamma c_\phi s_\theta)^2 - (1 - d^2)(1 - \Gamma^2)}}{1 + d\Gamma c_\phi s_\theta} \quad (8.1.18)$$

Two degree-of-polarization surfaces as plotted in Fig. 8.4. The partial polarization at the output increases when the polarized portion of the input light is aligned to the direction of maximum transmission and decreases when aligned to the direction of maximum extinction. This behavior follows what one would expect.

While the above examples used a single PDL vector aligned along S_1 , a Mueller matrix can be calculated for any PDL operator P or concatenation of operators. Such a generalized Mueller matrix is used to calculate the repolarization surfaces illustrated in Fig. 8.5. In both cases $\mathcal{D}_{\text{in}} = 0.25$ and PDL for each segment is 3 dB. While the single-segment case exhibits both re- and de-polarization, the three-segment case exhibits repolarization for all input states.

The effect of repolarization is statistical when many PDL elements are distributed along a birefringent cascade. This is the model of a transmission link. Repolarization can be an impairment for long-haul transmission systems that launch polarized light to mitigate polarization-dependent gain effects from the erbium amplifiers. As the light becomes repolarized, PDG impairments begin to accumulate. The statistics of repolarization are derived in [45].

8.1.4 PDL Evolution Equations

Polarization-dependent loss in a communications link comes from optical components or imperfections along the optical fiber. In either case a local PDL vector $\vec{\alpha}$ each source. Polarization-dependent loss through a chain of PDL element accumulates and is denoted by $\vec{\Gamma}$. The cumulative PDL vector will always try to track to the local PDL element, but if the elements have low PDL and are randomly distributed, the cumulative vector eventually decorrelates. In this section the evolution of $\vec{\Gamma}$ is derived; its driving term is the local PDL $\vec{\alpha}$.

Two slightly different derivations for the evolution of $\vec{\Gamma}$ through a chain of arbitrarily oriented PDL elements are given by Gisin and Huttner, and Mecozzi and Shtaif [44]. The Gisin and Huttner derivation [18, 31] tracks the evolution of the PDL vector as defined at the output. Mecozzi and Shtaif do the opposite, tracking the PDL vector as defined at the input. Consider the output polarization produced by a chain of PDL elements acting on an input state: $|t\rangle = AT|s\rangle$ where, as in (8.1.4), A is the common loss and T is an Hermitian operator. The output intensity is then

$$\langle t|t\rangle = A^2 \langle s|T^\dagger T|s\rangle \quad (8.1.19)$$

Since T is Hermitian, the operator $T^\dagger T$ may be written in spin-vector form: $T^\dagger T = p_o I + \vec{p} \cdot \vec{\sigma}$. Taking $\langle s|s\rangle = 1$, one finds

$$\langle t|t\rangle = A^2 (p_o + \vec{p} \cdot \hat{s}) \quad (8.1.20)$$

The maximum and minimum output intensities are for alignment and anti-alignment of the input state $\hat{s}$ with $\vec{p}$. Therefore $\vec{p}$ represents the cumulative PDL vector referenced to the input.

Reference to the output state comes from the complement of (8.1.19):

$$\langle s|s\rangle = A^{-2} \left(\langle t|(TT^\dagger)^{-1}|t\rangle \right)$$

Denoting $QQ^\dagger$ as the inverse of $TT^\dagger$, the spin-vector form is $QQ^\dagger = q_o I + \vec{q} \cdot \vec{\sigma}$. Still assuming the input intensity is normalized, the output intensity can be expressed as a function of the output polarization state:

$$\langle t|t\rangle = \frac{A^{-2}}{q_o + \vec{q} \cdot \hat{t}}$$

Now the minimum and maximum output intensities are for alignment and anti-alignment of the *output* state $\hat{t}$ with $\vec{q}$.

The derivation below follows the Mecozzi and Shtaif choice of reference frame: equations of motion are referred to the input.

Referring back to (8.1.20), the extrema in transmission are clearly

$$T_{\max} = A_N^2 (p_o + p), \quad \text{and} \quad T_{\min} = A_N^2 (p_o - p)$$

where $p = |\vec{p}|$. The PDL is therefore

$$\rho_{\text{dB}} = 10 \log_{10} \left(\frac{1 + p/p_o}{1 - p/p_o} \right) \quad (8.1.21)$$

In light of the analogous expression (8.1.12), one can expect that the cumulative PDL magnitude Γ is related to the spin-vector as $\Gamma = p/p_o$. The vector form follows this relation and will be used in a moment.

The evolution of p_o and $\vec{p}$ is determined by the evolution of $T^\dagger T$. In the continuum limit, the cumulative loss A and PDL operator T are

$$A = \exp \left(-\frac{1}{2} \int_0^z \alpha(z) dz \right), \quad \text{and} \quad T = \exp \left(\frac{1}{2} \int_0^z \vec{\alpha}(z) \cdot \vec{\sigma} dz \right) \quad (8.1.22)$$

where $\vec{\alpha}(z)$ represents the derivative in z of the local PDL vector. (Gisin and Huttner use a discretized version to emphasize the form of the derivatives to follow.) The output intensity is

$$\langle t | t \rangle = A^2 \langle s | T^\dagger T | s \rangle = A^2 \langle s | p_o I + \vec{p} \cdot \vec{\sigma} | s \rangle \quad (8.1.23)$$

Now, the evolution is determined by the differential change of $T^\dagger T$. Taking the derivative along z (from input to output) of both sides of (8.1.23) gives

$$\frac{dT^\dagger}{dz} T + T^\dagger \frac{dT}{dz} = \frac{d}{dz} (p_o I + \vec{p} \cdot \vec{\sigma})$$

with the initial conditions of $p_o = 1$ and $\vec{p} = 0$ at $z = 0$. The spatial evolution of T and its adjoint comes from (8.1.22) and (2.5.82) on page 63

$$\frac{dT^\dagger}{dz} T + T^\dagger \frac{dT}{dz} = \frac{1}{2} ((\vec{\alpha}(z) \cdot \vec{\sigma}) T^\dagger T + T^\dagger T (\vec{\alpha}(z) \cdot \vec{\sigma}))$$

Plugging in the spin-vector form of $T^\dagger T$ and using the anti-commutator relation $\{(a \cdot \vec{\sigma}), (b \cdot \vec{\sigma})\} = 2(a \cdot b)$ gives

$$\frac{d}{dz} (p_o I + \vec{p} \cdot \vec{\sigma}) = p_o (\vec{\alpha}(z) \cdot \vec{\sigma}) + (\vec{\alpha}(z) \cdot \vec{p})$$

Finally, separation of spin-vector from scalar terms gives the coupled evolution equations

$$\frac{dp_o}{dz} = \vec{\alpha}(z) \cdot \vec{p}, \quad \text{and} \quad \frac{d\vec{p}}{dz} = p_o \vec{\alpha}(z) \quad (8.1.24)$$

These equations are converted to equations for the cumulative PDL vector and depolarization transmission. Define the cumulative PDL vector as

$$\vec{\Gamma}(z) \equiv \frac{\vec{p}(z)}{p_o(z)} \quad (8.1.25)$$

and the depolarization transmission T_{depol} as

$$T_{\text{depol}} = A_N^2 p_o \quad (8.1.26)$$

Differentiation with respect to length and substitution of (8.1.24) yields the equations of motion for $\vec{\Gamma}$ and T_{depol} :

$$\frac{d\vec{\Gamma}}{dz} = \vec{\alpha} - (\vec{\alpha} \cdot \vec{\Gamma}) \vec{\Gamma} \quad (8.1.27a)$$

$$\frac{dT_{\text{depol}}}{dz} = (\vec{\alpha} \cdot \vec{\Gamma} - \alpha) T_{\text{depol}} \quad (8.1.27b)$$

These elegant equations predict the direction and length of the PDL vector in Stokes space as well as the depolarized transmission coefficient. Note that $0 \leq |\vec{\Gamma}| \leq 1$, so the PDL vector is bound by the unit sphere.

As for discussion, note how the local PDL element $\vec{\alpha}$ pulls the cumulative PDL vector. If the cumulative and local vectors are initially orthogonal at the start of the n^{th} segment, then $\vec{\alpha} \cdot \vec{\Gamma} = 0$. The differential equation (8.1.27a) then pulls $\vec{\Gamma}$ strongly in the direction of $\vec{\alpha}$. Moreover, since $\Gamma \leq \alpha$ always, $\vec{\Gamma}$ asymptotically approaches $\vec{\alpha}$, but never aligns perfectly unless they are aligned initially. This said, the magnitude of $\vec{\Gamma}$ may increase or decrease depending on the cascade.

In contrast, T_{depol} monotonically decreases in all cases. At best T_{depol} is stationary. Lost power is never recovered when propagating through PDL media. Since $0 \leq \Gamma \leq 1$, $\vec{\alpha} \cdot \vec{\Gamma} \leq \alpha$ always.

Figure 8.6 illustrates four examples of cumulative PDL evolution. The details are given in the caption. One point to note, however, is the behavior of the insertion loss in the case of many small-valued PDL elements. Figures 8.6(c,d) show a trend of linear decrease (on a log scale) of T_{depol} . The origin of this behavior is as follows. The general solution to (8.1.27b) is

$$T_{\text{depol}} = \exp \left(- \int_0^z (\vec{\alpha}(z) - \vec{\alpha}(z) \cdot \vec{\Gamma}(z)) dz \right)$$

In a long chain of PDL elements, the cumulative vector eventually loses track of the local PDL element vector. That is, $\vec{\alpha}$ and $\vec{\Gamma}$ decorrelate. Once the two vectors decorrelate, one can write that the statistical average of the inner product vanishes: $\langle \vec{\alpha}_j \cdot \vec{\Gamma}_j \rangle = 0$. Beyond this point, the mean value of the insertion loss goes as

$$T_{\text{depol}} \simeq \exp \left(- \int_0^z \alpha(z) dz \right)$$

The insertion loss plots in Figures 8.6(c,d) illustrate this trend. Moreover, the variance of T_{depol} is related to the variance $\text{var}(\vec{\alpha} \cdot \vec{\Gamma})$. The variance of the cosine function uniformly distributed in angle is $1/2$. When all PDL elements have the same value, the variance of T_{depol} is approximately

$$\text{var}(T_{\text{depol}}) \simeq \exp(-\langle \alpha \rangle z / 2)$$

where $\langle \Gamma \rangle$ is the mean Γ over the link.

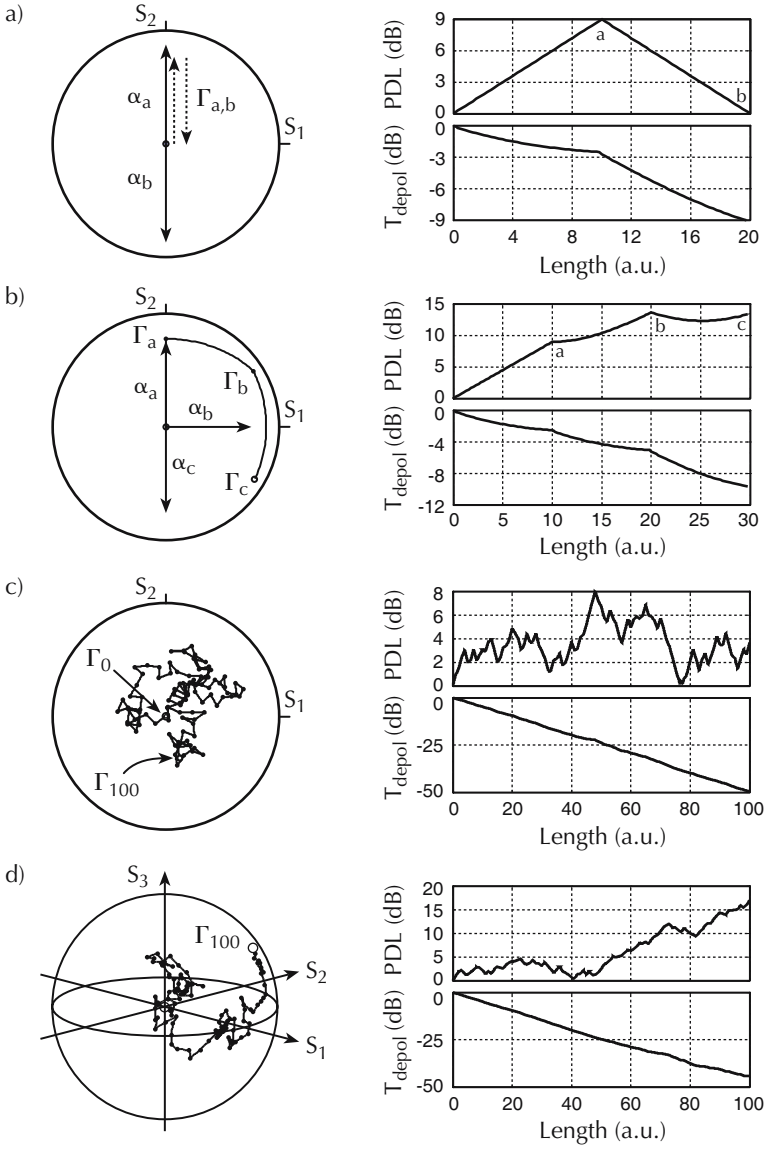


Fig. 8.6. PDL-vector evolution examples. a) Two orthogonal PDL segments $\vec{\alpha}_{1,2}$, 9 dB each. The PDL accumulates to a maximum of 9 dB through the first element and diminishes to zero at the termination of the second element. The insertion loss T_{depol} monotonically decreases to -9 dB. b) Three PDL segments $\vec{\alpha}_{1,2,3}$, 9 dB each, oriented at right angles. The cumulative PDL vector $\vec{\Gamma}$ tries to track the element vectors with increasing disparity. c) One hundred 1 dB randomly oriented PDL segments confined to the equator. The cumulative vector $\vec{\Gamma}$ does a random walk. The insertion loss decreases almost linearly (on log scale). d) One hundred 1 dB PDL segments with random orientation in all directions.

8.2 Polarization-Mode Dispersion

C. D. Poole and R. E. Wagner’s seminal paper, entitled, “Phenomenological Approach to Polarisation Dispersion in Long Single-Mode Fibres” [52] marks the historical dividing line between the classical treatment of birefringence and the modern development of polarization-mode dispersion (PMD). In 1986, Drs. Poole and Wagner, while working at Bell Laboratories on fiber communications, created a characteristic description of the frequency-dependence of concatenations of birefringent elements, since recognized as the point of discovery of PMD. Few discoveries, however, are made in a vacuum – indeed a scan of the technical literature shows titles including the term “Polarization-Mode Dispersion” dating back at least to 1978 [2, 56] – and one is inclined to ask about the context which led to the breakthrough.

The year 1986 was before the advent of the fiber-based erbium-doped optical amplifier; the pinnacle of long-distance communications rested at the time on coherent communications. Coherent communications mixes a local oscillator (LO) with the received signal to pull the signal out of the noise. As early as the 1970s, birefringence of single-mode fiber was recognized and that the output polarization state would drift with temperature or abruptly shift when hit was well understood. Since the LO needs to be aligned to the signal polarization, the matter of polarization tracking and control was under intense investigation. Many papers can be found on these subjects throughout the 1980s. In this sense, polarization effects were treated in the time domain, for the speed and control of polarization reside there.

As head of a research department at Crawford Hill Research, Wagner tasked Poole to investigate the overall link properties and characteristics that needed to be understood in order to make coherent communications practicable. A Bell Labs colleague of Dr. Poole’s at the time, N. S. Bergano, believed that the job would be completed somewhat quickly. Regarding polarization Dr. Bergano was to say, “its x , y , and a phase – how hard can it be?”.

Part of Dr. Poole’s work included building a relatively fast polarimeter to measure the speed of polarization change. When a fiber was terminated by the polarimeter and a probe diode laser was turned on, Poole noticed that the output polarization went through a transient but would settle out in a few minutes. Intrigued by this effect, Poole determined it was the diode laser ramping up in temperature which in turn produced a frequency sweep that was responsible for the changing output polarization. The unavoidable conclusion was that the output polarization state was frequency dependent. Determined to describe the effect as an eigenvalue problem, Poole found that the eigenvectors of a Jones transformation matrix always change to first-order in frequency. Nothing distinctive could be derived from that behavior. Poole and Wagner then asked if there exists an output state that is stationary to first-order in frequency – consequently the two Researchers discovered the principal characteristic of PMD.

The principal distinction between PMD research before and after Poole and Wagner's paper is that PMD enables a global description of the birefringence. Only local descriptions were available before 1986. Indeed, polarization optics had been studied for centuries, but always in the context of local birefringent behavior. Perhaps the closest earlier researchers got to the questions that encompass PMD are the several inventions of birefringent filters, including Lyot, Solc, Jones, Pancharatnam, and Harris. Even though, no one had put a global description together before and it is not too surprising that communications researchers were the first ones to make this observation.

With the advent of the optical amplifier, work on coherent communications went into decline. PMD, by contrast, has remained at the forefront of research ever since. Particularly important work includes the statistical treatment of PMD; the use of the statistics to develop link budgets for installed system designs; the measurement and mitigation of PMD in single-mode fiber; the measurement of PMD in components, installed fiber, operating links, and all manner of configurations; programmable PMD generation; and interaction of impairments such as PMD with PDL and chromatic dispersion, and PMD with nonlinear effects.

Moreover, the wealth of research developed in polarization tracking for coherent communications was redirected to active PMD compensation. The father of the optical PMD compensator is Fred Heismann, who through his work on lithium-niobate electro-optic polarization controllers [27, 28] demonstrated the first closed-loop PMDC [29]. Yet at the time of writing, the economics behind optical PMDCs appear unfavorable, even in light of proven high-quality, live-traffic-certified products [50]; and, simultaneously, lower-cost chip sets are being developed to make corrections electronically.

The following sections of this chapter cover the major highlights of the time, frequency, Fourier, and Stokes properties of PMD. These sections will be successful if they arm the reader with the tools necessary to read the literature and patents critically and informatively.

8.2.1 A PMD Primer

Polarization-mode dispersion is an optical effect present in concatenations of lossless birefringent elements. The following observation was first made by Poole [52]:

OBSERVATION: There always exists an orthogonal pair of polarization states output from a birefringent concatenation which are stationary to first order in frequency. These two states are called the Principal States of Polarization (PSP). A differential delay exists between signals launched along one PSP and its orthogonal complement.

Based on this observation, the PMD vector, which has a length and a pointing direction in Stokes space, is defined as follows [20]¹ :

DEFINITION PART 1: The pointing direction of the PMD vector is aligned to the slow PSP, the PSP that imparts more delay than the other.

DEFINITION PART 2: The length of the PMD vector is the differential-group delay between slow and fast PSP's.

The definition of the PMD vector follows immediately from the Poole observation, and that observation is clearly different from the usual input-to-output eigenstate definition for a birefringent concatenation.

The “usual” eigenstate of any concatenation is that polarization state which is the same at the output as at the input. (In contrast, the polarization state associated with the input and output PSP's are generally not the same.) The failure of the eigenstate description happens because the eigenstate almost always changes to first order with change in frequency. Since group delay requires a frequency derivative, the derivative of an eigenstate and its phase results in two terms: the derivative of the phase, which is delay, and the derivative of the eigenstate, which rotates the polarization. The lack of a stationary eigen-polarization prevents the separation of group delay from its eigenstate. In the context of eigenstates, differential-group delay is not uniquely determined.

Another way to express the need for a definition of the PMD vector is to say that a parallelism should exist for infinitesimal rotations of polarization state in length and in frequency. As seen many times earlier in this text, the differential change of output polarization state due to differential change in length of the last birefringent element in a concatenation is

$$\frac{d\hat{s}}{dz} = \beta_n \times \hat{s}$$

where β_n is the birefringent vector of the last element. This is the customary precession rule. What vector, then, does the output polarization precess about for differential change in frequency? That is, what is the vector V such that

$$\frac{d\hat{s}}{d\omega} = V \times \hat{s} \quad ?$$

Equivalent to a change in frequency would be a if all elements in the concatenation changed their length, or refractive index, or both, for the same frequency. That is, the optical phase for any section is

¹ Note: This definition of PMD follows that of Gordon and not that adopted by Poole. The Gordon definition aligns the PMD vector along the slow PSP and uses a right-hand coordinate system in Stokes space. A detailed comparison between the two systems is available [21].

$$\varphi = \frac{\omega}{c} nL$$

Proportional change of any parameter of the optical phase along an entire cascade leads to eigenstate change for first-order change in that parameter. For example, a change in temperature will change the refractive index, and eigenstates of the cascade with change to first order with it. This generalization of PMD was recognized by Shieh and Kogelnik [58]. It is customary, nonetheless, to use frequency as the underlying parameter.

It is important to observe that so far no specific context for PMD has been given other than lossless birefringent concatenations. PMD exists in optical fiber as well as birefringent crystals; it exists for any number of homogeneous sections, whether one or thousands; it exists for any type of statistical distribution, whether Maxwellian or not, and any delay within or accumulated across multiple sections. PMD also exists in the presence of PDL, although its description is more complicated. Whether PMD is a significant effect for an optical communications link or a single component depends on these particulars, but the particulars do not alter its definition.

It is not at all obvious that for any arbitrary lossless birefringent cascade of any length and composition a pair of principal polarization states exists. A good heuristic is to construct and to compare the differential-rotation rules for length and frequency, and to draw an analogy between the principal-state system and the eigenstate system.

To begin, an eigenstate ties the output to the input: the same polarization state must exist at both ends (at any frequency). Likewise, a principal state ties the output to the input but in a different way: the input polarization state must be oriented such that the output polarization state does not change to first order in frequency. Generally, the eigenstate and principal state are not the same. Figure 8.7 illustrates the eigenstates for a single homogeneous birefringent element. When an input polarization is oriented to the “fast” eigen-axis of the element (Fig. 8.7(a)), the refractive index seen by the light is the smaller of the two. The output polarization is the same as the input polarization, and a pulse of light exits at a certain time. Next, when the input polarization is oriented to the “slow” eigen-axis of the element (Fig. 8.7(b)), the refractive index seen by the light is the larger of the two. The output polarization is also the same as the input polarization, and an output pulse is delayed with respect to the fast-axis output pulse. The relative delay time is τ , and is called the differential-group delay (DGD). Note that when the input polarization is oriented either to the fast or slow axis, only one polarization state and one pulse is present at the output; this defines the eigenstate of the system.

In general an input polarization is not aligned to the birefringent axis of the element (Fig. 8.7(c)). In this case, the input state is projected onto the two orthogonal birefringent axes of the element and these projects propagate

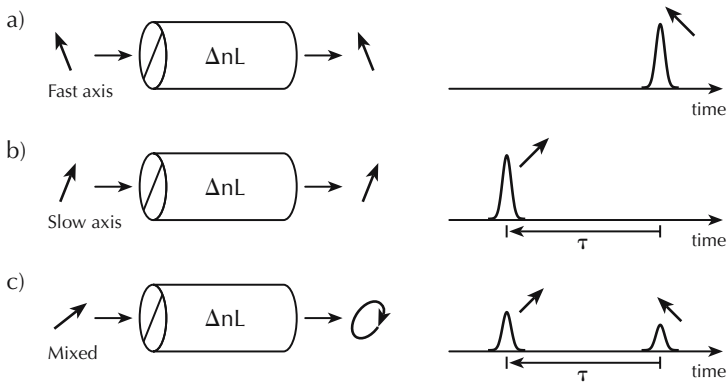


Fig. 8.7. Eigenstates and effect of single homogeneous birefringent element. a) Launch along fast axis gives an output polarization state that is the same as the input state. b) Launch along slow axis also leaves the input state unaltered. There is, however, a delay with respect to the fast axis. c) Arbitrary polarization state at the input is projected onto the two eigen-axes, and these two projects propagate independently.

independently to the output. Accordingly, an input pulse is split into two pulses, one delayed by τ with respect to the other. The relative intensity of the two pulses is dictated by the angle between the input polarization state and the birefringent axis of the element.

Next, consider two birefringent elements in cascade. The polarization evolution is best described in Stokes space. As in Fig. 8.8(a), consider one stage again. The input polarization $\hat{s}_{\text{in}}$ precesses about birefringent axis $\hat{r}_1$ to output state $\hat{s}_1(z_1)$, where the propagation axis is denoted as z . When a second stage is added (Fig. 8.8(b)), polarization state $\hat{s}_1(z_1)$ precesses about birefringent axis $\hat{r}_2$ to new output polarization $\hat{s}_2(z_2)$. Consider now a small increase in length Δz of the second stage. The output polarization comes from a continuation of the precession about $\hat{r}_2$ up to $z_2 + \Delta z$ (Fig. 8.8(c)). Thus the differential precession rule for change of length is $d\hat{s}/dz = \hat{r}_2 \times \hat{s}$.

Now, instead of an increment in length, consider an decrement in frequency.² At a first frequency ω , the output polarization from two stages is $\hat{s}_2(\omega)$ (Fig. 8.9(a)). When the frequency is decremented (Fig. 8.9(b)), the precession about $\hat{r}_1$ decreases, terminating the precession about $\hat{r}_1$ early. Accordingly, the radius of the precession circle centered on $\hat{r}_2$ is increased, and the precession angle about $\hat{r}_2$ is decreased due to the lower frequency. The output polarization is $\hat{s}_2(\omega - \Delta\omega)$. For this choice of input polarization state $\hat{s}_{\text{in}}$, the output polarization changes to first-order with frequency, Fig. 8.9(c). This

² In the following the frequency is decremented. The choice of increase or decrease in frequency is of no consequence; the decrease in frequency was selected only to make the illustrations clearer.

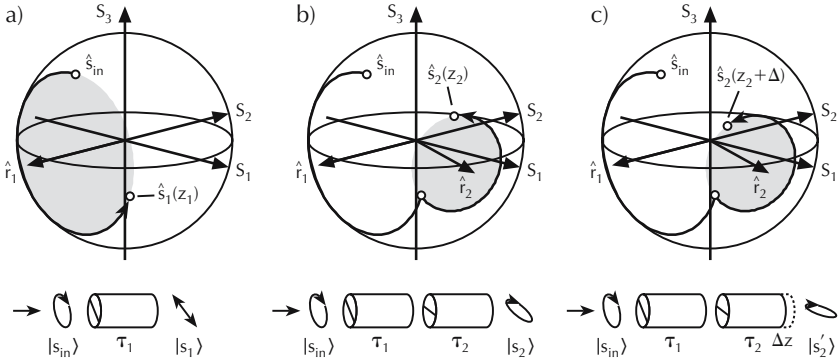


Fig. 8.8. Polarization evolution through one and two birefringent elements. a) Input state $\hat{s}_{in}$ precesses about birefringent axis $\hat{r}_1$ as the light travels along length z . b) The state output from the first stage is input to the second stage and precesses about $\hat{r}_2$. c) A small increase in length of the second stage continues the polarization precession to state $\hat{s}_2(z_2 + \Delta z)$ from $\hat{s}_2(z_2)$ about $\hat{r}_2$.

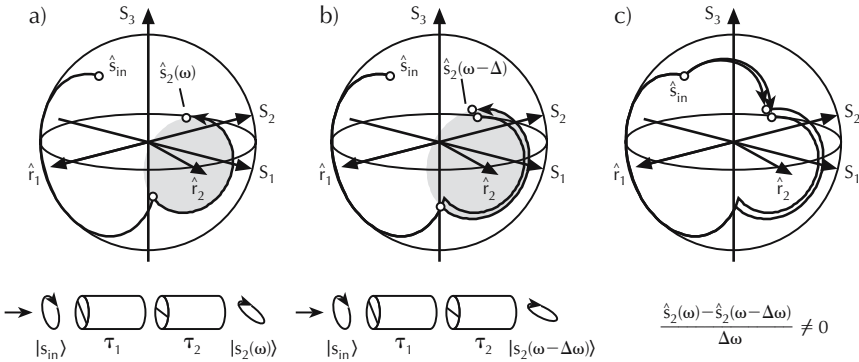


Fig. 8.9. Polarization evolution through two birefringent elements at two different frequencies. a) Evolution through two stages as in Fig. 8.8(b). b) A decrease in frequency reduces the precession through stages one and two. Reduced precession about $\hat{r}_1$ increases the radius of the precession circle centered on $\hat{r}_2$. c) For the same $\hat{s}_{in}$, the output polarization changes to first-order with frequency.

input state does not yield a principal state of polarization for the system at frequency ω .

A principal state of polarization for this system can be nonetheless found for a different input state.

Figure 8.10 shows the construction necessary for two equal-length stages. At frequency ω , input state $\hat{p}_{in}$ precesses about $\hat{r}_1$ until it reaches the equator. The polarization state then precesses about $\hat{r}_2$ to output state $\hat{p}_{out}$. Now, a decrement in frequency moves the intermediate polarization state $\hat{s}_1$ away

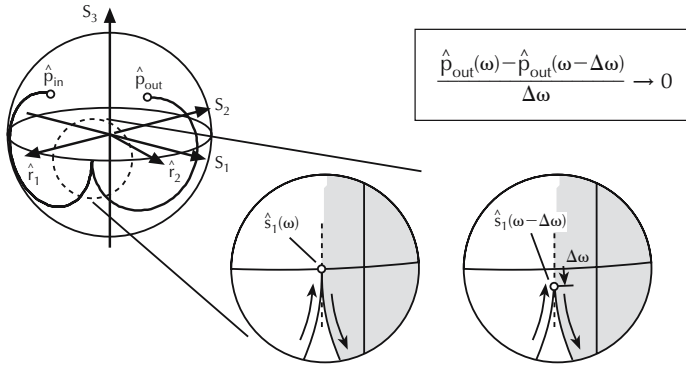


Fig. 8.10. Principal input and output states at ω for two equal-length birefringent stages. The insets show that a decrement in frequency does not change the radius of the precession circle about $\hat{r}_2$ to first order; and that decrease in precession about $\hat{r}_1$ is compensated by the requisite decrease in precession about $\hat{r}_2$ necessary to reach $\hat{p}_{\text{out}}$ to first order.

from the equator. However, as shown in the inset, the left-right motion of the polarization state as it precesses about $\hat{r}_1$ changes only to second order; the same holds for precession about $\hat{r}_2$. This is because the two precession circles share the same tangent line at their intersection. (That the two circles are tangent is the key point, the point of tangency happens to be at the equator in this example.) Accordingly, the precession circle centered on $\hat{r}_2$ does not change radius to first order. Second, the decrease in precession about $\hat{r}_1$ is compensated by a decrease in precession about $\hat{r}_2$ necessary to reach $\hat{p}_{\text{out}}$ at the output, to first order. Neither precession radii nor arc lengths change to first order with change in frequency. Therefore $\hat{p}_{\text{out}}$ is a principal state of the system. The stationary property is expressed as

$$\lim_{\Delta\omega \rightarrow 0} \frac{\hat{p}_{\text{out}}(\omega - \Delta\omega) - \hat{p}_{\text{out}}(\omega)}{\Delta\omega} = 0 \quad (8.2.1)$$

The output polarization state $\hat{p}_{\text{out}}$ is an output PSP. Its orthogonal state is also a PSP. The output principle states have corresponding input principal states, one of which is $\hat{p}_{\text{in}}$. The input and output principal states are related by the transformation matrix of the system:

$$|p_{\text{out}}\rangle = T |p_{\text{in}}\rangle$$

Unlike an eigenstate of T , the output PSP is generally not the same as the input PSP. However, the two output PSP's are orthogonal to one another, as are the two input PSP's.

Now, since the output PSP is stationary with frequency to first order, a group delay τ can be defined and is separable from the polarization state on which it travels. The existence of this PSP-dependent group delay was first

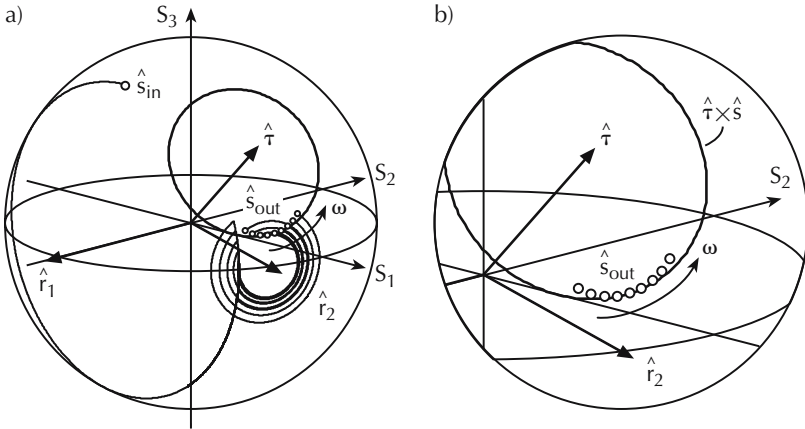


Fig. 8.11. The PMD vector $\vec{\tau}$ defines the precession rule in frequency for output polarizations. a) Evolution of polarization state through two stages from fixed $\hat{s}_{in}$ for range of frequencies. Precession circle $\vec{\tau} \times \hat{s}_{out}(\omega)$ is shown for comparison. b) Magnified view of output polarization state as function of frequency with precession circle. The two deviate only to second order.

experimentally demonstrated in the time-domain by Poole and Giles [51]. The above construction has not determined which of the two output PSP's is the slow axis and which is the fast axis, but nevertheless the frequency derivative of the phase of $\hat{p}_{out}$ imparted by T yields the differential-group delay denoted by τ . The DGD is a length and therefore always a positive quantity. By construction, the PMD vector points along the direction of the slow PSP and has length τ ; that is,

$$\vec{\tau} \equiv \tau \hat{p} \quad (8.2.2)$$

where $\hat{p}$ is the slow output PSP. The PMD vector defines the infinitesimal precession rule about which any output polarization state travels for small changes in frequency. That is,

$$\frac{d\hat{s}}{d\omega} = \vec{\tau} \times \hat{s} \quad (8.2.3)$$

where $\hat{s}$ is any polarization state at the output. This frequency-dependent precession rule was used by Curti *et al.* [1] as early as 1987 to measure the differential-group delay of a single-mode fiber.

Figure 8.11 illustrates a calculation of the output polarization states for nine different frequencies for an arbitrary but fixed input polarization state $\hat{s}_{in}$. Overlaid on the figure is the precession circle $\vec{\tau} \times \hat{s}$, where the output PSP is defined by $\hat{p}_{out}$ in Fig. 8.10. It is clear that the output states circle about $\vec{\tau}$ and deviate from the precession circle only to second-order.

That there is deviation at all between the output polarization states and the precession circle underscores a fundamental aspect of PMD. The PMD

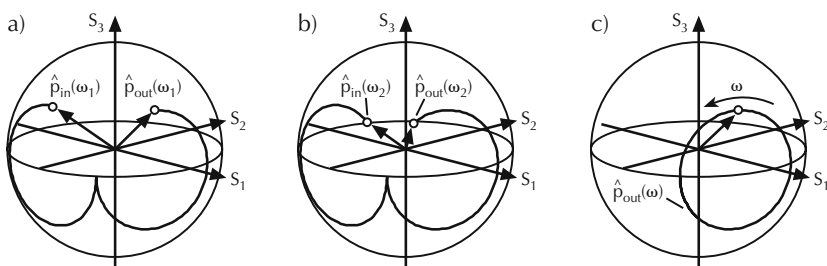


Fig. 8.12. The principal state of polarization changes with frequency. a) At ω_1 input and output PSP's are shown for a two-stage concatenation. b) At different frequency ω_2 new input and output PSP's are required to achieve a stationary output. c) The output PSP spectrum for all frequency. For two stages the spectrum is a circle. This spectrum is periodic.

vector is defined as a vector that is stationary to first order. However, the PMD vector is itself frequency dependent. The vector $\vec{\tau}$ in the infinitesimal precession rule (8.2.3) must be updated for each new frequency. One generally denotes the frequency dependence of the PMD vector as $\vec{\tau}(\omega)$.

Figure 8.12 illustrates the fundamental nature of PMD vector change. The PMD vector construction requires an output polarization that is stationary to first order with frequency for every frequency. Figure 8.12(a) shows the input and output PSP's at ω_1 , the same frequency as in Fig. 8.10. However, at a higher frequency ω_2 , more precession occurs about the birefringent axes. A new stationary state must account for this increase; this is shown in Figure 8.12(b). A locus of output PSP's is determined by a full frequency sweep. The locus is called the principal-state spectrum. The output PSP spectrum for two equal-length stages is shown in Figure 8.12(c). For two stages the PSP spectrum is always a circle in Stokes space. For more stages the spectrum shows more complicated motion.

That the PMD vector changes pointing direction with frequency is of great practical importance in optical communication systems. Generally there is also an associated change in the DGD with frequency. One should recognize that a principal state of polarization is as delicate and ephemeral as any state of polarization, and will change with any perturbation on the system.

The preceding work is now generalized to show the meaning and behavior of a PMD spectrum associated with any lossless system. Figure 8.13(a) illustrates the PMD vector in Stokes space for a particular frequency, $\vec{\tau}(\omega)$. The vector has a length $\tau(\omega)$ and a pointing direction $\hat{p}(\omega)$. The spectrum of this vector is composed of two components: the PSP spectrum (Fig. 8.13(b)), and the DGD spectrum (Fig. 8.13(c)). The PSP spectrum is a vector spectrum which indicates the principal state as a function of frequency. The DGD spectrum is a positive scalar spectrum which indicates the differential-group delay as a function of frequency. The PMD of any lossless system is entirely determined by these two spectra.

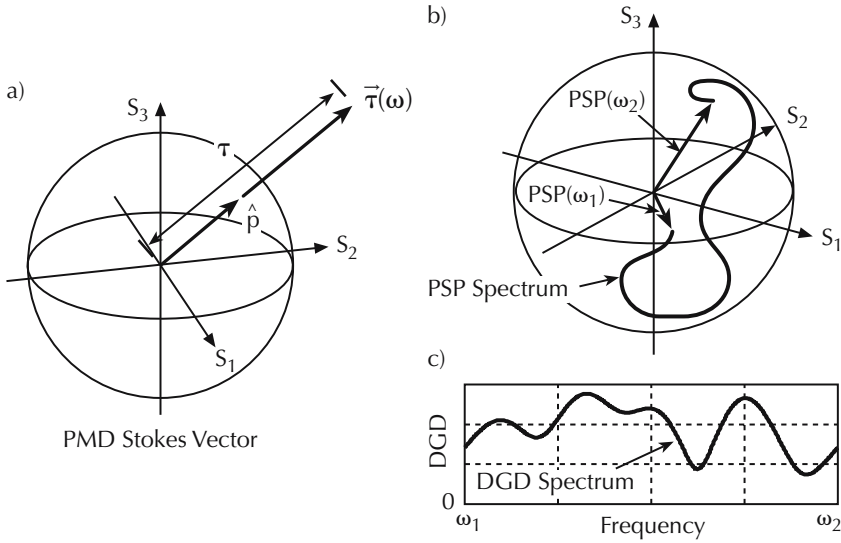


Fig. 8.13. The two components of a PMD spectrum. a) The PMD vector illustrated in Stokes space. The vector has a length τ and a pointing direction $\hat{p}$. These are functions of frequency. b) A PSP spectrum: $\hat{p}(\omega)$. The pointing direction changes with frequency on the unit sphere. c) A DGD spectrum: $\tau(\omega)$. The vector length, which is a positive scalar value, changes with frequency.

The connection between the two defining PMD spectra and the time domain is simple for narrow-band signals. Four narrow-band signals are illustrated in Fig. 8.14(a), having frequency centers at $\omega_1, \dots, \omega_4$. The overlaid DGD spectrum determines the delay between the two orthogonal polarization components of each signal. For instance, at ω_1 , the DGD spectrum has a high value which in turn imparts a large delay between the two components of the associated signal. At ω_4 , however, the DGD spectrum has a small value, so the temporal delay for orthogonal components of the associated narrow-band signal is small. Note, however, that the DGD spectrum provides no information as to the relative weight between split components of the signals. That is left to the PSP spectrum.

The effect of the PSP spectrum is illustrated in Fig. 8.14(b). At a particular frequency, the relative weight between orthogonal polarization components is determined by the projection of the input state onto the PSP for that frequency. While the illustration is not rigorous, the projection for the signal centered at ω_1 , which has a large differential delay, is about equal. In contrast, the projection for the signal centered at ω_4 , which has a small differential delay, is lopsided. The change of relative weights between these two signals is due to the change of the PSP pointing direction with frequency. Note, however, that the PSP spectrum provides not information as to the differential delay between

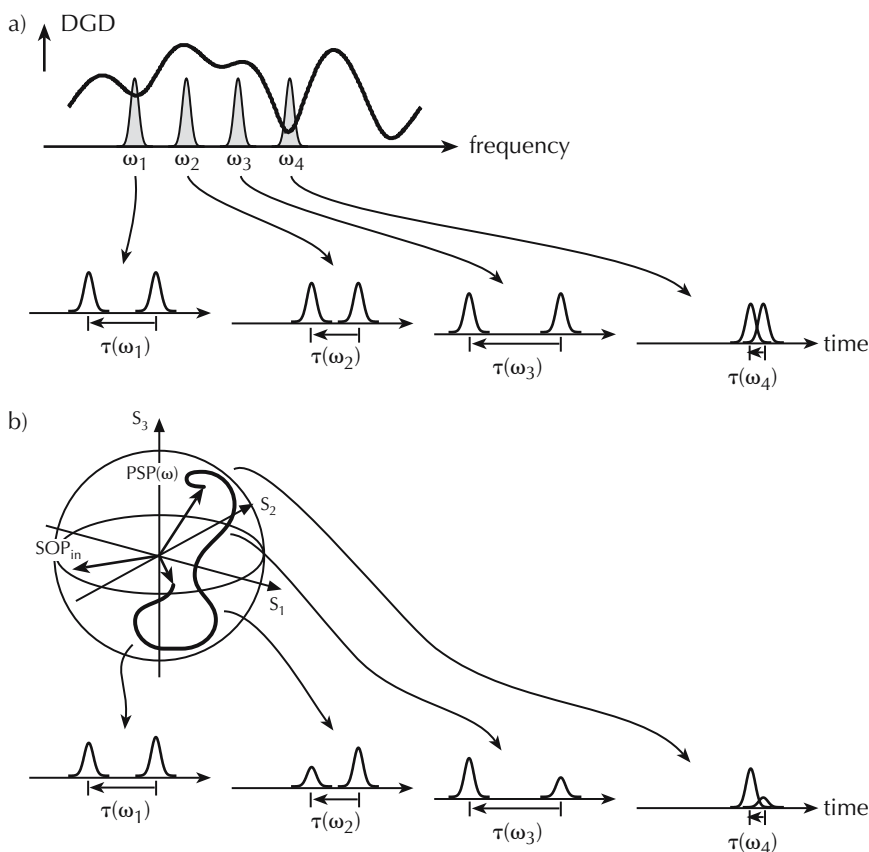


Fig. 8.14. Relation between frequency and time domains for various narrow-band signals affected by PMD in terms of the scalar and vector PMD spectra. a) The DGD spectrum determines the time delay between orthogonal polarization components of a narrow-band signal. For signal centered at ω_1 , the time delay is $\tau(\omega_1)$; the figure illustrates a large differential delay. For signal centered at ω_4 , the time delay is $\tau(\omega_4)$; the figure illustrates a small differential delay. b) The PSP spectrum determines the relative weight between orthogonal signal components for each frequency. The change in projection between the fixed input polarization and the PSP vector for different frequencies changes the relative weight between orthogonal signal components on each narrow-band signal.

split components of the signals. The DGD and PSP spectra are complementary and the two must be considered together.

More complicated pulse deformations occur when the PMD spectrum varies over the bandwidth of the signal. For the same PMD, the higher the data rate the broader the signal spectrum. In this case each spectral component can be analyzed as in Fig. (8.14), but then the interference between

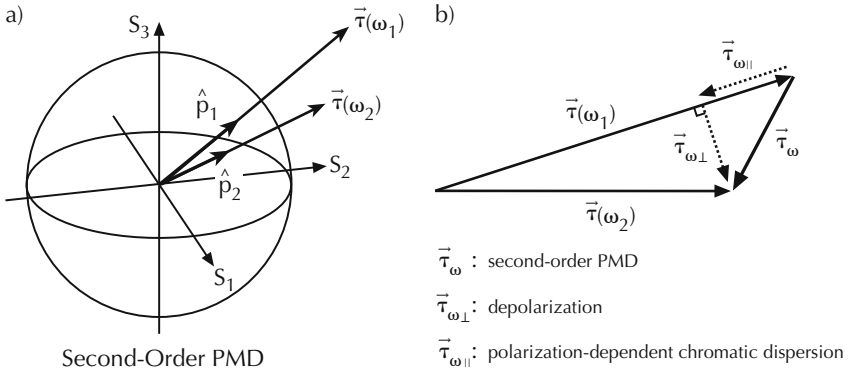


Fig. 8.15. Definition of the second-order PMD (SOPMD) vector $\vec{\tau}_\omega$. a) PMD vectors $\vec{\tau}(\omega_1)$ and $\vec{\tau}(\omega_2)$. The vectors have differential lengths and pointing directions. b) Second-order PMD is the vector difference $\vec{\tau}_\omega = (\vec{\tau}(\omega_2) - \vec{\tau}(\omega_1)) / \Delta\omega$ as $\omega_2 - \omega_1 \rightarrow 0$. The SOPMD vector can be resolved on the $\vec{\tau}(\omega_1)$ axis into perpendicular and parallel components. The perpendicular component is called depolarization and the parallel component is called polarization-dependent chromatic dispersion (PDCD).

these components must be accounted for. A measure of pulse-deformation complexity is found by the degree of PMD change over a signal bandwidth. If the PMD changes little, then only “first-order” PMD affects the pulse. If the PMD changes a lot, then “higher-order” PMD effects become pronounced. The higher-order PMD effects generate complicated pulse deformations.

One higher-order effect that draws much attention is second-order PMD. Second-order PMD (SOPMD) is a vector generated by a first-order change of the PMD vector with frequency. The SOPMD vector is illustrated in Fig. 8.15. In general, the PMD vector $\vec{\tau}$ is different for different frequencies. Consider two vectors in particular: $\vec{\tau}(\omega_1)$ and $\vec{\tau}(\omega_2)$, where $\omega_2 - \omega_1$ is a small change. The vector difference is the SOPMD vector, denoted by $\vec{\tau}_\omega$. Like any vector, $\vec{\tau}_\omega$ has a length and a pointing direction. The pointing direction is generally neither parallel nor perpendicular to either first-order PMD vector, and the length is zero only when $\vec{\tau}$ pirouettes about itself or when there is only one birefringent element. The length of the SOPMD vector in relation to the DGD determines the significance of the second-order vector.

The SOPMD vector is resolved onto two components called the depolarization component and the polarization-dependent chromatic dispersion (PDCD) component (Fig. 8.15(b)). The depolarization component $\vec{\tau}_{\omega\perp}$ runs perpendicular to $\vec{\tau}(\omega_1)$ and indicates the rate which the pointing direction of the PMD vector changes. The PDCD component $\vec{\tau}_{\omega\parallel}$ runs parallel to $\vec{\tau}(\omega_1)$ and indicates the change of DGD with frequency. The length of the PDCD component is in fact the frequency derivative of the DGD spectrum at ω_1 .

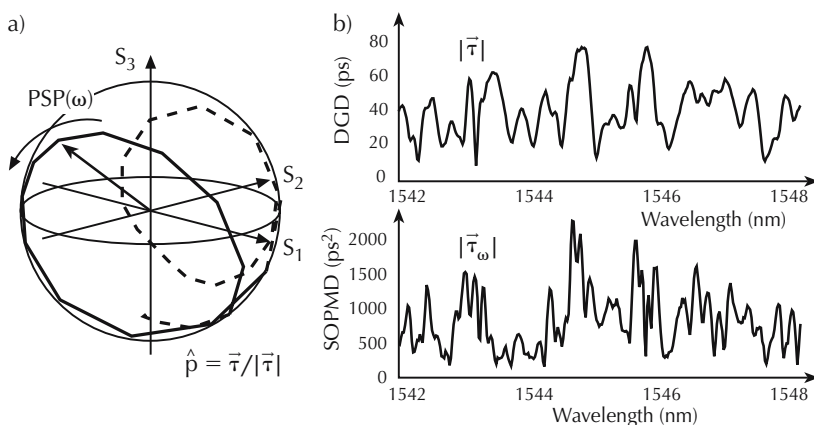


Fig. 8.16. Measurement of PSP, DGD, and magnitude-SOPMD spectra from a line of single-mode fiber. Courtesy D. Peterson, MCI [50].

The term “depolarization” has a connotation in the time domain, so why does SOPMD, best described in the frequency domain, have a depolarization component? The change in pointing direction of the PMD vector with frequency disperses a fixed input polarization into many states at the output. For each frequency there is one output state. When the signal is inverse Fourier transformed into the time domain, the dispersed output polarizations are folded into the time-domain signal so that each time interval contains many polarizations. A time average over these states reduces the degree-of-polarization as compared to the input; that is, the input state is depolarized by the PMD.

Second-order PMD is the frequency derivative of the PMD vector (8.2.2):

$$\vec{\tau}_\omega = \underbrace{\tau_\omega \hat{p}}_{\text{pdcd}} + \underbrace{\tau \hat{p}_\omega}_{\text{depolarization}} \quad (8.2.4)$$

where the two orthogonal vector components are identified in the expression. Statistically the depolarization component dominates the SOPMD vector.

Like first-order PMD, second-order PMD is itself dependent on frequency: $\vec{\tau}_\omega = \vec{\tau}_\omega(\omega)$. SOPMD therefore has a scalar and vector spectrum. Often the magnitude of first- and second-order PMD is plotted when comparing the two orders. Figure 8.16 shows measurements of the PSP, DGD, and magnitude SOPMD $|\vec{\tau}_\omega|$ as a function of frequency made on a line of single-mode fiber.

An alternative but equivalent view of PMD is to look at its response directly in the time domain. To do so, one looks at the impulse response of a cascade of birefringent sections. An impulse in time is a delta function whose

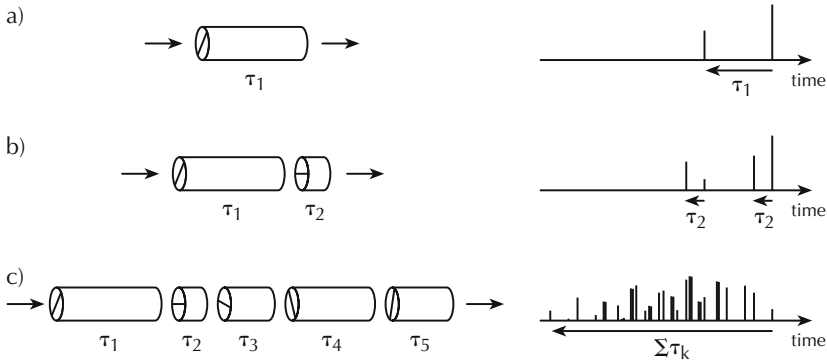


Fig. 8.17. Impulse response from one or more birefringent elements. The polarization of the impulses has been abstracted. a) Impulse response from one stage alone. An input impulse is split along the two birefringent axes and one pulse is delayed with respect to the other by τ_1 . b) Impulse response from two stages. The two impulses from the first stage are each divided into two parts, the slow components are then delayed by τ_2 . c) Impulse response from five stages generates 2^5 or 32 impulses. There is a first and last pulse, and the time response is FIR.

spectrum is uniform over all frequency. Since no two impulses can overlap unless they are precisely at the same time instant, no accounting for interference is required to determine the output response.

Figure 8.17 illustrates the impulse response from one, two, and five different-length birefringent elements. For one stage, Fig. 8.17(a), an impulse is split along the two birefringent axes and one impulse is delayed with respect to the other by τ_1 , the DGD of the element. When a second stage is added, Fig. 8.17(b), each impulse from the first stage is projected onto the birefringent axes of the second. The projection on the second stage is independent from the projection onto the first stage, so the relative impulse heights differ. The two components aligned to the slow axis of the second stage are delayed by τ_2 , yielding in general four impulses. Note that in the figure the state of polarization for each impulse is not indicated but only its relative weight and time position are shown. Through a cascade of five stages, Fig. 8.17(c), there are five projections and five delays, resulting in 2^5 or 32 distinct impulses. The impulse response can be complicated, but a characteristic of the response is that it is finite in duration. In general, PMD is a linear effect that acts as a finite-impulse response (FIR) filter. The temporal extent of the impulse response is $\sum \tau_k$.

The temporal extent can be compared with a pulse duration of a signal to determine the impairment of PMD. When a signal pulse, such as a non-return-to-zero ONE, is much longer in time than the FIR response of the birefringent cascade, there is little effect of PMD on the pulse. However, when the time extent of the signal pulse and the birefringent cascade are comparable, the signal pulse can be significantly distorted. Strictly, the polarization-dependent

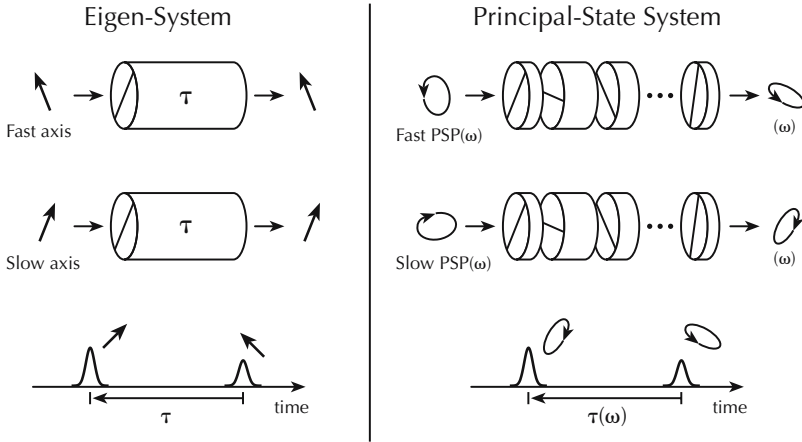


Fig. 8.18. Comparison between eigenstate and principal-state systems. Left: Eigen-system for a single birefringent stage. Input polarizations aligned to the two birefringent axes are output with no change in polarization. Temporally, one axis has a lower group index than the other, so the two eigenstates have a delay τ between them. Right: Principal-state system for an arbitrary birefringent cascade. For any frequency, there exists an input polarization such that the output polarization does not change to first-order in frequency. Along the pair of principal axes, there is a differential-group delay $\tau(\omega)$ between signals launched on orthogonal principal axes. Generally, the PSP and DGD change with frequency.

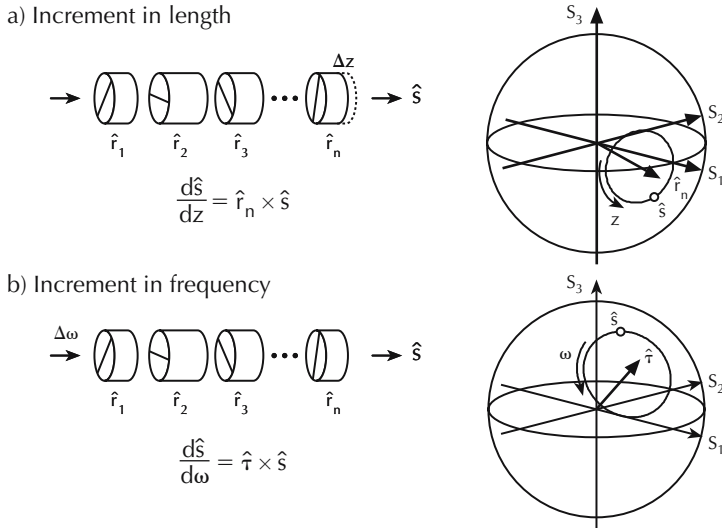


Fig. 8.19. Comparison of infinitesimal rotation for changes in length and frequency. a) Increment in length of the last element. The output polarization precesses about the birefringent axis of the last element. b) Increment in frequency. The output polarization precesses about the principal-state axis of the system.

convolution of the signal pulse with the FIR response of the cascade determines the shape of the output signal. Convolution in time is multiplication in frequency, and the frequency response of PMD has already been heuristically constructed.

To conclude this primer, Fig. 8.18 shows the analogy between the eigenstate system for a single birefringent element and the principal-state system for a birefringent cascade. Figure 8.19 illustrates the two infinitesimal rotation laws that govern PMD, one for change in length and the other for change in frequency. These rotation laws are derived and applied in the next sections in a more rigorous manner.

8.2.2 Fundamental Derivations

The output principal state vector is that vector which is stationary to first order in frequency. The principal state vector is well defined for a lossless birefringent concatenation of any length and composition. The issues at hand are to derive an equation whose eigenvectors are the principal states of the system and whose eigenvalues are the differential group delays along the PSP axes. Together the eigenvectors and values are used to define the PMD vector. This and the following section follows the spin-vector-based derivations set forth by Frigo [16], Gisin [19], and so well elucidated by Gordon and Kogelnik [20].

First, a remark about the analytic tools available for these PMD studies. The introduction of this chapter showed that while a Hermitian matrix fills all 16 entries of the corresponding Mueller matrix, a traceless Hermitian matrix fills only the lower right-hand 3×3 sub-matrix of the Mueller matrix. It was also shown in §2.1 that a unitary matrix also fills only the lower right-hand 3×3 sub-matrix of the corresponding Mueller matrix. Comparison of these matrices shows

$$\underline{H} \mapsto \mathbf{M}_{\underline{H}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix}, \text{ and } U \mapsto \mathbf{M}_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \end{pmatrix} \quad (8.2.5)$$

It will be shown that the PMD operator is an identically traceless Hermitian operator. In Stokes space, PMD represents a vector having a length and pointing direction. Unitary matrices act on PMD matrices without scattering Mueller entries outside of the lower 3×3 sub-matrix. The action of a unitary matrix is to rotate the PMD matrix in Stokes space through similarity transforms: $\underline{H}_T = U \underline{H} U^\dagger$. A complete representation of vectors and rotational transformations can therefore be described solely within the $O(3)$ group.

The Hermitian PMD operator and its traceless property is now derived. Table 8.2 lists the symbols used to describe PMD. An output polarization

state is related to the input polarization state via the system's transformation matrix T . Recall that for a lossless system, $T = \exp(-j\phi_o)U$, where U is a unitary matrix. An arbitrary input is transformed at the output as

$$|t\rangle = e^{-j\phi_o}U|s\rangle \quad (8.2.6)$$

The frequency derivative is

$$|t\rangle_\omega = (-j\phi'_o U + U_\omega) e^{-j\phi_o} |s\rangle$$

Substitution of (8.2.6) into the frequency derivative yields a expression that relates the output polarization change in frequency to the output polarization itself:

$$|t\rangle_\omega = -j(\phi'_o + jU_\omega U^\dagger)|t\rangle \quad (8.2.7)$$

The derivative of the common phase is a delay, that is, $\phi'_o = \tau_o$. Physically this is the common delay experienced by both polarization states. By analogy one expects $jU_\omega U^\dagger$ to have units of delay as well.

To proceed further, the properties of $jU_\omega U^\dagger$ must be determined. Recall that a unitary matrix is defined as $UU^\dagger = U^\dagger U = I$. Taking the frequency derivative and multiplying both sides by j yields

$$jU_\omega U^\dagger = -jUU_\omega^\dagger$$

Notice that

$$(jU_\omega U^\dagger)^\dagger = -jUU_\omega^\dagger$$

Therefore $jU_\omega U^\dagger$ is Hermitian: its eigenvalues are real. The question is whether this Hermitian operator has trace or not. The answer is that $jU_\omega U^\dagger$ is identically traceless. One way to show this is by Taylor expansion. Given that the determinant of a unitary operator is +1 for all frequencies, one has

$$\det(U(\omega)) = \det(U(\omega + \delta\omega)) = +1$$

The determinant of the Taylor expansion of $U(\omega + \delta\omega)$ must be unity to first order in frequency. For $U(\omega + \delta\omega) \simeq U(\omega) + \delta\omega U_\omega$,

$$\det(U(\omega + \delta\omega)) = \det(I + \delta\omega U_\omega U^\dagger) \det(U)$$

In general, the determinant of a matrix A plus the identity matrix yields

$$\det(I + A) = 1 + \text{Tr}(A) + \det(A)$$

For the present case,

$$\det(I + \delta\omega U_\omega U^\dagger) = 1 + \text{Tr}(U_\omega U^\dagger)\delta\omega + \det(U_\omega U^\dagger)(\delta\omega)^2$$

The coefficient of the determinant is second order in frequency while that of the trace is first-order in frequency. In order for the determinant of $U(\omega + \delta\omega)$ to be unity to first-order, the trace must vanish. Therefore

Table 8.2. Symbol Definitions for PMD

Symbol	Expression	Definition
$\vec{\tau} :$	$\vec{\tau} = \tau \hat{p}$	PMD vector in Stokes space. Length τ , pointing direction $\hat{p}$
$\tau :$	$\tau = \partial\varphi/\partial\omega$	Differential-group delay (DGD), $\tau \geq 0$
$\hat{p}, \hat{\tau} :$		Pointing direction of PMD vector, called principal state of polarization (PSP). PSP $\hat{p}$ is defined at output and aligned to slow PSP axis
$\varphi :$	$\varphi = \omega\Delta nL/c$	Birefringent phase
	$\varphi = \omega\tau$	As function of DGD
	$\varphi = \beta L$	As function of local birefringence β
$\vec{\tau}_n :$		Local PMD vector of n^{th} element
$\vec{\tau}(n) :$		Cumulative PMD vector through n^{th} element
$\langle\tau\rangle :$	$\langle \vec{\tau}(\omega) \rangle$	Mean PMD. Frequency average of DGD
$\langle\tau^2\rangle :$	$\langle\vec{\tau}(\omega) \cdot \vec{\tau}(\omega)\rangle$	Mean-square PMD. Frequency average of DGD ²

$$\text{Tr} (jU_\omega U^\dagger) = 0 \quad (8.2.8)$$

A Hermitian matrix with zero trace has important properties. First, its eigenvalues are equal in magnitude and opposite in sign. Second, the SU(2) matrix is equivalent to a vector in O(3). That vector is determined by the eigenvalue equation for $jU_\omega U^\dagger$:

$$jU_\omega U^\dagger |p_\pm\rangle = \pm \lambda |p_\pm\rangle \quad (8.2.9)$$

where $|p_\pm\rangle$ are the eigenvectors of $jU_\omega U^\dagger$. The eigenvalues $\pm\lambda$ are defined as $\lambda \equiv \pm\tau/2$, and thus

$$jU_\omega U^\dagger |p_\pm\rangle = \pm \tau/2 |p_\pm\rangle \quad (8.2.10)$$

Since the determinant is the product of the eigenvalues, $\det(jU_\omega U^\dagger) = -\tau^2/4$. Moreover, the determinant of a product of matrices is the product of the determinants, so the eigenvalues are

$$\tau = 2\sqrt{\det U_\omega} \quad (8.2.11)$$

Combining the eigenvalue equation (8.2.9) with (8.2.7), and choosing an output polarization along an eigenvector of $jU_\omega U^\dagger$, one finds

$$\begin{aligned} |p_\pm\rangle_\omega &= -j (\phi'_o + jU_\omega U^\dagger) |p_\pm\rangle \\ &= -j (\tau_o \pm \tau/2) |p_\pm\rangle \end{aligned}$$

The total group delay of the signals along the principal state axes is

$$\tau_g = \tau_o \pm \tau/2 \quad (8.2.12)$$

where τ_o is the common delay and $\pm\tau/2$ is the differential delay. The slow principal state is $|p_+\rangle$, with corresponding delay $\tau_o + \tau/2$, while the fast principal state is $|p_-\rangle$, with corresponding delay $\tau_o - \tau/2$. In the following, $|p_+\rangle$ is denoted simply by $|p\rangle$.

To verify that the output polarization $|p\rangle$ is stationary to first order for either principal state, the Jones vector is converted to a Stokes vector as

$$\begin{aligned} \vec{p}_\omega &= \langle p_\omega | \vec{\sigma} | p_\omega \rangle \\ &= j\tau_g \langle p | \vec{\sigma} | p \rangle - j\tau_g \langle p | \vec{\sigma} | p \rangle \end{aligned}$$

which is evidently zero.

In summary, the eigenvectors of $jU_\omega U^\dagger$ are the principal states of polarization of the system, and the eigenvalues are the differential group delays along the two axes. The PMD vector is defined to point along the slow output principal state and have a length of the total differential-group delay:

$$\vec{\tau} \equiv \tau \hat{p}, \quad \hat{p} \text{ is slow output PSP} \quad (8.2.13)$$

The Stokes vector interpretation of (8.2.13) is illustrated in Fig. 8.13(a).

8.2.3 Connection Between Jones and Stokes Space

That $jU_\omega U^\dagger$ is traceless Hermitian and has zero trace implies a connection between the $SU(2)$ Jones space and the $O(3)$ Stokes space (cf. §2.6). In particular, observe that for a Stokes vector $\vec{\tau}$, the following two eigenvalue equations are equal to within a factor of two:

$$\begin{aligned} jU_\omega U^\dagger |p\rangle &= (\tau/2) |p\rangle \\ (\vec{\tau} \cdot \vec{\sigma}) |p\rangle &= \tau |p\rangle \end{aligned}$$

The spin-vector operator and $jU_\omega U^\dagger$ are clearly related:

$$jU_\omega U^\dagger = \frac{1}{2} (\vec{\tau} \cdot \vec{\sigma}) \quad (8.2.14)$$

The spin-vector can be used to determine how an arbitrary output polarization state changes with frequency. In Stokes space, the frequency change is

$$\hat{t}_\omega = (\langle t |_\omega) \vec{\sigma} | t \rangle + \langle t | \vec{\sigma} (| t \rangle_\omega) \quad (8.2.15)$$

Employing the spin-vector form (8.2.14) and the eigenvalue equation (8.2.7), (8.2.15) generates

$$\begin{aligned}
\hat{t}_\omega &= j \langle t | \left(\tau_o + \frac{1}{2} (\vec{\tau} \cdot \vec{\sigma}) \right) \vec{\sigma} | t \rangle - j \langle t | \vec{\sigma} \left(\tau_o + \frac{1}{2} (\vec{\tau} \cdot \vec{\sigma}) \right) | t \rangle \\
&= \frac{j}{2} \langle t | (\vec{\tau} \cdot \vec{\sigma}) \vec{\sigma} - \vec{\sigma} (\vec{\tau} \cdot \vec{\sigma}) | t \rangle \\
&= \langle t | \vec{\tau} \times \vec{\sigma} | t \rangle
\end{aligned}$$

and thus

$$\frac{d\hat{t}}{d\omega} = \vec{\tau} \times \hat{t} \quad (8.2.16)$$

This is the infinitesimal frequency-change rule for an arbitrary output polarization state. The output state precesses about the PMD vector $\vec{\tau}$. Only if the output state is aligned or anti-aligned along $\vec{\tau}$ will its state not change with frequency. Illustrations of the precession rule in frequency are given by Figs. 8.11 and 8.19(b).

The PMD vector $\vec{\tau}$ is defined at the output of the system. There is a corresponding PMD vector at the input of the system. Denote $\vec{\tau}_t$ and $\vec{\tau}_s$ as the output and input PMD vectors, respectively. Since in general the output density matrix D_t of a unitary transformation is related to the input density D_s via $D_t = U D_s U^\dagger$, and the density operator is related to the spin-vector through (2.5.29) on page 56, the relation between input and output spin vectors is

$$(\vec{\tau}_t \cdot \vec{\sigma}) = U (\vec{\tau}_s \cdot \vec{\sigma}) U^\dagger$$

Isolating $(\vec{\tau}_s \cdot \vec{\sigma})$ and substitution of (8.2.14) yields

$$\frac{1}{2} (\vec{\tau}_s \cdot \vec{\sigma}) = j U^\dagger U_\omega \quad (8.2.17)$$

It makes sense that the input PMD vector is determined by writing the unitary matrices U_ω and $U^\dagger$ in reverse order to that for the output PMD vector. Moreover, as the operators are unitary, the DGD for both the input and output PMD vectors is identical.

For the unitary transformation in Jones space $|t\rangle = T|s\rangle$, the equivalent transformation in Stokes space is

$$\hat{t} = R \hat{s} \quad (8.2.18)$$

where a vector form of the unitary operator R is given in (2.6.25) on page 68, repeated here due to its significance:

$$R = (\hat{r} \hat{r} \cdot) + \sin \varphi (\hat{r} \times) - \cos \varphi (\hat{r} \times) (\hat{r} \times) \quad (8.2.19)$$

The precession vector $\hat{r}$ of operator R points along an eigenvector of U and the precession angle $\varphi = \omega \tau$ is the angle of the complex eigenvalue of U .

A expression analogous to $j U_\omega U^\dagger$ exists in Stokes space. Consider a frequency change for (8.2.18):

$$\hat{t}_\omega = R_\omega \hat{s}$$

where the input polarization state is fixed in frequency. Substitution of $\hat{s} = R^\dagger \hat{t}$ gives

$$\hat{t}_\omega = R_\omega R^\dagger \hat{t}$$

By comparison with the previously derived infinitesimal rotation rule (8.2.16) one is led to the identification that

$$\vec{\tau} \times = R_\omega R^\dagger \quad (8.2.20)$$

The value of $R_\omega R^\dagger$ will be clear when deriving the PMD concatenation rules. Yet even at this point its use is significant. Consider again the input and output PMD vectors $\vec{\tau}_s$ and $\vec{\tau}_t$. It was already determined by (8.2.17) that the lengths of these two vectors are identical. Since in general $\hat{t} = R\hat{s}$, one can choose the input and output polarizations parallel to the PSP's of $jU_\omega U^\dagger$. Accordingly,

$$\vec{\tau}_t = R \vec{\tau}_s \quad (8.2.21)$$

That is, the input and output PMD vectors are related by R . How is the second-order PMD component, $\vec{\tau}_{t\omega}$, related to $\vec{\tau}_{s\omega}$? Taking the frequency derivative of (8.2.21),

$$\vec{\tau}_{t\omega} = R_\omega \vec{\tau}_s + R \vec{\tau}_{s\omega}$$

along with the substitution of (8.2.21) and (8.2.20) yields

$$\begin{aligned} \vec{\tau}_{t\omega} &= \vec{\tau} \times \vec{\tau} + R \vec{\tau}_{s\omega} \\ &= R \vec{\tau}_{s\omega} \end{aligned} \quad (8.2.22)$$

Both the first- and second-order PMD vectors transform from input to output through R . Higher-order frequency derivatives quickly become more complex.

The spin-vector form ($\vec{\tau} \cdot \vec{\sigma}$) also assists in the evaluation of the three Stokes components of the vector $\vec{\tau}$. In particular, (2.5.29) on page 56 gives

$$\vec{\tau} \cdot \vec{\sigma} = \begin{pmatrix} \tau_1 & \tau_2 - j\tau_3 \\ \tau_2 + j\tau_3 & -\tau_1 \end{pmatrix} \quad (8.2.23)$$

Likewise, (8.2.20) is identified with (2.6.29) on page 70 to gives

$$\vec{\tau} \times = \begin{pmatrix} 0 & -\tau_3 & \tau_2 \\ \tau_3 & 0 & -\tau_1 \\ -\tau_2 & \tau_1 & 0 \end{pmatrix} \quad (8.2.24)$$

When the unitary matrix is written in Cayley-Klein form, given by (2.4.28) on page 51, the matrix entries of $jU_\omega U^\dagger$ are identified with (8.2.23) as

$$\tau_1 = 2j(a_\omega a_\omega^* + b_\omega b_\omega^*) \quad (8.2.25a)$$

$$\tau_2 = 2 \Im m(a_\omega b - b_\omega a) \quad (8.2.25b)$$

$$\tau_3 = 2 \Re e(a_\omega b - b_\omega a) \quad (8.2.25c)$$

The DGD τ can be determined either from $\tau^2 = \tau_1^2 + \tau_2^2 + \tau_3^2$ or from (8.2.11). In either case the resulting expression is

$$\tau = 2\sqrt{a_\omega a_\omega^* + b_\omega b_\omega^*} \quad (8.2.26)$$

In comparison with (8.2.26) it is interesting to note that $\sqrt{aa^* + bb^*} = 1$. The frequency derivative of the Jones matrix elements brings down group-delay coefficients which are responsible for the differential-group delay of the system.

Finally, the spin-vector form is used to relate the eigenvector of U to the output PMD vector. The exponential form of the unitary operator

$$U = \exp[-j\varphi(\hat{r} \cdot \vec{\sigma})/2] \quad (8.2.27)$$

has a frequency derivative of (cf. (2.6.19) on page 68)

$$U_\omega = -j(\varphi_\omega/2)(\hat{r} \cdot \vec{\sigma})U - j(\hat{r}_\omega \cdot \vec{\sigma})\sin(\varphi/2)$$

The product $jU_\omega U^\dagger$ is

$$jU_\omega U^\dagger = \varphi_\omega/2(\hat{r} \cdot \vec{\sigma}) + (\hat{r}_\omega \cdot \vec{\sigma})\sin(\varphi/2)[I\cos(\varphi/2) + j(\hat{r} \cdot \vec{\sigma})\sin(\varphi/2)]$$

Identification with (8.2.14) and use of spin-vector product identity (2.5.38) on page 57 generates the relation between the eigenvector $\hat{r}$ of U and the PMD vector $\vec{\tau}$ of $jU_\omega U^\dagger$:

$$\vec{\tau} = \varphi_\omega \hat{r} + \sin \varphi \hat{r}_\omega - (1 - \cos \varphi) \hat{r}_\omega \times \hat{r} \quad (8.2.28)$$

As discussed in §8.2.1, the eigenvector of U generally changes to first order with frequency. That first-order effect is accounted for by $\hat{r}_\omega$ in (8.2.28). An important special case is when the PMD vector refers to a single homogeneous birefringent section. In this case $\hat{r}_\omega = 0$ and the PMD vector of the element is aligned to the birefringent axis: $\vec{\tau} = \varphi_\omega \hat{r}$. The frequency derivative of the phase is the group delay of the element, or $\varphi_\omega = \tau$.

8.2.4 Concatenation Rules for PMD

The PMD concatenation rules are very helpful to gain physical insight and intuition on how PMD behaves. The concatenation rules track the PMD vectors themselves as more and more sections are added to a system. Early work on PMD statistics clearly deals with the concatenation of many (indeed infinitely many) birefringent sections, but the discussion of PMD statistics is left for the next chapter. The four works from the technical literature that present concatenation rules are from Curti *et al.* [4], who presented a two-stage concatenation and numerically extended the work to large numbers; Gisin [19], who derived the recurrence relation for an arbitrary number of sections; Karlsson [38], who recast Gisin's recurrence relation in spin-vector form; and Gordon and Kogelnik [20], who extended the prior work.

The concatenation rules are rules for adding two or more output PMD vectors. The rules could be written for the input PMD vectors just as well, but they are not for a mixture of input and output vectors.

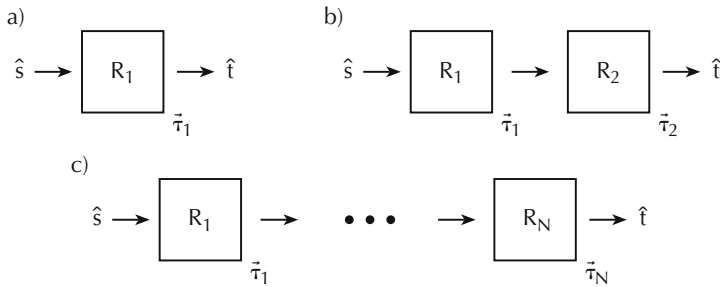


Fig. 8.20. Block diagrams of one, two, and N birefringent blocks in concatenation. It is assumed that the blocks are lossless. A birefringent block may have heterogeneous or homogeneous birefringence.

One Birefringent Section

Without even one section of birefringence there is no PMD at all. The PMD vector first appears after a single homogeneous birefringent section. The relation between a single birefringent element and the PMD vector is given in (8.2.28), where $\hat{r}_\omega = 0$. Thus,

$$\vec{\tau} = \tau \hat{r} \quad (8.2.29)$$

where $\hat{r}$ points in the direction of the slow birefringent axis of the element. The DGD is given by $\tau = \varphi_\omega$. $\vec{\tau}$ is the first-order PMD vector for the section. The second order vector is

$$\vec{\tau}_\omega = 0 \quad (8.2.30)$$

There is no second-order frequency dependence of the PMD vector. This is a unique case for PMD, as concatenations with two or more stages will always have finite $\vec{\tau}_\omega$ except, possibly, at particular frequencies when the second-order vector momentarily vanishes.

One Birefringent Block

A birefringent block is distinguished from a birefringent section in that the latter is homogeneous birefringence, i.e. with no intermediate polarization mode coupling, while the former can be birefringence of any composition and inhomogeneity. Any output polarization is related to the input polarization by $\hat{t} = R_1 \hat{s}$, while the first- and second-order PMD of the block is denoted simply by $\vec{\tau}_1$ and $\vec{\tau}_{1\omega}$ (Fig. 8.20(a)). The PMD vector can be determined from $\vec{\tau}_1 \times = R_{1\omega} R_1^\dagger$.

Two Birefringent Blocks

Denote the PMD vector generated by a first birefringent block as $\vec{\tau}_1$ and by a second birefringent block as $\vec{\tau}_2$ (Fig. 8.20(b)). When these two blocks are

concatenated the overall input-to-output transformation is $R = R_2 R_1$. The cumulative PMD vector $\vec{\tau}$ is related to the block transfer operators as

$$\begin{aligned}\vec{\tau} \times &= (R_2 R_1)_\omega (R_2 R_1)^\dagger \\ &= R_{2\omega} R_1 R_1^\dagger R_2^\dagger + R_2 R_{1\omega} R_1^\dagger R_2^\dagger \\ &= \vec{\tau}_2 \times + R_2 (\vec{\tau}_1 \times) R_2^\dagger \\ &= \vec{\tau}_2 \times + (R_2 \vec{\tau}_1) \times\end{aligned}$$

The last line is derived from the preceding one as $RvR^\dagger$ is a unitary transformation on v . The embedded expression for first-order PMD is

$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1 \quad (8.2.31)$$

The second-order expression comes from the frequency derivative

$$\vec{\tau}_\omega = \vec{\tau}_{2\omega} + R_2 \vec{\tau}_{1\omega} + R_{2\omega} \vec{\tau}_1$$

Using (8.2.31) to write $\vec{\tau}_1 = R_2^\dagger (\vec{\tau} - \vec{\tau}_2)$, the second-order expression reduces to

$$\vec{\tau}_\omega = \vec{\tau}_{2\omega} + R_2 \vec{\tau}_{1\omega} + \vec{\tau}_2 \times \vec{\tau} \quad (8.2.32)$$

The second-order expression for $\vec{\tau}_\omega$ is almost like that for $\vec{\tau}$ except for the additional $\vec{\tau}_2 \times \vec{\tau}$ on the right-hand side. The additional vector generates a pulling of the second-order cumulative vector $\vec{\tau}_\omega$ in a direction defined by $\vec{\tau}_2 \times \vec{\tau}$, which is orthogonal to both the local PMD component $\vec{\tau}_2$ and the cumulative first-order vector $\vec{\tau}$.

Since these derivations have applied to heterogeneous birefringent blocks, the block PMD vectors and the transformation operators R are generally a function of frequency. Accordingly, the first-order cumulative vector is more accurately written as

$$\vec{\tau}(\omega) = \vec{\tau}_2(\omega) + R_2(\omega) \vec{\tau}_1(\omega) \quad (8.2.33)$$

For small frequency changes, (8.2.33) establishes a precession rule. The PMD vector $\vec{\tau}_1$ precesses about the axis $\hat{r}_2$ of the second PMD block through angle φ_2 . This is the effect of R_2 on $\vec{\tau}_1$. The rotated vector is then added to $\vec{\tau}_2$. As Gordon and Kogelnik point out, “The rule is very similar to that for impedances of a transmission line: to get the PMD vector of an assembly, transform the PMD vectors of each individual section to a common reference cross section and take the sum of all the vectors” [20].

A significant special case of (8.2.33) is when R_2 and R_1 refer to homogeneous birefringent sections. In this case there is no frequency dependence of $\hat{r}_{1,2}$ or $\vec{\tau}_{1,2}$, and the precession behavior is more clear since the birefringent axes are fixed. The two-section precession rule (8.2.31) is expanded to

$$\vec{\tau} = \vec{\tau}_2 + \left[(\hat{r}_2 \hat{r}_2 \cdot) + \sin \varphi_2 (\hat{r}_2 \times) - \cos \varphi_2 (\hat{r}_2 \times) (\hat{r}_2 \times) \right] \vec{\tau}_1$$

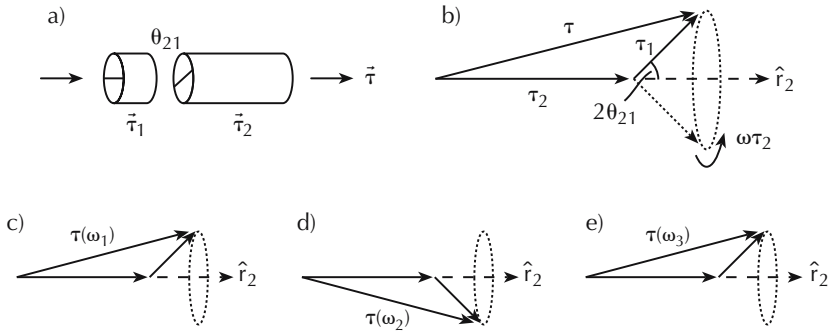


Fig. 8.21. PMD concatenation rule for two homogeneous birefringent sections. a) Concatenation of two birefringent sections having section DGD's of τ_1 and τ_2 , and relative angle between birefringent axes θ_{21} . b) PMD vector addition. $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with retardance $\omega\tau_2$. The cumulative PMD vector $\vec{\tau}$ is the vector sum of the components. The pointing direction of $\vec{\tau}$ is the output PSP of the cascade. c-e) Component PMD vectors at three different frequencies. The PSP spectrum is periodic and the DGD spectrum is constant.

where $\hat{r}_2$ points in the direction of $\vec{\tau}_2$ and φ_2 is the birefringent phase of the second section. This motion and the associated physical construct is illustrated in Fig. 8.21(a,b).

Two homogeneous birefringent sections with mode-mixing at a well-defined junction are illustrated in Fig. 8.21(a). The mode-mixing angle is the difference between the angles of the birefringent axes of the two sections. The two-section concatenation rule is illustrated in Fig. 8.21(b), which is drawn in Stokes space. The base of $\vec{\tau}_1$ is jointed to the tip of $\vec{\tau}_2$. The angle between the two vectors is $2\theta_{21}$, twice the physical angle at the mode-mixing junction; this angle is fixed in frequency. When frequency is changed, the PMD vector of the first section $\vec{\tau}_1$ precesses about the axis of the second PMD vector, $\hat{r}_2$. The angle of precession is $\omega\tau_2$, which is solely dictated by the length of the second PMD vector. The precession rate is τ_2 . Over all frequencies the tip of $\vec{\tau}_1$ traces a circle; the circular motion is periodic with a free-spectral range of $\text{FSR} = 1/\tau_2$. Figures 8.21(c-e) illustrate this motion. At a first frequency the cumulative PMD vector $\vec{\tau}$ points upward; at a second frequency it points downward; and at a third frequency it points up again. Since $\vec{\tau}$ points in the direction of the slow output PSP, the circle traced by $\vec{\tau}_1$ in frequency is the PSP spectrum of the two-section concatenation. The length of the cumulative vector $\vec{\tau}$ is the vector sum of the components. In this case there are only two fixed-length components, so $\vec{\tau}$ completes the triangle rule and remains constant in length over frequency.

N Birefringent Blocks

The concatenation rule for N birefringent blocks is derived from repeated application of (8.2.31) and (8.2.32). Denoting $\vec{\tau}_k$ the PMD vector of the k^{th} block and $\vec{\tau}(k)$ the cumulative PMD vector through the k^{th} block, the cumulative first- and second-order PMD vectors are boot-strapped from the origin

$$\begin{aligned} \vec{\tau}(1) &= \vec{\tau}_1 & \vec{\tau}_\omega(1) &= 0 \\ \vec{\tau}(2) &= \vec{\tau}_2 + R_2 \vec{\tau}(1) & \vec{\tau}_\omega(2) &= \vec{\tau}_2 \times \vec{\tau}(2) \\ \vec{\tau}(3) &= \vec{\tau}_3 + R_3 \vec{\tau}(2) & \vec{\tau}_\omega(3) &= \vec{\tau}_3 \times \vec{\tau}(3) + R_3 \vec{\tau}_\omega(2) \\ &\vdots & &\vdots \\ \vec{\tau}(n) &= \vec{\tau}_n + R_n \vec{\tau}(n-1) & \vec{\tau}_\omega(n) &= \vec{\tau}_n \times \vec{\tau}(n) + R_n \vec{\tau}_\omega(n-1) \end{aligned} \quad (8.2.34)$$

Another form for $\vec{\tau}(n)$ is to write recursively

$$\vec{\tau}(n) = \vec{\tau}_n + R_n (\vec{\tau}_{n-1} + R_{n-1} (\vec{\tau}_{n-2} \cdots R_3 (\vec{\tau}_2 + R_2 \vec{\tau}_1)) \cdots) \quad (8.2.35)$$

This form gives some physical insight. Starting from the beginning of the cascade, component vector $\vec{\tau}_1$ is placed on the tip of $\vec{\tau}_2$ and precesses about its axis at rate τ_2 . This is the action of R_2 . Together these two components are placed on the tip of $\vec{\tau}_3$ and they precess about the $\hat{r}_3$ axis at rate τ_3 . This is the action of R_3 . Keep in mind that while $\vec{\tau}_2$ and $\vec{\tau}_1$ precess about $\vec{\tau}_3$, $\vec{\tau}_1$ continues to precess about $\vec{\tau}_2$. The procedure is repeated through the n^{th} section.

Figure 8.22 illustrates the precession for three homogeneous birefringent sections in cascade. Unlike the two-section case, the motion of the tip of $\vec{\tau}_1$ is more complicated, tracing a “folded-eight” curve in Stokes space and having a frequency-dependent vector sum. The vector sum $\vec{\tau}$ is the cumulative PMD vector of the cascade.

Finally, a compact form of the cumulative first- and second-order PMD vectors is written as

$$\vec{\tau} = \sum_{k=1}^n R(n, k+1) \vec{\tau}_k \quad (8.2.36)$$

and

$$\vec{\tau}_\omega = \sum_{k=1}^n R(n, k+1) (\vec{\tau}_{n\omega} + \vec{\tau}_n \times \vec{\tau}(n)) \quad (8.2.37)$$

where

$$R(n, k) = R_n R_{n-1} \cdots R_k \quad (8.2.38)$$

Note that the evaluation of $\vec{\tau}_\omega$ requires the concurrent evaluation of $\vec{\tau}$. Expressions (8.2.36-8.2.37) offer a fast way to evaluate numerically the first- and second-order PMD vectors for an arbitrary cascade. The evaluation occurs directly in Stokes space and $\vec{\tau}_\omega$ is determined without a numerical derivative.

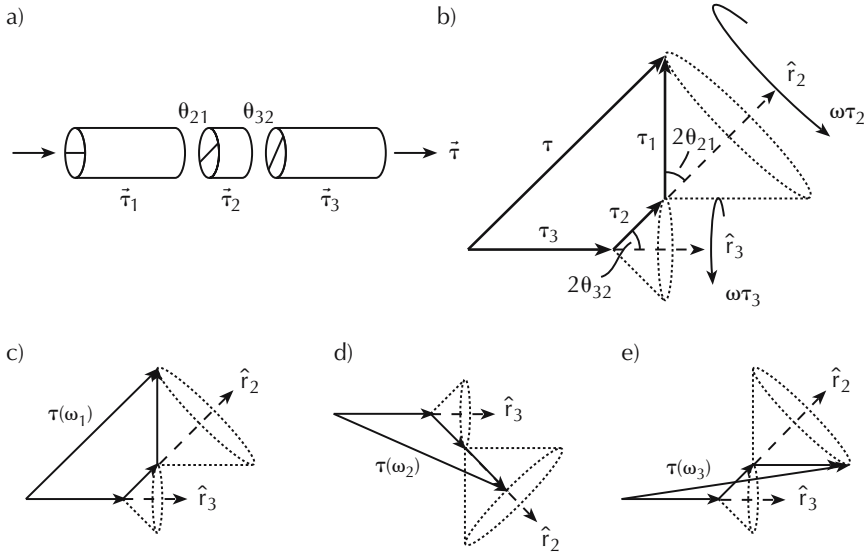


Fig. 8.22. PMD concatenation rule for three homogeneous birefringent sections. a) Concatenation of three birefringent sections having section DGD's of $\tau_{1,2,3}$ and relative angle between birefringent axes θ_{21} and θ_{32} . The drawing is for $\tau_1 = \tau_3 = 2\tau_2$. b) PMD vector addition. $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with retardance $\omega\tau_2$; combined, the vectors precess about $\vec{\tau}_3$ with retardance $\omega\tau_3$. The motion at the tip of $\vec{\tau}_1$ is more complicated than the two-section case, and the length of the cumulative PMD vector now changes with frequency. c-e) Component PMD vectors at three different frequencies. The PSP spectrum is periodic but more complicated than for two sections. The DGD spectrum varies with frequency and is also periodic.

An alternative method, the multiplication of Jones matrices where each matrix represents a homogeneous element and conversion to PMD vectors via (8.2.25-8.2.26), requires substantially more computations and requires evaluation at two frequencies in order to estimate $\vec{\tau}_\omega$.

8.2.5 PMD Evolution Equations

The PMD concatenation equations developed in the preceding abstracted the underlying birefringence and focused solely on how PMD vectors combine in length and precess in frequency. Studies of optical pulse propagation in fiber and statistical properties of PMD require an explicit connection between the cumulative PMD vector and the underlying local birefringence. There are at least two ways to derive the PMD evolution equation, one due to Poole *et al.* [53], and another due to Gisin *et al.* [19] and Gordon *et al.* [20]. The latter derivation type makes the connection between the PMD concatenation rules and the underlying birefringence. The Poole derivation is detailed below.

A note on notation. The PDL evolution equations denote $\vec{\alpha}(z)$ as the local PDL vector $\vec{\alpha}$ per unit length (cf. §8.1.4). For PMD, the local birefringence per unit length is customarily denoted β (as in $\exp(-j\beta z)$ where $\beta = \omega\Delta n/c$). The connection to the local PMD vector $\vec{\tau}_\ell$ is $\vec{\tau}_\ell = \vec{\beta}_\omega L$ where L is the length of the segment.

The Poole derivation starts with the established precession rules in length and frequency for an arbitrary polarization state:

$$\frac{\partial \hat{s}}{\partial z} = \vec{\beta} \times \hat{s}, \quad \text{and} \quad \frac{\partial \hat{s}}{\partial \omega} = \vec{\tau} \times \hat{s}$$

where $\vec{\beta}$ is the local birefringence vector and $\vec{\tau}$ is the cumulative PMD vector up to location z . Taking the frequency derivative of the first equation and the length derivative of the second gives

$$\begin{aligned} \frac{\partial^2 \hat{s}}{\partial z \partial \omega} &= \vec{\beta}_\omega \times \hat{s} + \vec{\beta} \times \hat{s}_\omega \\ \frac{\partial^2 \hat{s}}{\partial \omega \partial z} &= \vec{\tau}_z \times \hat{s} + \vec{\tau} \times \hat{s}_z \end{aligned}$$

Under the assumption of continuity of the function $\hat{s}(z, \omega)$, the left-hand sides are equal. Using the vector identity $(\vec{\beta} \times \vec{\tau}) \times \hat{s} = \vec{\beta} \times (\vec{\tau} \times \hat{s}) - \vec{\tau} \times (\vec{\beta} \times \hat{s})$ results in the PMD evolution equation

$$\frac{\partial \vec{\tau}}{\partial z} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau} \quad (8.2.39)$$

The local birefringence changes the cumulative PMD vector in both an additive and multiplicative sense: additive through $\vec{\beta}_\omega$ and multiplicative through $|\beta\tau|$. Moreover, $\vec{\beta}$ drives the average direction of $\vec{\tau}_z$ while $\vec{\beta} \times \vec{\tau}$ drives $\vec{\tau}_z$ in a perpendicular direction.

Figure 8.23 illustrates the motion of $\vec{\tau}$ through two birefringent sections, where $\tau(z=0) = 0$. Through the first section the cross-product vanishes, so $\vec{\tau}$ directly tracks β_1 . However, once the light enters the second section, the local birefringence and cumulative PMD vector are no longer parallel. The motion of the PMD vector is helical about a center axis $\vec{\beta}_2$. The $\vec{\beta}_{2\omega}$ term pulls the average direction of $\vec{\tau}$ along $\vec{\beta}_2$ while the $\vec{\beta} \times \vec{\tau}$ term drives the helical motion.

The physical interpretation of the motion of $\vec{\tau}$ in section β_2 is as follows. Recall that the length of the PMD vector is the differential-group delay. The DGD is, roughly, a measure of the number of full-wave slips between orthogonal polarizations that has occurred. For each accumulated full-wave slip along β_2 there is an increment of the DGD by the associated delay. For instance, at $1.55 \mu\text{m}$ a one-wave slip is a delay of about 5 fs. The projection of $\vec{\tau}$ onto the $\vec{\beta}_2$ axis is approximately the number of full-wave phase slips that has occurred in the section. Clearly the longer the section the longer the projected vector.

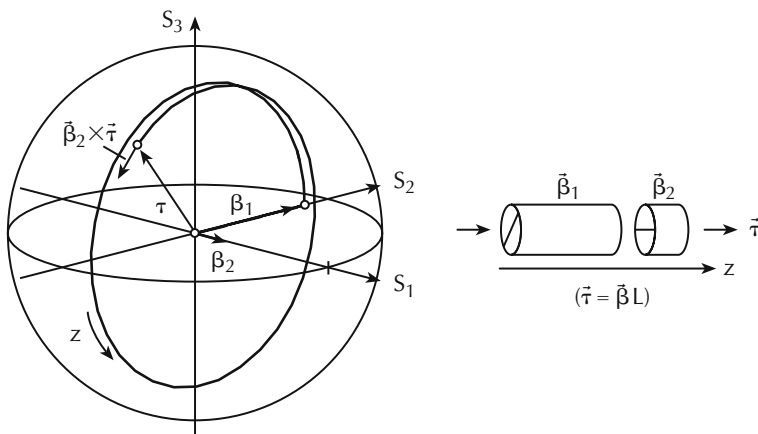


Fig. 8.23. Cumulative PMD vector $\vec{\tau}$ evolution through birefringent sections β_1 and β_2 . The cumulative vector follows β_1 directly. Once the light enters β_2 the cross-product between local and cumulative vectors is non-zero and precession commences. $\vec{\tau}$ traces a helical path in the direction of β_2 . The birefringence $\beta_{2\omega}$ is the driving term of the equation while $\beta_2 \times \vec{\tau}$ forces the cumulative vector in a direction perpendicular to β_2 .

Since $\vec{\tau}$ is a Stokes vector as well, it must track the principal state of polarization as the light propagates. As the phase slips through one full wave the principal state (as well as the polarization state) precesses about $\vec{\beta}_2$. A full revolution is made for length Δz such that

$$\beta_2 \Delta z = 2\pi$$

If the helical motion had zero pitch then the motion of $\vec{\tau}$ about β_2 would be a pure precession. But given that the length of the PMD vector tracks the number of orthogonal-wave slips the helix pitch is greater than zero.

Figure 8.23 shows less than two full revolutions of $\vec{\tau}$ about $\vec{\beta}_2$. This corresponds to less than two full-wave phase slips, or less than 10 fs. A birefringent crystal or fiber segment that has a substantial DGD, say 1 ps, has about 200 revolutions of $\vec{\tau}$ in Stokes space at $1.55 \mu\text{m}$. A 100 ps delay has accordingly some 20,000 revolutions. This is the origin of an order-of-magnitude distinction that is common when dealing with PMD. The DGD would have to be written to an accuracy better than one part in 20,000 in order to capture the fraction of a phase slip a long birefringent concatenation imparts. Yet only two or three significant figures are generally reported. Three significant figures leaves unresolved hundreds of revolutions for a 100 ps delay. It also leaves unresolved the fraction of a phase slip that occurs for even a 1 ps delay. But the fraction of a phase slip is essential information to properly track the PMD vector evolution. The criticality of the fractional phase slip is illustrated by the following example.

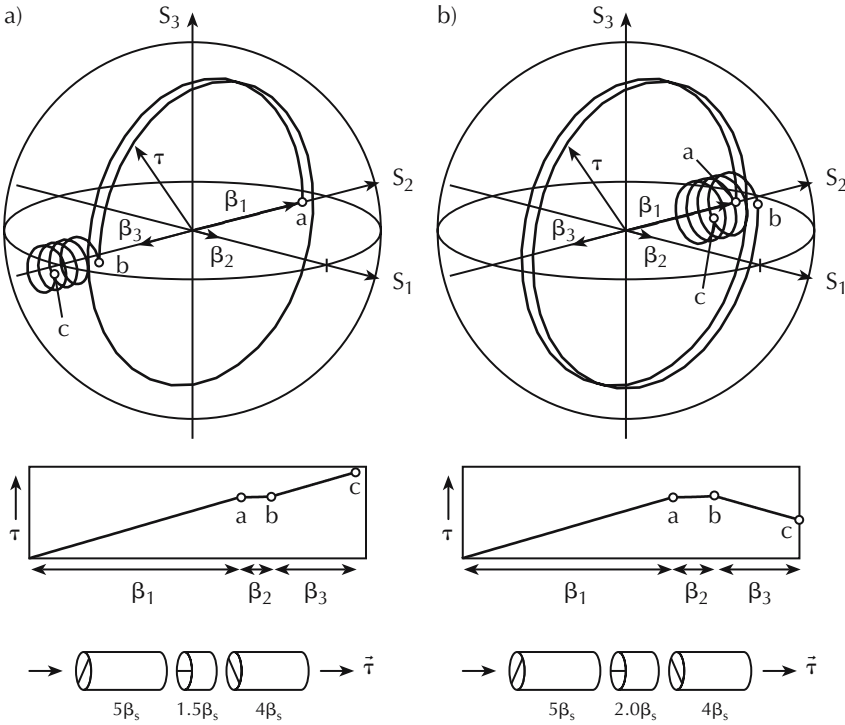


Fig. 8.24. The importance of residual birefringent phase on the evolution of the cumulative PMD vector. The three-segment cascades are identical but for the second segment: in (a) the segment imparts 1.5 revolutions while in (b) the segment imparts 2.0 revolutions. a) Trajectory of three-segment cascade where center segment has 1.5 birefringent phase slips. The cumulative DGD increases monotonically. a) Trajectory of three-segment cascade where center segment has 2.0 birefringent phase slips. The cumulative DGD decreases when it enters the third segment. The output PSP's of the two cascades are nearly in opposite directions.

Figure 8.24 shows the importance of the fractional phase slip, also known as the residual birefringent phase, in the evolution of the PMD vector. The figure illustrates the evolution through two concatenations of three segments each. The concatenations are almost the same; the only difference is that the second segment in Fig. 8.24(a) has 1.5 revolutions while the one in Fig. 8.24(b) has 2.0 revolutions. In the first sequence the PMD vector output from $\vec{\beta}_2$ points in the direction of $\vec{\beta}_3$; the cumulative PMD vector continues to increase in length through the third segment. In the second sequence, the extra half-wave rotation of $\vec{\tau}$ through the second segment orients the resultant PMD vector in the opposite direction from $\vec{\beta}_3$. Propagation through the third segment still pulls $\vec{\tau}$ toward $\vec{\beta}_3$ but does so first by decreasing the PMD vector length. The resultant cumulative vector length and pointing direction are very different

for the two cases; yet the only underlying difference is a half-wave shift along the second birefringent segment. This discussion highlights the importance of the residual birefringent phase; this phase will present itself time and again in the context of the Fourier spectrum of the DGD and programmable PMD generation.

As a point of comparison to PDL evolution, the cumulative PMD vector can point in any direction in Stokes space even if the underlying birefringent vectors of the segments all lie in the same plane. For PDL, however, if the underlying PDL vectors all lie in the same plane, the cumulative PDL vector cannot leave that plane. This is illustrated in Fig. 8.6(c) on page 311. The difference stems from the $-(\vec{\alpha} \cdot \vec{\Gamma})\vec{\Gamma}$ term in the PDL evolution equation which pulls the cumulative PDL vector toward the local PDL vector, while the $\vec{\beta} \times \vec{\tau}$ term in the PMD evolution equation drives the cumulative PMD vector in a helical motion about the local birefringence. The helical motion is in a plane nearly orthogonal to the local birefringent vector (it has a small longitudinal component along the local vector). Hence the cumulative PMD vector will likely lie off of a plane of the local birefringent vectors.

8.2.6 Time-Domain Representation

The predominant representation of PMD thus far has been in the frequency domain. This is a natural consequence of the frequency-centric definition of the PMD vector. However, the parallel representation is the PMD impulse response in the time domain. The impulse response is generally a more difficult characteristic with which to make computations, but a richness of intuition is gained by understanding the parallels between the two domains.

Between the extremes of sine-wave response and impulse response lies the signal response of a communications channel, particularly the distortion imparted on a signal due to PMD. The signal response is fundamentally the convolution of the input waveform with the impulse response. What makes the calculation tricky is that co-polarized signal-image components that result from the convolution interfere coherently; the temporal location of the impulses matter to within a fraction of a wave. When in one case two co-polarized signal images are in phase and add constructively, a dephasing by π leads to destructive interference between the signal images. While the impulse weights change with changes in mode mixing, the temporal locations of the impulse response change only when the composition of the PMD concatenation changes. This implies that for a specific concatenation, the impulse response may extend well into the duration of a signal pulse; how the signal is distorted depends on which co-polarized signal images make constructive interference and which make destructive interference. In some cases the signal will look undistorted while in others it will look quite distorted. How the PMD impacts the signal depends on the expression of this coherent interference.

The following three subsections present three views of the time-domain response of a signal to PMD: simple distortions, signal-moments analysis, and impulse-moments analysis.

Signal Distortion

The impact of PMD on a signal is calculated by multiplying the waveform spectrum with the PMD spectrum or convolving the waveform with the PMD impulse response. Either way a reconciliation must be made as its natural to describe the signal waveform in the time domain and PMD in the frequency domain.

In general, the time- and frequency-domain representations of the input signal are

$$e_s(t) = f(t) |s\rangle, \quad \text{and} \quad E_s(\omega) = f(\omega) |s\rangle \quad (8.2.40)$$

where $f(t)$ is the waveform envelope of the electric field $e_s(t)$ and $|s\rangle$ is its input polarization state; and where $E_s(\omega)$ and $f(\omega)$ are their Fourier transform equivalents³.

The output electric field is related to the input via the PMD transformation matrix $T(\omega)$:

$$\begin{aligned} E_t(\omega) &= T(\omega) E_s(\omega) \\ &= f(\omega) T(\omega) |s\rangle \end{aligned} \quad (8.2.41)$$

where $f(\omega)$ is a complex-valued function of frequency and $T(\omega)$ is an operator. In the time domain, the output waveform envelope is governed by the polarization transfer function

$$e_t(t) = \begin{pmatrix} f(t) * t_{11}(t) & f(t) * t_{12}(t) \\ f(t) * t_{21}(t) & f(t) * t_{22}(t) \end{pmatrix} |s\rangle \quad (8.2.42)$$

where $t_{ij}(t)$ is the inverse Fourier transform of the respective matrix element in $T(\omega)$. It is important to recognize that the waveform envelope and

³ The Fourier transform pair used herein follows Haus [22]:

$$\begin{aligned} e(t) &= \int_{\mathbb{R}} E(\omega) e^{j\omega t} d\omega, \quad E(\omega) = \frac{1}{2\pi} \int_{\mathbb{R}} e(t) e^{-j\omega t} dt \\ W &= \int_{\mathbb{R}} e^\dagger(t) e(t) dt = 2\pi \int_{\mathbb{R}} E^\dagger(\omega) E(\omega) d\omega \end{aligned}$$

where the last equation is Parseval's theorem. Dirac delta functions are defined by the following integrals:

$$2\pi \delta(t - t') = \int_{\mathbb{R}} e^{j\omega(t-t')} d\omega, \quad \text{and} \quad 2\pi \delta(\omega - \omega') = \int_{\mathbb{R}} e^{j(\omega-\omega')t} dt$$

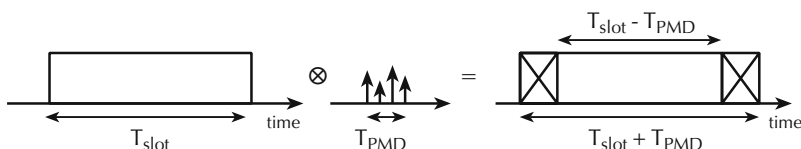


Fig. 8.25. Polarization-independent convolution of waveform envelope with PMD impulse response. The convolution broadens the output signal by T_{PMD} and reduces the duration of the undistorted portion of the signal to $T_{\text{slot}} - T_{\text{PMD}}$. The details of the distortion in the transition regions depend on the interference of co-polarized signal images. These details can be calculated only when the precise nature of the PMD and signal characteristic are known.

polarization state of the output field are not necessarily separable; that is, $e_t(t) = f_t(t) |t\rangle$ is not necessarily true. Rather, the waveform and polarization state are entangled. Such is the effect of depolarization: while each spectral component has a distinct polarization state, the inverse Fourier transform into the time domain folds the polarization states together so that on every time interval multiple polarization states exist; the degree of polarization is accordingly reduced (cf. §1.5.3).

The elements that affect the output waveform are present in (8.2.42): the input state $|s\rangle$, the impulse response of the PMD $t_{ij}(t)$, and its convolution with the waveform. It is hard to make generalizations about a system than can be arbitrarily complex. Instead, the following presents a few case studies to show different aspects of general PMD, first-order PMD, second-order PMD, and higher-order PMD.

- For an arbitrary birefringent concatenation with low overall PMD, signal distortion starts at the edges. Figure 8.25 illustrates schematically the effect of a signal “one” convolved with a simple PMD impulse response; the illustration is polarization independent but two of the four impulses are orthogonal to the others. Pulse images in the same polarization state coherently combine either constructively or destructively depending on the relative phase of the corresponding impulses. The regions of these combinations are at either transition edge of the signal. The temporal extent of the transition regions equals the width of the PMD impulse response T_{PMD} . Due to the convolution, the overall signal duration is increased by T_{PMD} and the duration of the center part of the signal is $T_{\text{slot}} - T_{\text{PMD}}$.

When there is a large number of impulses, and there are 2^N impulses for N birefringent segments, the gaussian distribution of impulses broadens the transition edges by an amount related to the standard deviation of the distribution. The calculation by Gisin and Pellaux [19], detailed in the third subsection, shows that the standard deviation of the impulse response equals the mean DGD.

- First-order PMD is distinct from all other forms because the output waveform is the sum of two identically shaped, orthogonally polarized, time-shifted

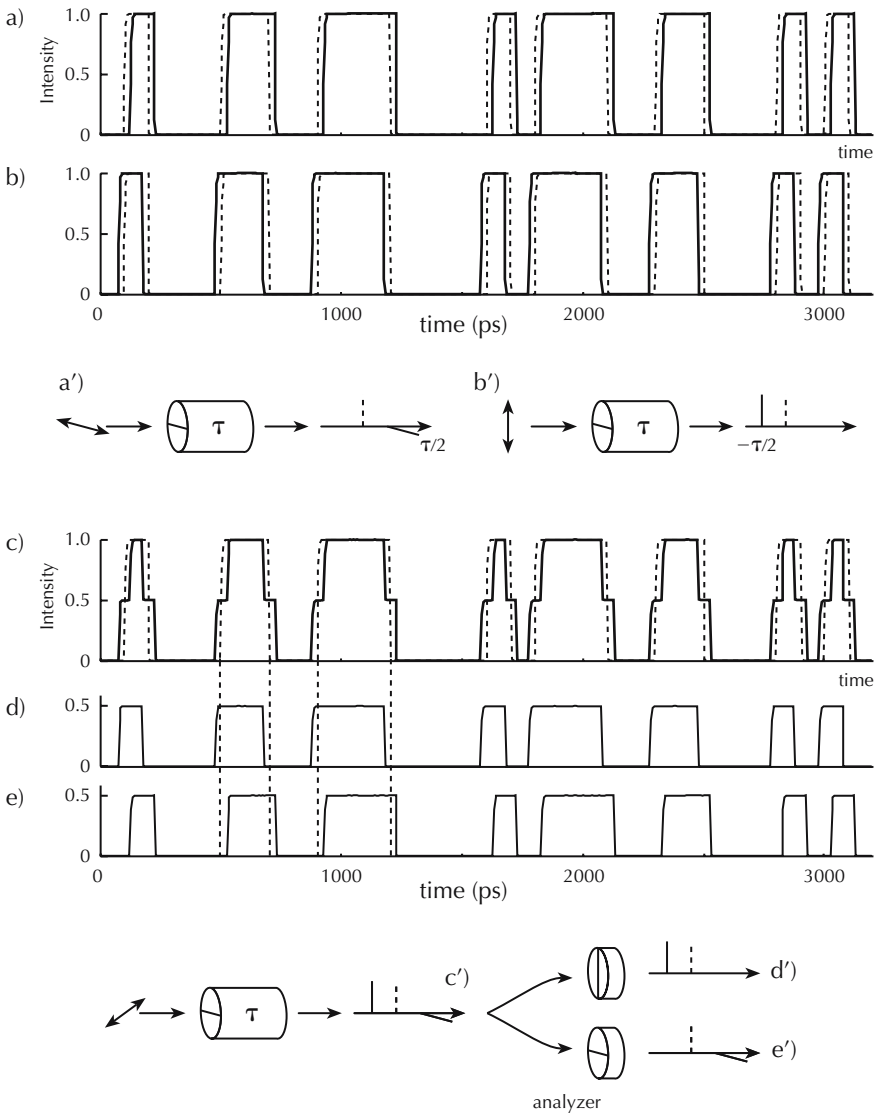


Fig. 8.26. Time response to first-order PMD. a–b) Output signal waveforms delayed and advanced by $\pm\tau/2$; the corresponding launch conditions are (a') and (b'). The output signals are undistorted. c) Distorted output waveform due to launch at 45° with respect to either birefringent axis. When analyzed by output polarizer aligned to either slow or fast axis, the original waveform is recovered, (d) and (e).

copies of the input. Pure first-order PMD comes only from a single birefringent section. Figure 8.26 illustrates the extrema conditions. When the input polarization state is aligned to either the slow or fast birefringent axis,

Figs. 8.26(a') and (b') respectively, the output signal is either delayed or advanced with respect to the average delay, (a) and (b). All the light remains in a single polarization state.

Alternatively, when the input polarization state is equally divided between slow and fast axes (Fig. 8.26(c')), the signal is equally divided between the two axes and time-shifted relative to one another. The square-law detected output is distorted as shown in Fig. 8.26(c). Now, when an analyzer is placed at the output and aligned to the fast axis (d'), all the light from the slow axis is clipped; only the signal along the fast axis emerges. The shape of the analyzed output waveform is identical to the input waveform but with 3 dB less intensity (d). Similarly, when an analyzer is aligned to the slow axis (e'), the analyzed output waveform is also identical to the input but time-delayed by τ (e).

Finally, the polarization dependence of the distortion is evident from (a–c): launch along either birefringent axis leaves the signal undistorted, while the equally-mixed state launch maximally distorts the signal. The distortion is first-order launch-state dependent.

With this one example complete, the reader is warned that DGD is not PMD; DGD is one aspect of PMD. It is all too often forgotten or ignored that second-order and higher-order PMD have significant and characteristically different effects on a signal. Examples of significant activities that have been conducted under the “PMD is DGD” misconception are optical and electronic PMD compensator development; PMD emulator construction; PMD measurement; and PMD outage probability calculations. At least second-order PMD must be considered, and in fact the mean DGD of a link must also be folded into the analysis. Anything short of this richer set of considerations will likely render the calculation or product ineffective for industry applications.

- The first venture into second-order PMD (SOPMD) comes from two birefringent sections alone. These two birefringent sections impart depolarization as well as DGD onto the signal; the second component of SOPMD, the polarization-dependent chromatic dispersion (PDCD), is identically zero for two sections. The distortion of a signal due to second-order PMD is typified by overshoot and false floors. Moreover, the output waveform can *never* be resolved into two identical, time-shifted copies of the input, as was the case for first-order PMD alone.

Figure 8.26 showed that launch of a signal along a birefringent axis (or equivalently, an input PSP) into a one-stage system left the output signal undistorted. This is not true when SOPMD is present. Figure 8.27(a) shows the output distortion when the signal is launched along an input PSP. The output exhibits overshoot and false floors. When an analyzer aligned to either output PSP is placed after the two birefringent sections one observes light along both polarizations. Fig. 8.27(b) shows the analyzed components of the output signal, and (c) shows an excerpt. The evident effect is that the light

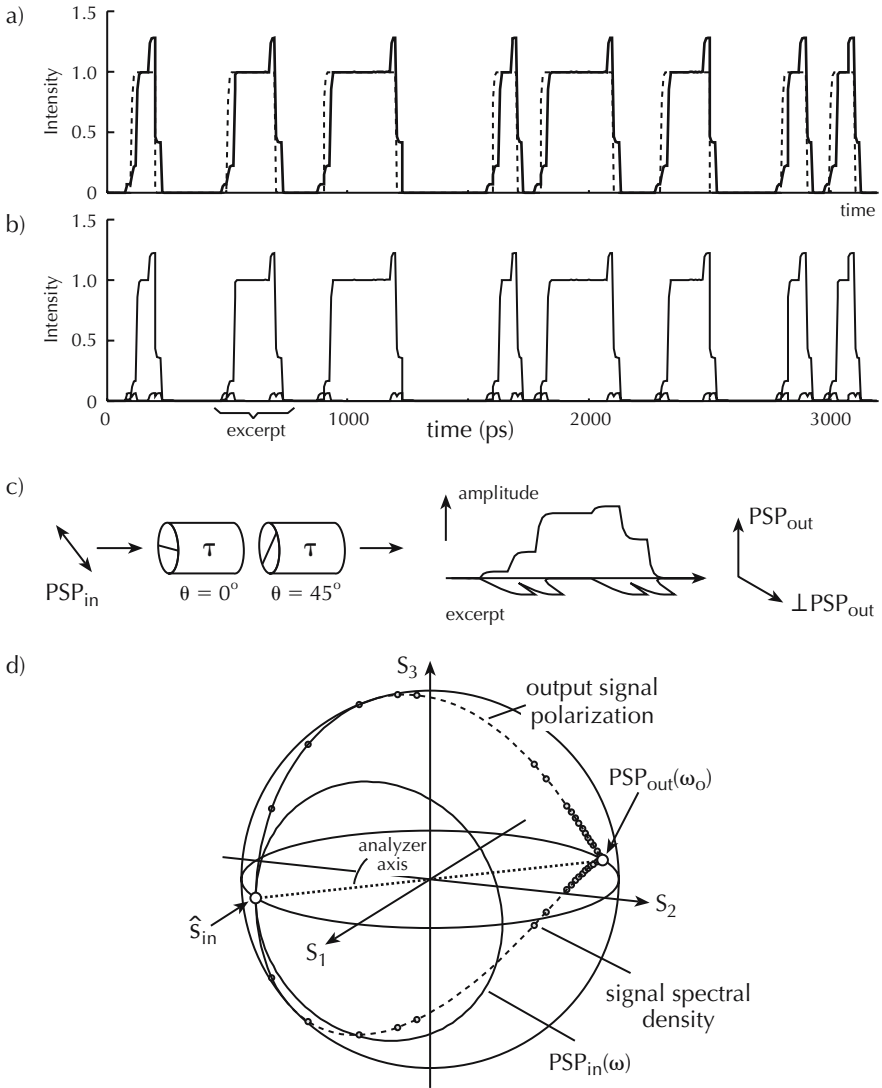


Fig. 8.27. Time-response to second-order PMD with only depolarization. a) Distorted output signal when launched state is aligned to the input PSP at a center frequency. b) Output signal analyzed by a polarizer aligned along the center-frequency output PSP and its orthogonal state. Light comes through the polarizer in both states; the output cannot be constructed from two time-shifted identical copies of the input. c) Blowup of excerpted signal, plotted in amplitude rather than intensity to highlight the relation of the waveforms. The transition edges of the input signal have in part been rotated down to the orthogonal polarization axis. d) Input launch state $\hat{s}_{\text{in}}$, input PSP spectrum, and output signal polarization spectrum.

on the orthogonal PSP axis coincides with the transition edges of the input waveform.

Along the transition edges the high-frequency components of the signal spectrum have their phases aligned; one can say the high-frequency Fourier components of the signal dominate at the edges, while the low-frequency components dominate at the centers of the “ones” and “zeros.” With this in mind, the polarization spectra of the input PSP and output signal polarization are shown in Fig. 8.27(d). The launch polarization state is fixed in frequency, but the input PSP is not. At a center frequency the input PSP coincides with the launch state (here, not in general), but the PSP vector traces a circle in frequency. Accordingly, frequency components of the input signal are mapped to onto a locus of polarization states at the output; this is called polarization-state dispersion. The output signal spectrum is drawn in the figure. Overlaid with the output signal spectrum are small circles that indicate the amplitude of the signal spectrum. The signal spectrum is densely packed about $\text{PSP}_{\text{out}}(\omega_o)$, but at higher and lower relative frequencies the output polarization makes large excursions. Thus the high frequency components of the signal are misaligned to the output analyzer (when aligned to the output PSP) and come through. Those high-frequency components appear on the transition edges of the signal.

The impulse response of two birefringent sections is comprised of two impulses aligned along one axis and two impulses aligned along an orthogonal axis. This is shown in Fig. 8.28(a) for two equal delays. The output signal is the convolution of the input signal with this impulse response. The effect of interference, resulting in coherent addition or cancellation, of co-polarized signal images can now be well illustrated. The three columns in the center of Fig. 8.28 show how co-polarized signal images add for three different residual birefringent phases ϕ in the first delay section. When $\phi = 0^\circ$, the impulses along the u -polarization have zero phase difference. Therefore when the convolved signal images overlap the underlying carriers are in phase and add (Fig. 8.28(c-d)). The phase relation is indicated by + signs. However, along the v -polarization the impulses are out of phase by 180° ; when the signal images overlap the fields subtract, indicated by the – sign, but when they do not overlap the partial signal images come through (Fig. 8.28(d)). A square-law detector adds the orthogonal components in quadrature. The waveform in Fig. 8.28(e) for $\phi = 0^\circ$ shows the result.

In the case when $\phi = 90^\circ$, the phase difference between both pairs of co-polarized impulses is zero. In this case the signal images along the u -axis add as do the signal images along the v -axis. The resultant waveform is shown in Fig. 8.28(e). Finally, when $\phi = 180^\circ$ the u -axis impulses are out of phase while the pair on the v -axis are in phase. The result is the mirror image of $\phi = 180^\circ$ about τ_o .

Interference of co-polarized signal images plays a central role in how the PMD manifests itself on an input signal. A slight phase change clearly can change the enter shape of the signal. When the width of the PMD impulse

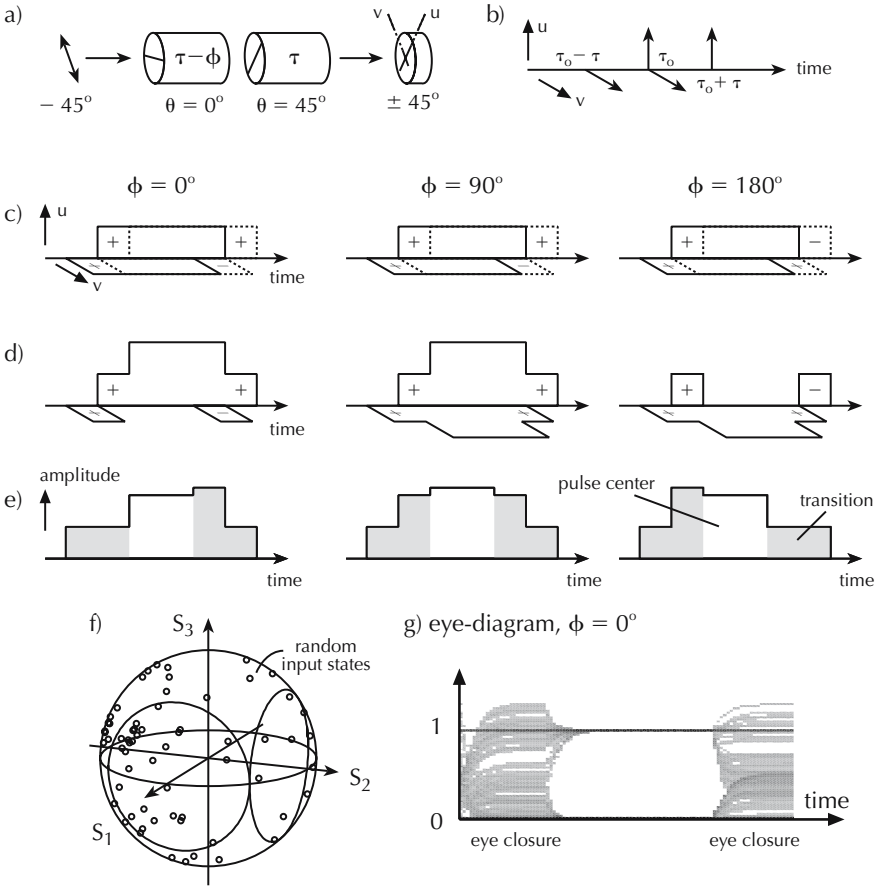


Fig. 8.28. Interference of an impulse response from two stages. a) Two stages generate an impulse with two co-polarized impulses along a u -axis and two more co-polarized impulses along a v -axis. c–e) Four signal images as convolved onto the impulse response. When two co-polarized impulses are aligned in phase the signal images add (denoted by +); when the same impulses are out of phase the signal images subtract (denoted by –). The complete output signal as measured by a square-law detector adds the coherently-added components in quadrature. f) Randomly generated launch states, and g) eye diagram of cumulative distortion.

response is on the order of the signal duration, the expression of coherent addition deep within a signal “one” or “zero” depends on this phase; when the phases cancel, the distortion will take place closer to the transitions, but when the phases add the distortion will be toward the center. In either case the temporal impulse locations do not change more than a fraction of a wave.

One aspect that is similar to first-order PMD is that signal distortion is critically dependent on launch state. Every launch state produces a different

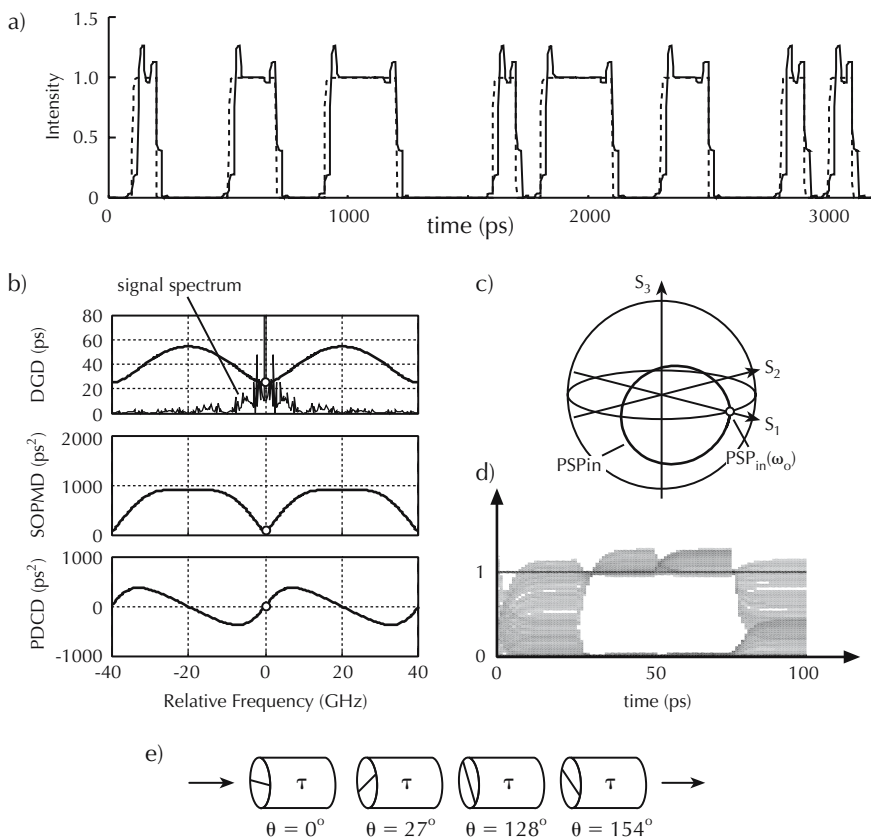


Fig. 8.29. Signal distortion for DGD with low average SOPMD. a) Comparison of distorted signal to input signal. Launch state is aligned to the input PSP at band center, indicated in (c). Four 25 ps delay sections are concatenated with mode-mixing angles shown in (e). Resultant scalar PMD spectra are shown in (b). For comparison, the signal spectrum is overlaid with the DGD spectrum. The eye diagram (d) is calculated from uniformly distributed random launch states.

distortion effect, although degeneracies may exist. Figure 8.28(g) shows the calculation of an eye diagram for 64 randomly and uniformly distributed input states. Since each delay section is 25 ps long, the outer 25 ps of the pulse are completely blurred. The overshoot is also evident. This gives a flavor why any measurement of bit-error rate or eye-margin penalty should be made while scrambling the input polarization and the measurement must last until the Poincaré sphere is reasonably covered by the input state.

- Extension beyond two birefringent stages leads toward anecdotal examples or a statistical treatment of PMD effects. A couple of important examples still remain to be shown, even though they are two of a subset in a larger class.

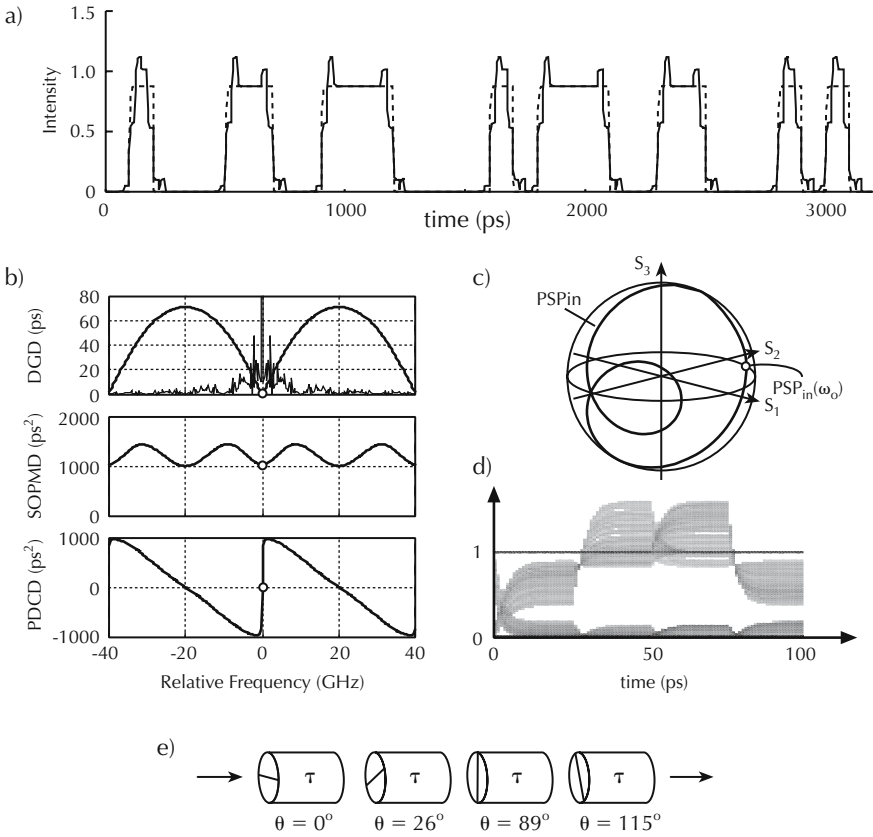


Fig. 8.30. Signal distortion for low average DGD with finite SOPMD. a) Comparison of distorted signal to input signal. Launch state is perpendicular to the input PSP at band center, the latter indicated in (c). Four 25 ps delay sections are concatenated with mode-mixing angles shown in (e). Resultant scalar PMD spectra are shown in (b). For comparison, the signal spectrum is overlaid with the DGD spectrum. The eye diagram (d) is calculated from uniformly distributed random launch states. The SOPMD creates large waveform distortions in the center region of the eye.

The examples are signal distortion for DGD with low average SOPMD, and for low average DGD with finite SOPMD. Figures 8.29 and 8.30 illustrate the two cases. Both calculations show that even when one PMD component is diminished with respect to the other, complicated distortions still occur.

The first example is DGD with low average SOPMD: (Fig. 8.29. The four birefringent sections, each 25 ps in this example, and their relative alignment is shown in (e). This example is special because the SOPMD vanishes at band center (b). The SOPMD can vanish when the output PSP pirouettes about a stationary point as the frequency changes. The PSP vector eventually stops

its pirouette and continues on a course along the Poincaré sphere; the input PSP spectrum is shown in (c). The launch state chosen for this example is the input PSP at the pirouette position. The scalar spectra of the PMD condition is shown in (b)⁴; the signal spectrum is overlayed with the DGD spectrum for comparison. The output waveform is shown in (a); observe the large overshoots and false floors even though the SOPMD is significant only at higher waveform frequencies. An eye diagram for uniformly distributed random input launch states is shown in (d). Since the four section delay totals 100 ps, the width of the impulse response is 100 ps, 50 ps of which extends into the interior of the signal waveform. This is why the center region of the waveform is distorted. However, the particular location of the signal spectrum with respect to the PMD spectrum prevents the marginal impulses at either end of the PMD impulse response to have strong weight. Thus the eye center is not closed.

The second example is low average DGD with finite SOPMD: Fig. 8.30. Here, the DGD vanishes at band center, while the SOPMD remains significant. The DGD can vanish at one frequency when the component PMD vectors form a closed loop in Stokes space. Since in this case there are four component vectors, the closed loop is a square or rhombus. The depolarization must be finite in this case, or else the closed loop of PMD component vectors would not open back up (or close on itself in the first place). Nonetheless, even with the signal spectrum aligned to the vanishing point of the DGD, the signal distortion is significant. The launch state associated with the distortion in (a) is perpendicular to the input PSP at band center. The input PSP spectrum is shown in (c). Of particular interest is the eye diagram generated for a large number of randomly selected launch states. Even when the average DGD is low, significant distortion appears all across the pulse. This is caused by the marginal impulses in the PMD impulse response having significant weight. The eye center is still not closed, though, because the average DGD is low.

Signal Moments and Distortion

One can make some general statements about the effects of PMD by looking at the moments of the signal waveform in the time domain.

Karlsson calculates the first and second moments of a signal waveform when effected by PMD [38]. Gordon offers details of Karlsson's presentation [20]. The main results are that: 1) the differential-group delay is the maximum possible delay of a narrowband signal launched into the fast and slow input PSP's, all other launch conditions yield a first-moment less than the DGD; 2) there is a minimum, non-zero pulse broadening in the presence of second-order PMD principally due to the rotation of the PSP's in frequency.

The first- and second-moments of the waveform are calculated by

⁴ Appendix C shows how to efficiently calculate the vector and scalar spectra for an arbitrary birefringent concatenation.

$$\langle t \rangle = \frac{1}{W} \int_{\mathbb{R}} t e^\dagger(t) e(t) dt = \frac{2j\pi}{W} \int_{\mathbb{R}} E^\dagger(\omega) E_\omega(\omega) d\omega \quad (8.2.43a)$$

$$\langle t^2 \rangle = \frac{1}{W} \int_{\mathbb{R}} t^2 e^\dagger(t) e(t) dt = \frac{2\pi}{W} \int_{\mathbb{R}} E_\omega^\dagger(\omega) E_\omega(\omega) d\omega \quad (8.2.43b)$$

The first moment at the input is simply

$$\begin{aligned} \langle t_s \rangle &= \frac{2j\pi}{W} \int_{\mathbb{R}} E_s^\dagger E_{s\omega} d\omega \\ &= \frac{2j\pi}{W} \int_{\mathbb{R}} f_\omega^*(\omega) f(\omega) d\omega \end{aligned}$$

If the waveform spectrum is real and symmetric then $\langle t_s \rangle = 0$. Since the output-field spectrum is $E_t = \exp(-j\phi_o) U E_s$, where ϕ_o is the common phase through the system and U is the unitary operator for the cascade, the integrand to $\langle t_t \rangle$ at the output is

$$E_t^\dagger E_{t\omega} = -j E_s^\dagger \left(\tau_o + \frac{1}{2} (\vec{\tau}_s \cdot \vec{\sigma}) \right) E_s + E_s^\dagger E_{s\omega}$$

where $\vec{\tau}_s$ is the input principal state (recall $2jU_\omega U^\dagger = \vec{\tau}_s \cdot \vec{\sigma}$) and τ_o is the common delay through the system. Note that τ_o and $\vec{\tau}$ are both functions of frequency. The first moment of the output waveform is therefore

$$\begin{aligned} \langle t_t \rangle &= \langle t_s \rangle + \frac{2\pi}{W} \int_{\mathbb{R}} E_s^\dagger \left(\tau_o + \frac{1}{2} (\vec{\tau}_s \cdot \vec{\sigma}) \right) E_s d\omega \\ &= \langle t_s \rangle + \frac{2\pi}{W} \int_{\mathbb{R}} |f(\omega)|^2 \left(\tau_o + \frac{1}{2} (\vec{\tau}_s \cdot \hat{s}) \right) d\omega \end{aligned} \quad (8.2.44)$$

The mean signal delay τ_g through the medium is simply the difference of first moments:

$$\tau_g = \langle t_t \rangle - \langle t_s \rangle \quad (8.2.45)$$

This can be expressed in spectrally averaged form as

$$\tau_g = \left\langle \tau_o(\omega) |f(\omega)|^2 \right\rangle_\omega + \frac{1}{2} \left\langle \hat{s}(\omega) \cdot \vec{\tau}_s(\omega) |f(\omega)|^2 \right\rangle_\omega \quad (8.2.46)$$

where $\hat{s}(\omega)$ allows for the possibility of frequency dependence of the input state. Physically, the mean signal delay is the normalized spectral average of the common delay weighted by the waveform envelope, plus the spectral average of the input launch polarization as projected onto the input PSP, again weighted by the waveform envelope.

An important observation is that the phase of the waveform envelope $f(t)$ is eliminated in (8.2.46); initial chirp or chromatic dispersion of the pulse does not affect its average position at the output.

Consider two extreme cases for (8.2.46): monochromatic input, and any input to a single birefringence segment. For monochromatic input, the waveform input is a sine wave, so in frequency $f(\omega) = \delta(\omega - \omega_o)$. The mean signal delay for any concatenation is

$$\tau_g = \tau_o + \frac{1}{2} \tau (\hat{s} \cdot \hat{r}_s) \quad (8.2.47)$$

where each term is evaluated at ω_o . This expression is the main result which connects the mean signal delay to the spectral description of PMD. The mean delay at ω_o is $\tau_g = \tau_o \pm \tau/2$ when the launch state is parallel or perpendicular to the input PSP. Any intermediate launch condition produces a first-moment that is between these extrema. Accordingly, one can say that the DGD is the maximum delay at a particular frequency between the fast and slow axes of a cascade. This interpretation was used in the time-frequency correspondence figure (Fig. 8.18 on page 326).

When there is only one homogeneous birefringent element then τ_o and $\vec{\tau}_s$ are stationary with frequency; the mean signal delay has the same form as (8.2.47). This is an important connection to the impulse response of a cascade which is considered in the next section.

The pulse spreading between the output and input can be measured by the second moments of the signal. The pulse spread $\Delta\tau$ is defined as

$$\Delta\tau = \sqrt{(\langle t_t^2 \rangle - \langle t_s^2 \rangle) - (\langle t_t \rangle - \langle t_s \rangle)^2} \quad (8.2.48)$$

The second moment of the input waveform is

$$\langle t_s^2 \rangle = \frac{2\pi}{W} \int_{\mathbb{R}} E_{s\omega}^\dagger E_{s\omega} d\omega = \frac{2\pi}{W} \int_{\mathbb{R}} |f_\omega(\omega)|^2 d\omega \quad (8.2.49)$$

At the output, $E_{t\omega} = T_\omega E_s + T E_{s\omega}$. The difference between output and input second moments is

$$E_{t\omega}^\dagger E_{t\omega} - E_{s\omega}^\dagger E_{s\omega} = E_s^\dagger T_\omega^\dagger T_\omega E_s + E_s^\dagger T_\omega^\dagger T E_{s\omega} + E_{s\omega}^\dagger T^\dagger T_\omega E_s$$

The frequency derivative of the transformation matrix is

$$T_\omega = -j e^{-j\phi_o} U \left(\tau_o + \frac{1}{2} \vec{\tau}_s \cdot \vec{\sigma} \right)$$

where the substitution $jU^\dagger U_\omega = \frac{1}{2} \vec{\tau}_s \cdot \vec{\sigma}$ was made. Since $jU^\dagger U_\omega$ is Hermitian, the matrix product is $T_\omega^\dagger T_\omega = \left(\tau_o + \frac{1}{2} \vec{\tau}_s \cdot \vec{\sigma} \right)^2$, or

$$T_\omega^\dagger T_\omega = \tau_o^2 + \frac{1}{4} \tau^2 + \tau_o (\vec{\tau}_s \cdot \vec{\sigma})$$

where the identity $(\vec{\tau}_s \cdot \vec{\sigma})(\vec{\tau}_s \cdot \vec{\sigma}) = \tau^2$ was used. Therefore

$$E_s^\dagger T_\omega^\dagger T_\omega E_s = |f|^2 \left(\tau_o^2 + \frac{1}{4} \tau^2 + \tau_o (\vec{\tau}_s \cdot \hat{s}) \right)$$

For the remaining two terms, one has $T_\omega^\dagger T = j \left(\tau_o + \frac{1}{2} \vec{\tau}_s \cdot \vec{\sigma} \right)$. The sum is

$$E_s^\dagger T_\omega^\dagger T E_{s\omega} + E_{s\omega}^\dagger T^\dagger T_\omega E_s = j (f^* f_\omega - f_\omega^* f) \left(\tau_o + \frac{1}{2} \vec{\tau}_s \cdot \hat{s} \right)$$

Expansion of the complex waveform envelope into amplitude and phase, $f(\omega) = a(\omega) \exp(-j\phi(\omega))$, one finds that

$$j(f^* f_\omega - f_\omega^* f) = 2|f|^2 \phi_\omega$$

where ϕ_ω is the frequency derivative of the waveform phase. Now that each integrand has been reduced, the complete second moment of the pulse spread is

$$\begin{aligned} \langle t_t^2 \rangle - \langle t_s^2 \rangle &= \frac{2\pi}{W} \int_{\mathbb{R}} |f(\omega)|^2 (\tau_o^2 + \frac{1}{4}\tau^2 + \tau_o (\vec{\tau}_s \cdot \hat{s})) d\omega \\ &+ \frac{4\pi}{W} \int_{\mathbb{R}} |f(\omega)|^2 \phi_\omega(\omega) (\tau_o + \frac{1}{2} \vec{\tau}_s \cdot \hat{s}) d\omega \quad (8.2.50) \end{aligned}$$

The first term on the right-hand side is similar in form to (8.2.46), where the waveform envelope intensity is the weighting factor to the spectral average of the delay components. Like the first moment, this term of the second moment does not depend on the phase of the waveform. The second term, however, includes the derivative of the waveform phase. Therefore, the pulse spread is related not only to the PMD but to the phase across the waveform as well. For example, for a linear delay across the waveform, $\phi_\omega(\omega) = \gamma\omega$, and the resultant ω multiplier in the integrand of the second term is equivalent to a time-derivative of the waveform intensity. The pulse spread then depends on the temporal details of the signal shape as well as the intensity spectrum.

Consider an example of one section of birefringence, the PMD τ is only first order, and the signal is real ($\phi_\omega(\omega) = 0$). The difference between output and input first and second moments is then

$$\begin{aligned} \langle t_t^2 \rangle - \langle t_s^2 \rangle &= \tau_o^2 + \frac{1}{4}\tau^2 + \tau_o (\vec{\tau}_s \cdot \hat{s}) \\ \langle t_t \rangle - \langle t_s \rangle &= \tau_o + \frac{1}{2} (\vec{\tau}_s \cdot \hat{s}) \end{aligned}$$

where $\hat{s}$ is the input launch state. Expansion of the frequency-independent dot product between launch state and input PMD vector gives $(\vec{\tau}_s \cdot \hat{s}) = \tau \cos \theta$, where θ is the Stokes angles between vectors. The pulse spread is therefore

$$\Delta\tau = \frac{1}{2} \tau \sin \theta \quad (8.2.51)$$

For comparison, the signal delay is

$$\tau_g = \tau_o + \frac{1}{2} \tau \cos \theta$$

Figure 8.31 illustrates the example. The maximum arrival-time deviation from τ_o is when the signal is launched along the fast or slow axis of the birefringent element (Fig. 8.31(a)). In this case $\cos \theta = \pm 1$. There is no pulse spreading with this condition. This is reasonable since all the energy is inserted into one eigenstate; no differential delay is experienced. However, if the launch is at 45° with respect to the input, then $\tau_g = \tau_o$ and the pulse spreading achieves its maximum at $\Delta\tau = \tau/2$ (Fig. 8.31(b)). Clearly when the DGD of the system approaches the time-slot duration of the signal then all communication can be lost.

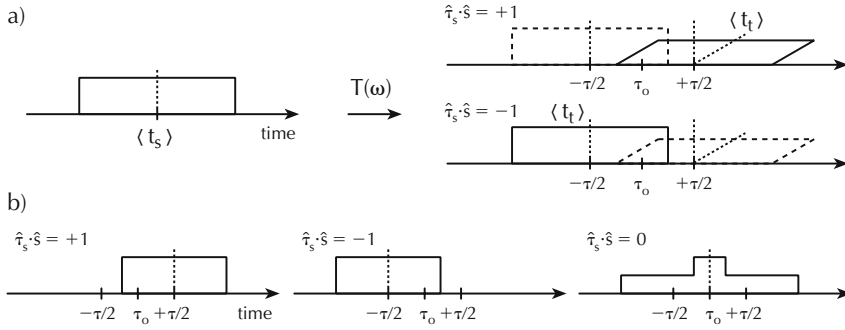


Fig. 8.31. Illustration of signal envelope through a very simple high-birefringence (first-order PMD only) system. a) When the signal is launched along a birefringent axis it is either advanced or delayed with respect to the isotropic travel time τ_o . b) Three launch conditions: aligned to slow axis, aligned to fast axis, and equally mixed between slow and fast. The waveform shapes represent amplitude and not polarization. The mixed launch imparts pulse spreading.

Consider an example of two sections of birefringence (cf. Fig. 8.21). This is the simplest form of second-order PMD: only depolarization exists. The output PMD vector is calculated from the concatenation rule: $\vec{\tau}_t = \vec{\tau}_2 + R_2 \vec{\tau}_1$. The moment calculations use the input PSP so that the launch state $\hat{s}$ can be projected. The corresponding input PMD vector is $\vec{\tau}_s = \vec{\tau}_1 + R_1^\dagger \vec{\tau}_2$. The motion of the input PSP is illustrated in Fig. 8.21, but with the terms $\vec{\tau}_1$ and $\vec{\tau}_2$ interchanged. The input PSP traces a circle in Stokes space over frequency while the cumulative DGD remains fixed; the period of the circle is $\Delta\omega = 2\pi/\tau_1$. The mean delay of the output waveform is

$$\tau_g = \tau_o + \frac{2\pi}{W} \int_{\mathbb{R}} |f|^2 \frac{1}{2} (\vec{\tau}_s(\omega) \cdot \hat{s}) d\omega$$

Defining the integral

$$I_s = \frac{2\pi}{W} \int_{\mathbb{R}} |f|^2 (\hat{r}_s(\omega) \cdot \hat{s}) d\omega$$

where $\hat{r}_s$ is the pointing direction of the input PSP, itself a function of frequency, the mean delay is then recast as

$$\tau_g = \tau_o + \frac{1}{2} \tau I_s$$

The cumulative PMD τ was extracted from the integral I_s because its magnitude is fixed in frequency, as is the case for two birefringent stages. In a similar manner, the square of the pulse spreading at the output is

$$\begin{aligned} \Delta\tau^2 &= \tau_o^2 + \frac{1}{4} \tau^2 + \tau_o \tau I_s - \tau_g^2 \\ &= \frac{1}{4} \tau^2 (1 - I_s^2) \end{aligned}$$

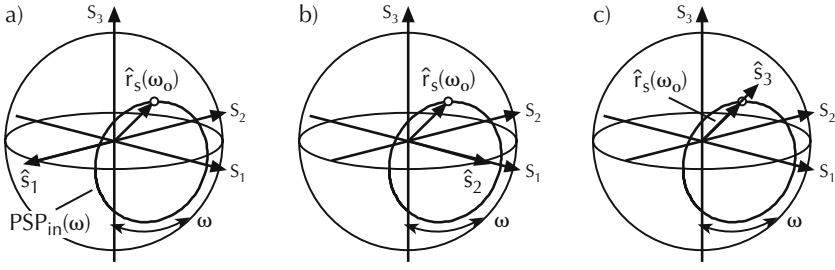


Fig. 8.32. Three launch conditions into a depolarizing system. Two birefringent sections generate precession of the input PMD vector about $\vec{\tau}_1$. a) Launch state for maximum pulse spread, $\hat{s}_1 = \pm \vec{\tau}_1 \times \vec{\tau}_2$. b) Intermediate launch state where $\hat{s}_2 \cdot \hat{r}_s$ is constant in frequency. c) Launch state for largest minimum pulse spread, $\hat{s}_3 = \hat{r}_s(\omega_0)$.

Since I_s enters the equation as a squared quantity, the pulse spread can only decrease with I_s . The value of I_s depends on the launch state at the input and the degree of rotation of $\hat{r}_s$. Figure 8.32 illustrates three launch states, the first and last states being the extrema. The maximum pulse spreading occurs when $I_s = 0$. State $\hat{s}$ can always be selected to drive I_s to zero. For a symmetric spectrum centered at ω_0 and launch state $\hat{s}_1 = \pm \vec{\tau}_1 \times \vec{\tau}_2$ (Fig. 8.32(a)), where $\vec{\tau}_1$ and $\vec{\tau}_2$ are evaluated at ω_0 , the product $\hat{r}_s \cdot \hat{s}_1$ is antisymmetric. Integral I_s therefore vanishes and the pulse spread is $\Delta\tau^2 = \tau^2/4$.

An interesting albeit non-extrema case is where the inner product is fixed in frequency. This occurs when $\hat{s}$ is aligned with $\vec{\tau}_1$ (Fig. 8.32(b)). In this case $I_s = \cos\theta$. Only when the two birefringent axes are aligned does the pulse spread reach zero. But this is simply the case of a single birefringent segment made up of two parts. The general case is when there is mode mixing between sections. Thus in the general $\hat{s}_2$ is not a state that produces minimum pulse spreading.

The launch state for minimum pulse spreading is illustrated in Fig. 8.32(c). In general, I_s will be less than unity, so the pulse always experiences non-zero minimum spread. This contrasts with the first-order PMD situation where launch along a PSP ensures zero spreading. The problem here is that the PSP's move with frequency, so there is no single PSP to launch into.

One can calculate the largest minimum pulse spread. This case coincides with maximum depolarization, which is when $\tau_1 = \tau_2$ and $\hat{r}_s \cdot \hat{s}_1 = 0$. The PMD vector at the input is

$$\vec{\tau}_s/\tau_1 = \hat{r}_1 - \sin\omega\tau_1 \hat{r}_1 \times \hat{r}_2 - \cos\omega\tau_1 \hat{r}_1 \times \hat{r}_1 \times \hat{r}_2$$

and the launch state is $\hat{s}_3 = (\hat{r}_1 + \hat{r}_2)/\sqrt{2}$. Since the length of the PMD vector is $\tau_s = \sqrt{2}\tau_1$, the frequency-dependence of the inner product is

$$\hat{s}_3 \cdot \hat{r}_s = \frac{1}{2} (1 + \cos\omega\tau_1)$$

In the regime where the bandwidth of the waveform is greater than $1/\tau_1$, then the frequency average of the inner product is driven to one-half. In this case

the minimum pulse spread is approximately (ignoring details of the waveform spectrum) $\Delta\tau^2 \simeq (3/16)\tau^2$. The comparison between the two extrema is

$$\begin{aligned} \text{maximum pulse spread: } \hat{s} &= \pm \vec{\tau}_1 \times \vec{\tau}_2, \quad \Delta\tau = \tau/2 \\ \text{minimum pulse spread: } \hat{s} &= \hat{r}(\omega_o), \quad \Delta\tau \rightarrow \sqrt{3}\tau/4 \end{aligned} \quad (8.2.52)$$

Nearly the same pulse spreading as found for the “best” and “worst” launch conditions. Clearly depolarization can have a strong effect on the waveform.

PMD Impulse Response

Gisin and Pellaux deduced the formal connection between the PMD impulse response and the DGD spectrum of a lossless birefringent cascade [19]. Their result shows that the root-mean-square (rms) DGD is equal to the standard deviation of the impulse response width. The derivation is elegant and yet result seems to be under represented in the subsequent literature.

The derivation constructs a recurrence relation for the frequency average of the mean-squared DGD, denoted $\langle\tau^2(\omega)\rangle$, and a separate recurrence relation for the second-moment of the impulse response, denoted $\langle h^2(t)\rangle$, from the same concatenation. The recurrence relations are then shown to be equivalent.

Consider a concatenation of N homogeneous birefringent sections, each section having DGD of τ_n and birefringent-vector orientation $\hat{r}_n$. In the frequency domain, the two-block PMD concatenation rule (8.2.33) on page 335 is written to relate the last element, element N , to the preceding $N-1$ elements as

$$\vec{\tau}(N) = \vec{\tau}_N + R_N \vec{\tau}(N-1) \quad (8.2.53)$$

where $\vec{\tau}(N)$ denotes the cumulative PMD through section N and $\vec{\tau}_N$ denotes the N^{th} local birefringent element. The cumulative vector $\vec{\tau}(N)$ is clearly a function of frequency ω while the local elements, such as $\vec{\tau}_N$, are not. The dot product of $\vec{\tau}$ with itself provides the DGD squared, in general $\tau^2 = \vec{\tau} \cdot \vec{\tau}$. The DGD squared for $\vec{\tau}(N)$ from recurrence relation (8.2.53) is

$$\tau^2(N) = \tau_N^2 + \tau^2(N-1) + 2\vec{\tau}_N \cdot \vec{\tau}(N-1)$$

The DGD squared spectrum, $\tau^2(N; \omega)$, is then averaged over all frequency to find its mean-square. The average is written

$$\langle\tau^2(N)\rangle_\omega = \tau_N^2 + \langle\tau^2(N-1)\rangle_\omega + 2\langle\vec{\tau}_N \cdot \vec{\tau}(N-1)\rangle_\omega$$

The value τ_N^2 comes through the average as its just a number. The last term on the right-hand side may be further reduced by employing (8.2.53) to the $N-1$ term:

$$\begin{aligned} \langle\vec{\tau}_N \cdot \vec{\tau}(N-1)\rangle_\omega &= \langle\vec{\tau}_N \cdot (\vec{\tau}_{N-1} + R_{N-1} \vec{\tau}(N-2))\rangle_\omega \\ &= \vec{\tau}_N \cdot \vec{\tau}_{N-1} + \vec{\tau}_N \cdot \langle R_{N-1} \vec{\tau}(N-2)\rangle_\omega \end{aligned} \quad (8.2.54)$$

The evaluation of the last term on the right-hand side is at the heart of the derivation. Expansion of the last term to include the rotation operator R_{N-1} gives

$$\begin{aligned} \langle R_{N-1} \vec{\tau}(N-2) \rangle_\omega &= \langle \hat{r}_{N-1} \hat{r}_{N-1} \cdot \vec{\tau}(N-2) \rangle_\omega \\ &\quad + \langle \sin(\omega \tau_{N-1}) \hat{r}_{N-1} \times \vec{\tau}(N-2) \rangle_\omega \\ &\quad - \langle \cos(\omega \tau_{N-1}) \hat{r}_{N-1} \times \hat{r}_{N-1} \times \vec{\tau}(N-2) \rangle_\omega \end{aligned} \quad (8.2.55)$$

The last two frequency-average terms include sine and cosine terms that average to zero. Even though $\vec{\tau}(N-2)$ itself generally varies with frequency, for a concatenation with enough segments the frequency average will eventually drive these terms to zero. In contrast, the disposition of the first term on the right-hand is not so clear, so it remains. Insertion of (8.2.55) into (8.2.54) and some manipulation produces

$$\langle \vec{\tau}_N \cdot \vec{\tau}(N-1) \rangle_\omega = \tau_N (\hat{r}_N \cdot \hat{r}_{N-1}) (\tau_{N-1} + \langle \hat{r}_{N-1} \cdot \vec{\tau}(N-2) \rangle_\omega)$$

The complete recurrence relation in the frequency domain is therefore

$$\begin{aligned} \langle \tau^2(N) \rangle_\omega &= \tau_N^2 + \langle \tau^2(N-1) \rangle_\omega \\ &\quad + 2 \tau_N (\hat{r}_N \cdot \hat{r}_{N-1}) (\tau_{N-1} + \langle \hat{r}_{N-1} \cdot \vec{\tau}(N-2) \rangle_\omega) \end{aligned} \quad (8.2.56)$$

The form of this expression has fully identified the effect of $\vec{\tau}_N$ on $\vec{\tau}(N-1)$. It is this recurrence relation that will be compared to the impulse-response relation.

Now for the impulse response. Each impulse has a position in time and a weight. As there is no loss, the sum of all the weights remains fixed regardless of the number of sections in the concatenation. The time-position and weight of each impulse depends on the path the light takes. If there are n sections, there are 2^n paths and 2^n impulses at the output. Each one has to be enumerated to construct the impulse response. The time position of the k^{th} pulse with respect to the common delay through the concatenation is

$$t_k = \epsilon_1(k)\tau_1 + \epsilon_2(k)\tau_2 + \dots + \epsilon_N(k)\tau_N \quad (8.2.57)$$

where $\epsilon_i = \pm 1$ depending whether the impulse travels along the slow or fast axis of the i^{th} segment⁵ To enumerate each path only once the binary equivalent of the decimal index is useful. If for each index integer $k \in [0, N-1]$ the integer is converted to its binary form $k \mapsto b_0 b_1 b_2 \dots b_N$, then the path selector ϵ is defined by $\epsilon_i = 1 - 2b_i$. The path selector is further indexed by k so that $\epsilon_i(k)$ is generated by the i^{th} binary digit of the binary representation of the index k .

⁵ The factor of one-half is dropped to conform with the PMD vector definition $\vec{\tau} = \tau \hat{p}$ rather than $\vec{\tau} = \tau/2 \hat{p}$.

The weight of the k^{th} impulse is determined by the projection of adjacent birefringent axes and whether the impulse is going between fast axis in one segment to the fast axis in the next, or between fast and slow, slow and fast, or slow and slow. The weight of the impulse is

$$w(k) = \frac{1}{2} \prod_{n=2}^N \frac{1}{2} \left(1 + \epsilon_{n-1}(k) \epsilon_n(k) \hat{r}_{n-1} \cdot \hat{r}_n \right) \quad (8.2.58)$$

where the first one-half factor comes from a 45° launch into the first element, that is $\hat{s}_{\text{in}} \cdot \hat{r}_1 = 0$. The weights are normalized such that

$$\sum_{k=1}^{2^N} w(k) = 1$$

The impulse response of the birefringent concatenation is

$$h(t, N) = \sum_{k=1}^{2^N} w(k) \delta(t - t_k)$$

where $\delta(t - t_o)$ is the dirac delta function that has zero value everywhere but for t_o . In the following the explicit time dependence of h will be dropped. The average position of the impulse response relative to the common delay is zero:

$$\langle h(N) \rangle_t = \int_{\mathbb{R}} t h(t, N) dt = \sum_{k=1}^{2^N} w(k) t_k = 0 \quad (8.2.59)$$

The pulse train is symmetric about $t = 0$. The second moment of $h(t, N)$ requires more work. The second moment is

$$\langle h^2(N) \rangle_t = \int_{\mathbb{R}} t^2 h(t, N) dt = \sum_{k=1}^{2^N} w(k) t_k^2 \quad (8.2.60)$$

The square of the k^{th} time position is expanded with (8.2.57) to give

$$\begin{aligned} t_k^2 &= \sum_{i=1}^N \sum_{j=1}^N \epsilon_i(k) \epsilon_j(k) \tau_i \tau_j \\ &= \sum_{i=1}^N \tau_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N \epsilon_i(k) \epsilon_j(k) \tau_i \tau_j \end{aligned}$$

where the first term on the right-hand side of the second line is the sum along the diagonal of the $N \times N$ matrix while the second term is the sum over the upper triangle of the matrix. Substitution back into (8.2.60) gives

$$\langle h^2(N) \rangle_t = \sum_{i=1}^N \tau_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N \tau_i \tau_j \sum_{k=1}^{2^N} \epsilon_i(k) \epsilon_j(k) w(k)$$

where the normalization of $w(k)$ was applied to the first right-hand-side term. The sum over k in the second term has an significant simplification. The binary-weight product $\epsilon_i(k)\epsilon_j(k)$ changes sign when counting over k with a frequency that depends on j . For instance with $N = 4$, $\epsilon_i(k)\epsilon_j(k)$ changes sign for every increment of k when $j = 4$, but changes sign for every two increments when $j = 3$. Concurrently, $\epsilon_{n-1}(k)\epsilon_n(k)$ changes sign with a rate when counting over k related to n . The combined result is that all terms of $\epsilon_{n-1}(k)\epsilon_n(k)$ that change sign at a counting-rate faster than $\epsilon_i(k)\epsilon_j(k)$ vanish from the sum over k while the remaining terms add to a unity coefficient. That is, all $n > j$ terms vanish.

The second-moment of $h(N)$ simplifies to

$$\langle h^2(N) \rangle_t = \sum_{i=1}^N \tau_i^2 + 2 \sum_{i=1}^N \sum_{j=i+1}^N \tau_i \tau_j (\hat{r}_i \cdot \hat{r}_{i+1}) \cdots (\hat{r}_{j-1} \cdot \hat{r}_j)$$

To construct a recurrence relation, the sum $\langle h^2(N) \rangle_t$ must be related to the partial sum $\langle h^2(N-1) \rangle_t$. Recall that $\langle h^2(N) \rangle_t$ can be viewed as the element sum over a symmetric $N \times N$ matrix, the first right-hand-side term being the sum along the diagonal and the second term being the sum of the upper triangle. The difference between $\langle h^2(N) \rangle_t$ and $\langle h^2(N-1) \rangle_t$ is therefore the element sum of the N^{th} column. Accordingly, the recurrence relation is

$$\langle h^2(N) \rangle_t = \tau_N^2 + \langle h^2(N-1) \rangle_t + 2 \sum_{i=1}^{N-1} \tau_i \tau_N (\hat{r}_1 \cdot \hat{r}_2) \cdots (\hat{r}_{N-1} \cdot \hat{r}_N)$$

or more succinctly,

$$\langle h^2(N) \rangle_t = \tau_N^2 + \langle h^2(N-1) \rangle_t + 2\tau_N S_N \quad (8.2.61)$$

where

$$S_N = (\hat{r}_N \cdot \hat{r}_{N-1}) (\tau_{N-1} + (\hat{r}_{N-1} \cdot \hat{r}_{N-2}) (\tau_{N-2} + \cdots)) \quad (8.2.62)$$

Comparison of the frequency-average recurrence relation (8.2.56) and the time-average recurrence relation (8.2.61) shows that $\langle \tau^2(N) \rangle_\omega = \langle h^2(N) \rangle_t$. The second-moment of the impulse response equals mean-square of the DGD spectrum. This result of Gisin is the formal tie between the impulse and frequency response of birefringent concatenation.

Figure 8.33 shows three concatenation calculations in both the frequency and time domains. The number of segments is four, six, and eight. The left-hand figures plot $\tau(\omega)$ as a function of frequency; the right-hand figures plot $h(t)$, or the impulse response of the cascade. The impulse response is

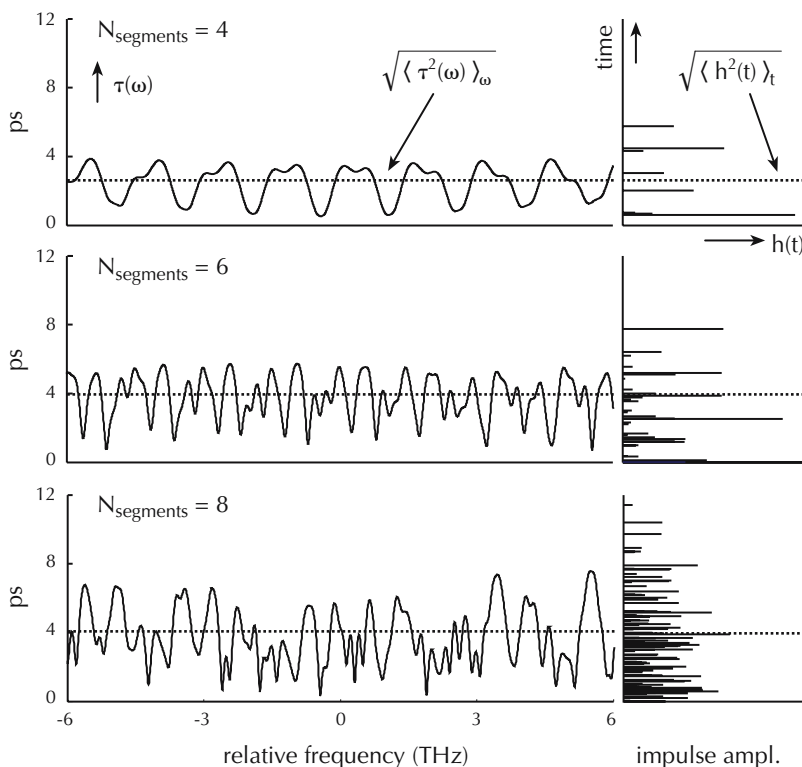


Fig. 8.33. Root-mean-square of DGD spectrum equals the standard deviation of its impulse response for lossless birefringent cascades of various lengths (after [19]). Calculated DGD spectra (left) and impulse responses for $t \geq 0$ (right) are shown for $N_{\text{seg}} = 4, 6, 8$. The dotted line shows the calculated rms DGD (left) and standard deviation (right) of the respective representations. The amplitude of the impulses as illustrated is the square-root of $w(k)$ to indicate all impulses more clearly. The birefringent vector orientations are randomly selected over a uniform distribution on the Poincaré sphere; the DGD values are randomly selected on a Maxwellian distribution where the underlying iid Gaussian distributions have a standard deviation of unity.

symmetric about $t = 0$ because $\hat{s}_{\text{in}} \cdot \hat{r}_1$ was set to zero – the plots shows only the positive side of the response. The calculated rms DGD in frequency and the standard deviation of the impulse spread in time are indicated by the dotted lines. It is remarkable how quickly the two averages converge as the number of segments increases.

Armed with the Gisin and Pellaux result a full circle can be closed on the triad of time-domain analyses of the preceding. The Signal Distortion section that started on page 343 covered the temporal extent of the PMD impulse response and the interference of co-polarized signal images that result from the

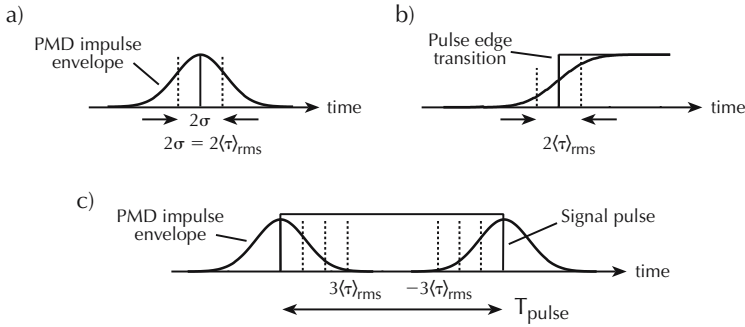


Fig. 8.34. Pulse-width broadening due to PMD. a) For a long, sufficiently random link, the PMD impulse response converges in distribution to a gaussian. Gisin and Pellaux show that the standard deviation is the rms DGD of the link $\sqrt{\langle\tau^2\rangle}$. b) Leading transition edge is broadened by the impulse response. c) Three standard deviations covers 99.9% of the gaussian. The center of a pulse can be distorted when three standard deviations of the impulse response equal half the pulse width.

convolution. In that section only simple impulse responses were considered. Now, the effects on a signal by the most general of impulse responses can be intuited.

For a long birefringent concatenation sufficiently random, the impulse response envelope converges in distribution to a gaussian shape; this is illustrated in Fig. 8.34(a). As a measure, the width of the gaussian is twice its standard deviation; but it is now known that the standard deviation is the rms DGD of the link. When a transition edge is convolved by the gaussian impulse response, the edge is characteristically broadened by the gaussian width which is twice the rms DGD (Fig. 8.34(b)). Moreover, one standard deviation covers only 84% of the envelope; two and three standard deviations include 97.7% and 99.9% of the gaussian envelope – the transition edges are likely to be broadened further than one standard deviation.

Since the impulse response is the inverse Fourier transform of the entire PMD spectrum, the impulse response does not depend on the particular DGD, SOPMD, and higher-order PMD experienced by the signal spectrum. The signal carrier may be frequency shifted to observe different aspects of the PMD spectrum, but the only effect on the convolution is how the signal images interfere. Therefore, the overlap between a signal “one” and the impulse-response gaussian envelope at the transition edges (Fig. 8.34(c)), remains fixed for any carrier frequency. The tails of the gaussians have reasonable probability three to four standard deviations into the signal; therefore some conditions of interference can excite distortion this far into the signal pulse. One can say roughly that when $T_{\text{pulse}}/2 \simeq 3\langle\tau\rangle_{\text{rms}}$, the center of the pulse can be effected by PMD. In the next chapter it is shown that for long fibers $\langle\tau\rangle \simeq 1.085\langle\tau\rangle_{\text{rms}}$, that is, the mean and rms with within 10% of one another. As a measure of complete channel loss due to distortion, the metric of

$$T_{\text{pulse}} \simeq 3 \langle \tau \rangle \quad \text{Metric of complete distortion}$$

is commonly used in the industry. For instance, a 10 Gb/s NRZ communication link has 100 ps time slots. The signal can be completely distorted when running on a fiber with accrued mean PMD of $\langle \tau \rangle \simeq 33$ ps.

Concluding Remarks on Time-Domain Studies

The preceding has intended to familiarize the reader with the cornerstone features of the time-domain representation. Many researchers have gone further than this text in the study of particular aspects. Heffner offers several papers on the time/frequency equivalence of PMD and particularly the impact of input signal chirp and coherence on the output distortion [23–25]. The moments analysis of Karlsson is extended by Shieh [57]. The correspondence between time and frequency domains in the study of second-order PMD is principally given by Gisin *et al.* [3], and separately by Penninckx *et al.* [13, 14]. The impact of higher orders of PMD has been studied by Gisin [30] and separately by Kogelnik *et al.* [40]. The degree of polarization as a function of PMD and optical data spectrum is analyzed by Nezam *et al.* [46].

Many researchers have studied the Jones-matrix representation of SOPMD and PMD in general. Significant works come from Eyal [9], Kogelnik [41], and Vincetti [47] for SOPMD; and Heismann [26], Penninckx [49], and Vincetti [12, 48] for general PMD.

8.2.7 Fourier Analysis of the DGD Spectrum

The Fourier analysis of the DGD spectrum offers yet another view into the nature of PMD. This Fourier analysis is not the PMD impulse response nor its Fourier transform into the associated vector and scalar spectra, but strictly a Fourier analysis of one scalar spectrum. The analysis can be extended to include the magnitude-SOPMD scalar spectrum. The purpose of this analysis is twofold. First, to go beyond second-order PMD, one can take higher and higher derivatives of the PMD vector, as reported by Kogelnik [40], or one can look at the degree of variation over a frequency window, as first analyzed by Poole and Favin [54]. In the former case, increased precision is found but over an infinitesimally narrow bandwidth; in the latter case, approximation is made but how the PMD spectrum overlays with the finite-width spectrum of a signal can be better understood.

The Poole and Favin analysis showed that in the long-length regime the expected number of level crossings $\langle N_m \rangle$ of the DGD spectrum through a given level over an interval $\Delta\omega$ is

$$\langle N_m \rangle = \frac{\langle \tau \rangle \Delta\omega}{4}$$

The number of level crossings increase linearly with the mean DGD of the link. Therefore the variation of the DGD spectrum must increase linearly

with mean DGD as well. For instance, if one conservatively wants to place a channel between two adjacent level crossing, one would take $\langle N_m \rangle = 1$ and estimate the maximum $\langle \tau \rangle$ as $\langle \tau \rangle_{\max} \simeq 4/\Delta\omega$. For a channel bandwidth of $\Delta f = 20$ GHz, the maximum $\langle \tau \rangle$ is $\langle \tau \rangle_{\max} \simeq 33$ ps.

The following analysis is presented for the short-length regime [5, 6, 64]. Its extension to the long-length regime may be possible but to date this has not been done. The purpose of presenting this analysis is to emphasize the origin of oscillatory variation in the DGD spectrum and to exhibit the spin-vector formalism used to arrive at the conclusions. In light of the Gisin and Pellaux time-response derivation based on recurrence relations, it is believed a similar approach can be used to extend the Fourier analysis into the long-length regime.

The problem at hand is similar to that of angular momentum. Analyses of angular momentum relate to coupled spinning objects or particles and the total overall momentum. Often one looks for the probability density of the overall angular momentum given all possible orientations of the component spins. Alternatively, the extrema can be determined. The quantized angular momentum analysis of coupled subatomic spins, such as electron and nuclear spins, determines the quantized levels of the total angular momentum and the state densities.

PMD concatenations are similar because component PMD vectors precess, or spin, about the axes of adjacent component vectors as frequency is swept. While the probability density of the overall PMD pointing direction or PMD vector length can be evaluated, the focus of the present analysis is to determine the Fourier components embedded in the variation of the PMD vector length when frequency is swept. The embedded Fourier components depend only on the delays of the component PMD vectors and not on their relative orientation or the frequency. The amplitude and phase of the Fourier components, however, do determine on these details.

A note on nomenclature. The Fourier content determined in the following refers to the oscillatory rate of a DGD spectrum, that is, the frequency of variation. The use of “frequency” as related to Fourier content differs from the use of “frequency” as related to the carrier frequency of a probe signal that measures the PMD. The optical carrier frequency is completely different from the frequencies of the Fourier content of the DGD spectrum.

The following analysis builds DGD spectra and the respective Fourier components from two, three, and four birefringent stages. The concatenation rules for each are illustrated in Fig. 8.35 and the corresponding DGD spectra and Fourier analysis are illustrated in Fig. 8.36.

The simplest concatenation is that of two vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ (Fig. 8.35(a)). The angle between the two vectors in the diagram is determined by the mode mixing angle between the stages and is not a function of frequency. The output vector is

$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1 \quad (8.2.63)$$

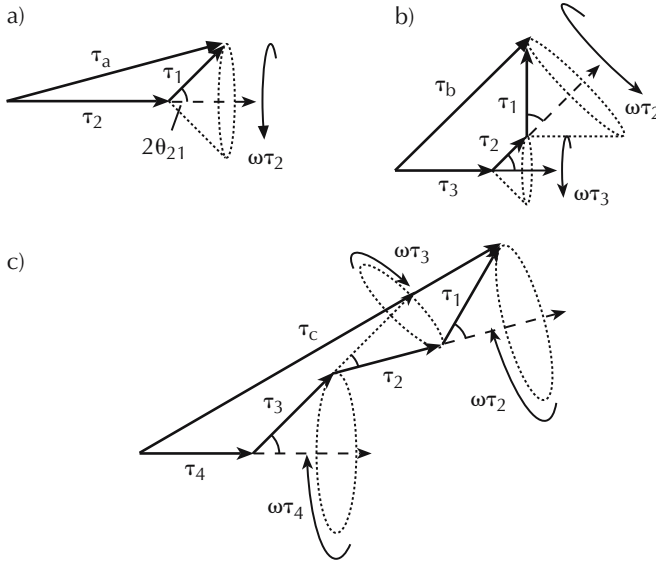


Fig. 8.35. Component PMD vector concatenations for two, three, and four stages. a) Two stages. Angle θ_{21} is determined by the mode mixer and is frequency independent. Vector $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ at rate τ_2 . The PMD vector is the sum of its component vectors; the length τ_a is the DGD. b) Three stages, two different precessions: $\omega\tau_2$ and $\omega\tau_3$. c) Four stages, three different precessions: $\omega\tau_2$, $\omega\tau_3$, and $\omega\tau_4$.

where $\vec{\tau}_k = \tau_k \hat{r}_k$, $k = 1, 2$. Component $\vec{\tau}_1$ precesses about $\vec{\tau}_2$ with birefringent phase $\varphi = \omega\tau_2$; the free-spectral range is $\text{FSR} = 1/\tau_2$. A note on the order of precession. The figure shows $\vec{\tau}_1$ precessing about $\vec{\tau}_2$, although in the concatenation $\vec{\tau}_1$ comes first. Physically, the cumulative PMD vector $\vec{\tau}$ is defined at the output. Looking from the output toward the input, one sees $\vec{\tau}_2$ immediately and $\vec{\tau}_1$ through the aperture of the second element that generates $\vec{\tau}_2$. When the frequency is changed, the appearance of Stokes orientation $\vec{\tau}_1$ is rotated due to the birefringence of $\vec{\tau}_2$, precisely in the same way a polarization state is altered due to the birefringence of $\vec{\tau}_2$. This gives the precession of $\vec{\tau}_1$ about $\vec{\tau}_2$.

The magnitude-squared of the DGD spectrum is

$$\vec{\tau} \cdot \vec{\tau} = \tau_2^2 + \tau_1^2 + 2\tau_2\tau_1 \hat{r}_2 \cdot \hat{r}_1 \quad (8.2.64)$$

The dot product in the last term,

$$\hat{r}_2 \cdot \hat{r}_1 = \cos \theta_{21}, \quad (8.2.65)$$

is frequency independent and θ_{21} is a Stokes angle. Since there is no frequency dependence in the DGD spectrum, the spectrum can be characterized as

$$\vec{\tau} \cdot \vec{\tau} = a_0 \quad (8.2.66)$$

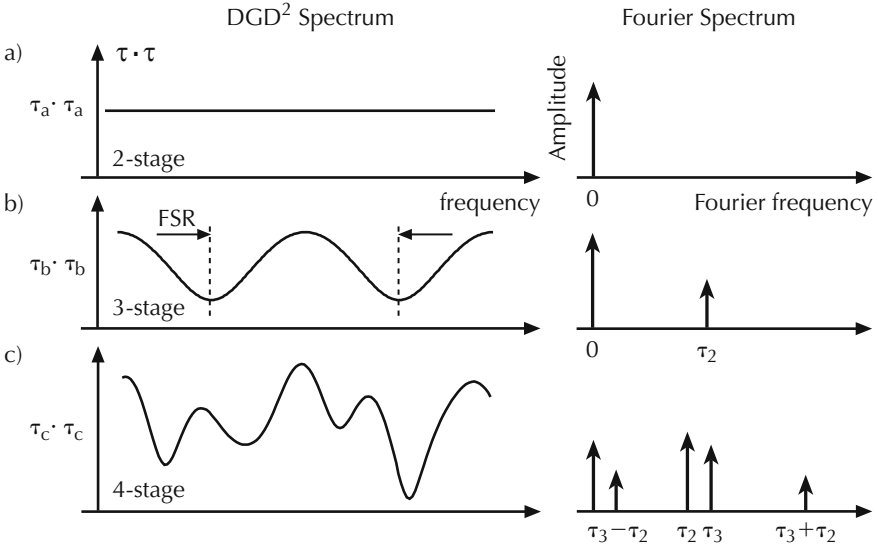


Fig. 8.36. Magnitude-square DGD spectra and associated Fourier analysis corresponding to precession diagrams in Fig. 8.35. The magnitude-square DGD spectra plot $\tau \cdot \tau$ as a function of carrier frequency. The Fourier analyses show the Fourier frequencies that are present in the respective $\tau \cdot \tau$ spectra. Only Fourier amplitudes are shown, although the Fourier phases are not necessarily zero. a) Two-stage spectrum has no oscillatory Fourier components and is governed by (8.2.66). Only a DC Fourier component is present. b) Three-stage spectrum has single oscillatory component with well-defined FSR and is governed by (8.2.70). The Fourier spectrum has a DC plus a single oscillatory component at τ_2 ; only the center stage dictates the frequency of oscillation. c) Four-stage spectrum has four oscillatory components, governed by (8.2.75). The Fourier spectrum has one constant plus four oscillatory components, including sum and different terms. Only the center two stage delays contribute to the oscillation.

where a_0 is a real number and the Fourier content is only DC, see Fig. 8.36(a).

The next case is the concatenation of three component vectors $\vec{\tau}_1$, $\vec{\tau}_2$, and $\vec{\tau}_3$ (Fig. 8.35(b)). As illustrated, the birefringent axes between the stages are not aligned, which results in mode mixing. The resultant PMD vector is

$$\vec{\tau} = \vec{\tau}_3 + R_3 \vec{\tau}_2 + R_3 R_2 \vec{\tau}_1 \quad (8.2.67)$$

The figure shows the motion of the three vector components. Vector $\vec{\tau}_1$ precesses about the $\vec{\tau}_2$ axis with birefringent phase $\varphi_2 = \omega\tau_2$. Vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ combined precess about the $\vec{\tau}_3$ axis with birefringent phase $\varphi_3 = \omega\tau_3$. The length and pointing direction of $\vec{\tau}$ exhibit a more complicated motion than that of the two-stage example and are in general frequency dependent. The magnitude-squared of the DGD spectrum is

$$\begin{aligned}
\vec{\tau} \cdot \vec{\tau} &= \tau_3^2 + \tau_2^2 + \tau_1^2 + \\
&2\tau_3\tau_2 \hat{r}_3 \cdot \hat{r}_2 + 2\tau_2\tau_1 \hat{r}_2 \cdot \hat{r}_1 + \\
&2\tau_3\tau_1 \hat{r}_3 \cdot (R_2\hat{r}_1)
\end{aligned} \tag{8.2.68}$$

In light of (8.2.65), the first five terms on the right-hand side are frequency independent. The last term, however, generates one non-DC Fourier component. The last term expands to

$$\begin{aligned}
\hat{r}_3 \cdot (R_2\hat{r}_1) &= \cos \theta_{32} \cos \theta_{21} + \\
&\sin \varphi_2 \hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_1) - \cos \varphi_2 \hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_2 \times \hat{r}_1)
\end{aligned} \tag{8.2.69}$$

The last two terms on the right-hand side add to yield a single oscillatory term governed by φ_2 (see Appendix A). Combining Eqs. (8.2.65), (8.2.69), and using the identity

$$A \cos \phi \pm B \sin \phi = \sqrt{A^2 + B^2} \cos(\phi \mp \tan^{-1} B/A)$$

yields the Fourier form of the magnitude-squared DGD spectrum:

$$\vec{\tau} \cdot \vec{\tau} = a_0 + a_1 \cos(\varphi_2 - \xi) \tag{8.2.70}$$

where, as before, a_0 and a_1 are real numbers that are independent of frequency. The spectrum is periodic where the periodicity is determined solely by the center section τ_2 , see Figs. 8.36(b). The Fourier-component phase shift ξ is determined from (8.2.69):

$$\xi = \tan^{-1} \left[\frac{\hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_1)}{\hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_2 \times \hat{r}_1)} \right] \tag{8.2.71}$$

where in general the phase changes as the relative angles between PMD components change. Only when all three birefringent axes lie in the same plane, which leads to $\hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_1) = 0$, does phase shift ξ vanish.

The coupling of Fourier-component phase shift ξ to the mode mixing angles $\hat{r}_2 \cdot \hat{r}_1$ and $\hat{r}_3 \cdot \hat{r}_2$ is an interesting effect that is illustrated in Fig. 8.37. Both three-stage examples in the figure show a range of DGD spectral shapes when the center section is rotated as indicated. When the birefringent axes $\hat{r}_{1,2,3}$ of all three sections lie in the same plane, such as the equator, then Fourier-component phase shift ξ is identically zero. In this case the frequency location of the maximum DGD value does not change even though the shape of the spectrum changes (Fig. 8.37(a)). However, when any one of the birefringent axes lies out of the plane of the other two then coupling between phase shift ξ and mode mixing occurs (Fig. 8.37(b)). This effect is observable, for instance, when a zero-order quarter-wave waveplate is inserted to either side of the center element. In a communication link, the bulk components such as isolators can make the apparent birefringent axes fall outside of a common plane.

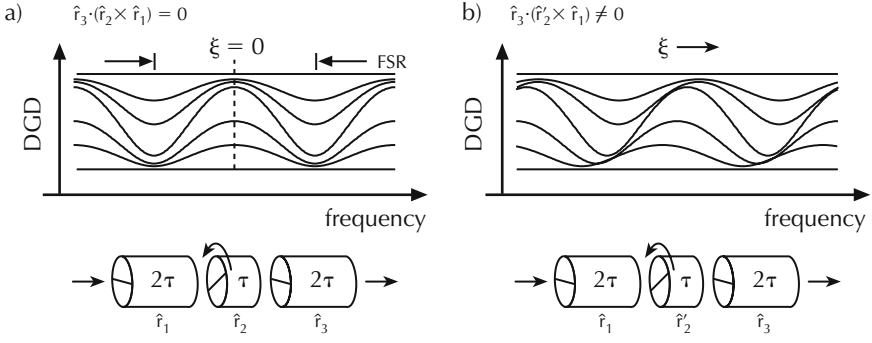


Fig. 8.37. Fourier-component phase shift ξ as decoupled (a) and coupled (b) to mode mixing. a) Locus of DGD spectra for a three-stage system as the center section is rotated. All three birefringent axes lie in the same plane in Stokes space. The frequency location of the maximum DGD is fixed for each spectrum. b) Locus of DGD spectra when one birefringent axes lies out of the plane defined by the other two. Fourier-component phase shift ξ is coupled to the mode mixing.

The following analysis is simplified by assuming all birefringent axes lie in the same plane. This is called the co-planar assumption. Violation of this assumption will not change the location or number of Fourier frequencies, only their phase.

The concatenation of four sections $\vec{\tau}_1$, $\vec{\tau}_2$, $\vec{\tau}_3$, and $\vec{\tau}_4$ is illustrated in Fig. 8.35(c). The resultant PMD vector is

$$\vec{\tau} = \vec{\tau}_4 + R_4 \vec{\tau}_3 + R_4 R_3 \vec{\tau}_2 + R_4 R_3 R_2 \vec{\tau}_1 \quad (8.2.72)$$

The motion of the vector sum is more complicated yet. Vector $\vec{\tau}_1$ precesses about the $\vec{\tau}_2$ axis with phase $\varphi_2 = \omega\tau_2$; vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ combined precess about $\vec{\tau}_3$ with phase $\varphi_3 = \omega\tau_3$; and vectors $\vec{\tau}_{1,2,3}$ combined precess about $\vec{\tau}_4$ with phase $\varphi_4 = \omega\tau_4$. The magnitude-squared DGD spectrum takes the form

$$\begin{aligned} \vec{\tau} \cdot \vec{\tau} = & \sum_{k=1}^4 \tau_k^2 + 2 \sum_{k=1}^3 \tau_{k+1} \tau_k \hat{r}_{k+1} \cdot \hat{r}_k + \\ & 2 \sum_{k=1}^2 \tau_{k+2} \tau_k \hat{r}_{k+2} \cdot (R_{k+1} \hat{r}_k) + 2\tau_4 \tau_1 \hat{r}_4 \cdot (R_3 R_2 \hat{r}_1) \end{aligned} \quad (8.2.73)$$

The first term on the right-hand side of (8.2.73) is scalar; the second term, identified with (8.2.65), generates no frequency-dependent terms; and the third term, identified with (8.2.69), generates the frequency-dependent terms $\cos \varphi_2$ and $\cos \varphi_3$ (assuming coplanar mode-mixing vectors). The last term generates additional frequency-dependent components. That term expands to

$$\begin{aligned}
\hat{r}_4 \cdot (R_3 R_2 \hat{r}_1) &= \cos \theta_{43} \cos \theta_{32} \cos \theta_{21} \\
&- \hat{r}_3 \cdot (\hat{r}_2 \times \hat{r}_2 \times \hat{r}_1) \cos \theta_{43} \cos \varphi_2 \\
&- \hat{r}_4 \cdot (\hat{r}_3 \times \hat{r}_3 \times \hat{r}_2) \cos \theta_{21} \cos \varphi_3 \\
&+ \hat{r}_4 \cdot (\hat{r}_3 \times \hat{r}_2 \times \hat{r}_1) \sin \varphi_3 \sin \varphi_2 \\
&+ \hat{r}_4 \cdot (\hat{r}_3 \times \hat{r}_3 \times \hat{r}_2 \times \hat{r}_2 \times \hat{r}_1) \cos \varphi_3 \cos \varphi_2
\end{aligned} \tag{8.2.74}$$

The mixing products $\sin \varphi_3 \sin \varphi_2$ and their cosine complements resolve themselves into sum and difference terms, e.g.

$$2 \sin \varphi_3 \sin \varphi_2 = \cos(\varphi_3 - \varphi_2) - \cos(\varphi_3 + \varphi_2)$$

The Fourier content of a four-stage concatenation therefore takes the form

$$\begin{aligned}
\vec{\tau} \cdot \vec{\tau} &= a_0 + a_1 \cos \omega \tau_2 + a_2 \cos \omega \tau_3 + \\
&\quad a_3 \cos \omega(\tau_3 - \tau_2) + a_4 \cos \omega(\tau_3 + \tau_2)
\end{aligned} \tag{8.2.75}$$

where, as before, coefficients a_k are real, frequency-independent numbers that are determined solely by the mode mixing between stages. An exemplar spectra and its Fourier spectrum are illustrated in Figs. 8.36(c).

Several observations about (8.2.75) are made. First, the Fourier content of the magnitude-squared DGD spectrum is determined only by the center two delay stages. In general the first and last section for any number of stages do not contribute to the Fourier content; this observation is proven below. Second, sum and difference Fourier terms are present in addition to the Fourier components associated with the center two stages. Here there are four Fourier components in total. Third, there is a Fourier-component generator function that generates these components. The generator function $\mathcal{G}(N)$ is

$$\mathcal{G}(N) = \hat{r}_N \cdot R(N, 2) \hat{r}_1 \tag{8.2.76}$$

where the cumulative operator $R(N, 2)$ is defined by (8.2.38) on page 337. The first four generator functions are

$$\begin{aligned}
\mathcal{G}(1) &= 1 \\
\mathcal{G}(2) &= g_0 \\
\mathcal{G}(3) &= g_0 + g_1 \cos \varphi_2 \\
\mathcal{G}(4) &= g_0 + g_1 \cos \varphi_2 + g_2 \cos \varphi_3 + \\
&\quad g_3 \cos(\varphi_3 - \varphi_2) + g_4 \cos(\varphi_3 + \varphi_2)
\end{aligned} \tag{8.2.77}$$

where the value g_k for one generator function has no relation to the value g_k of another generator function.

As a last part to this section, the absence of Fourier components generated from the first and last stages is shown [61]. The independence of the first stage has already been demonstrated: it is clear from any of the vector diagrams

in Fig. 8.35 that no vector precesses about $\vec{\tau}_1$, so there is no sensitivity of $\vec{\tau}$ to $\omega\tau_1$. The insensitivity to $\vec{\tau}_N$ is proven by separating the last stage from the preceding $N - 1$ stages:

$$\begin{aligned}\vec{\tau} \cdot \vec{\tau} &= (\vec{\tau}_N + R_N \vec{\tau}_{(N-1)}) \cdot (\vec{\tau}_N + R_N \vec{\tau}_{(N-1)}) \\ &= \tau_N^2 + \tau_{(N-1)}^2 + 2\tau_N \tau_{(N-1)} \hat{r}_N \cdot \hat{r}_{(N-1)}\end{aligned}\quad (8.2.78)$$

Since the last term on the right-hand side has the form $\mathcal{G}(2)$, there is in fact no φ_N Fourier component generated by the last stage. Geometrically this makes sense because rotation about $\vec{\tau}_N$ pirouettes the remaining vector structure, changing its pointing direction but not its length.

8.3 Combined Effects of PMD and PDL

When birefringent elements and partial polarizers are interspersed along a concatenation, the resulting polarization effects are more complicated than PDL or PMD alone. The resulting effects form a superset of PDL and PMD: communication impairments due to combined PDL and PMD can be more severe than either isolated effect. Since there is no separate name for the combined phenomena, “PMD and PDL” here denotes the aggregate effect.

The principal results from combined PMD and PDL are

1. The principal-states of polarization are not orthogonal to one another.
2. The output polarization state does not follow simple precessional motion as a function of frequency.
3. The polarization-dependent loss is frequency dependent. A new term, differential-attenuation slope (DAS) is introduced to characterize this effect.

These results conspire to create anomalous distortions. For instance, the differential-group delay of a PMD and PDL combination can be greater than the sum of the individual delays [32]. Or, a pulse can suffer spreading even with zero net DGD [30]. Moreover, neither an optical PDL nor PMD compensator can be perfectly realized.

The original work on PMD and PDL is by Frigo [16], who derived the equation of motion of the output polarization vector as function of frequency, albeit in the context of coupled-mode equations. Eyal [10] later recast the Frigo equation into unit-vector form to derive the complex motion of the output polarization state with frequency. The authors Gisin, Huttner, and Geiser [18, 31] discovered many anomalous effects due to the interaction of PMD and PDL. R. C. Jones also contributed on this subject: his 1941 paper proves a theorem, theorem 3, that any number of PMD and PDL elements (although he called them retarders and partial polarizers) can be replaced with just four elements – two PMD elements, one PDL element, and one

optically active polarization rotator [36]. This theorem has practical use when separating PMD and PDL effects.

Further research on the interaction of PMD and PDL in an optical communications link has been reported by Mollenauer [62, 63], Feced [11], and Eyal [8]. A very interesting measurement method to test for maximum eye-opening excursion has been invented by Kuperman *et al.* [42].

Three principal equations are derived in the following: the change of output polarization state with frequency; the change of output polarization state through propagation; and the cumulative PDL vector equation of motion. Table 8.3 compares the expressions for pure PMD and those including PDL.

8.3.1 Frequency-Dependence of the Polarization State

In general, the transformation matrix T between input and output polarization states is neither unitary nor Hermitian in the presence of combined PMD and PDL. Due to the birefringence in the medium, T is frequency dependent: $|t\rangle = T(\omega)|s\rangle$. As before, for $|s\rangle$ fixed in frequency, the frequency-change of the output state is $|t\rangle_\omega = T_\omega|s\rangle$. As long as a perfect polarizer is not placed between input and output, T^{-1} exists and the change of output state is expressed as

$$|t\rangle_\omega = T_\omega T^{-1}|t\rangle \quad (8.3.1)$$

Even without any particular reference to PMD and PDL parameters, the matrix $T_\omega T^{-1}$ can always be decomposed into a Hermitian, skew-Hermitian, and trace components, as was shown by (2.5.72) on page 61. The decomposition of $T_\omega T^{-1}$ takes the form

$$jT_\omega T^{-1} = \frac{1}{2} \left(\vec{\Omega}_r \cdot \vec{\sigma} + j \vec{\Omega}_i \cdot \vec{\sigma} + a_0 I \right) \quad (8.3.2)$$

The factor of j in front of $T_\omega T^{-1}$ is added to keep the notation parallel to the Hermitian operator $jU_\omega U^\dagger$ defined for the description of PMD. The scalar a_0 is in general complex; the real component is the common phase through the system and the imaginary part is the common loss. The vectors $\vec{\Omega}_r$ and $\vec{\Omega}_i$ are real-valued Stokes vectors. These vectors relate to the system birefringence and differential attenuation, but not in a straight-forward way. The decomposition parameters are related to $jT_\omega T^{-1}$ via (2.5.68) on page 61 for the trace and (2.5.73) for the real and imaginary parts.

Since T is not necessarily unitary, the length of the output Stokes vector $\vec{t}$ is generally not the same as the input vector $\vec{s}$. One must be careful to distinguish the output vector length t and the unit vector $\hat{t}$. With this remark in mind, the Stokes vector $\vec{t}_\omega$ is calculated in the same way as (8.2.15) on page 330. This makes

$$\vec{t}_\omega = \langle t | (T_\omega T^{-1})^\dagger \vec{\sigma} + \vec{\sigma} (T_\omega T^{-1}) | t \rangle \quad (8.3.3)$$

Substitution of $T_\omega T^{-1}$ (8.3.2) into (8.3.3) generates the Frigo equation of motion:

$$\frac{d\vec{t}}{d\omega} = \vec{\Omega}_r \times \vec{t} + a_{i,0} \vec{t} + t \vec{\Omega}_i \quad (8.3.4)$$

where $a_{i,0}$ is the imaginary part of the trace of $T_\omega T^{-1}$. This equation describes a complex behavior of the output Stokes vector $\vec{t}$. The first term on the right-hand side generates a precessional motion: $\vec{t}$ precesses about $\vec{\Omega}_r$ as a function of frequency. In the absence of PDL, $\vec{\Omega}_r = \vec{\tau}$, the PMD vector. In the presence of PDL, $\vec{\Omega}_r$ includes PDL as well as PMD terms. The second term on the right-hand side describes the growth or decay of $\vec{t}$ along its own axis. The imaginary part of the trace of $T_\omega T^{-1}$ governs this behavior. Also, these first two terms run perpendicular to one another. Finally, the third term on the right-hand side pulls $\vec{t}$ toward $\vec{\Omega}_i$. The pulling behavior has been seen before in §8.1.2 in regard to PDL. In sum, there are three distinct axes along which the output state is changed.

There is a competition setup between $\vec{\Omega}_r$ and $\vec{\Omega}_i$. If the former is the dominant term, then it acts to retard the growth or decay of $\vec{t}$ by generating a motion perpendicular to $\vec{t}$. If the latter term dominates, then $\vec{t}$ grows or decays without bound.

Further insight is found by decomposing (8.3.4) into coupled unit-vector and vector-length equations of motion [15]. The decomposition requires two identifications. First, by definition $\vec{t} = t\hat{t}$, so the frequency derivative is

$$\frac{d\vec{t}}{d\omega} = t \frac{d\hat{t}}{d\omega} + \hat{t} \frac{dt}{d\omega}$$

Note that the first term is perpendicular to $\vec{t}$ while the second term is parallel (that the first term is perpendicular is a consequence of $\hat{t}$ being a unit vector). Second, the vector $\vec{\Omega}_i$ is decomposed into components parallel and perpendicular to $\vec{t}$:

$$\begin{aligned} \vec{\Omega}_i &= \vec{\Omega}_{i,\parallel} + \vec{\Omega}_{i,\perp} \\ &= (\vec{\Omega}_i \cdot \hat{t}) \hat{t} - (\vec{\Omega}_i \cdot \hat{t}) \hat{t} + \vec{\Omega}_i (\hat{t} \cdot \hat{t}) \\ &= (\vec{\Omega}_i \cdot \hat{t}) \hat{t} + (\hat{t} \times \vec{\Omega}_i) \times \hat{t} \end{aligned}$$

Substitution into the equation of motion makes

$$t \frac{d\hat{t}}{d\omega} + \hat{t} \frac{dt}{d\omega} = t \vec{\Omega}_r \times \hat{t} + (a_{i,0}t) \hat{t} + t (\vec{\Omega}_i \cdot \hat{t}) \hat{t} + t (\hat{t} \times \vec{\Omega}_i) \times \hat{t}$$

The parallel and perpendicular components must separately satisfy this equation. Therefore the decomposition produces the two coupled equations

$$\frac{dt}{d\omega} = a_{i,0}t + (\vec{\Omega}_i \cdot \hat{t}) t \quad (8.3.5a)$$

$$\frac{d\hat{t}}{d\omega} = \vec{\Omega}_r \times \hat{t} - (\vec{\Omega}_i \times \hat{t}) \times \hat{t} \quad (8.3.5b)$$

These equations exhibit interesting properties. First, unit-vector equation is not coupled to the vector-length equation. Therefore these equations can be solved in sequence: first solve for $\hat{t}$, then for t . Second, the unit-vector equation of motion has both double- and triple-vector products. This result was first published by Eyal [10]. Third, the change of length equation is driven explicitly by $a_{i,0}$ and $\vec{\Omega}_i$, both directly related to PDL. If the PDL were zero, these terms would be absent and the vector length would be invariant. With PDL present, t changes with frequency – this is the origin of the differential-attenuation slope, or DAS. Even when the common loss generated by $a_{i,0}$ is transformed out, the differential attenuation embedded in $\vec{\Omega}_i$ induces DAS.

8.3.2 Non-Orthogonality of PSP's

The principal states of polarization for PMD and PDL are found in the same way as the PSP's are for pure PMD. Recall from (8.2.9) on page 329 that the PSP's are defined by the eigenvalue equation of the operator $jU_\omega U^\dagger$. When this equation is satisfied, the output polarization state is stationary to first-order in frequency.

In an entirely analogous way, the eigenvalue equation for $jT_\omega T^{-1}$ is defined. Substitution of the spin-vector form of $jT_\omega T^{-1}$ into (8.3.1) yields

$$|t\rangle_\omega = -\frac{j}{2} \left(\vec{\Omega}_r \cdot \vec{\sigma} + j \vec{\Omega}_i \cdot \vec{\sigma} \right) |t\rangle - j(a_0/2) |t\rangle \quad (8.3.6)$$

To make the output state stationary, the spin-vector operator must collapse to a complex scalar value:

$$\vec{\Omega} \cdot \vec{\sigma} |\tilde{p}_\pm\rangle = \pm \lambda |\tilde{p}_\pm\rangle \quad (8.3.7)$$

where $\vec{\Omega} = \vec{\Omega}_r + j\vec{\Omega}_i$. That the eigenvalues λ are equal and opposite is a direct consequence of the zero trace of $\vec{\Omega}$. Moreover, since the operator $\vec{\Omega} \cdot \vec{\sigma}$ is non-Hermitian, the eigenvalues are in general complex. Gisin *et al.* identify the real and imaginary parts of the eigenvalues as [31]

$$\lambda = \tau + j\eta \quad (8.3.8)$$

The real part τ is the familiar differential-group delay magnitude; the imaginary part η is the differential-attenuation slope (DAS), which is the frequency derivative of the differential attenuation along the two eigenvectors.

The eigenvalue λ and the operator $\vec{\Omega} \cdot \vec{\sigma}$ are related by

$$\lambda^2 = \vec{\Omega} \cdot \vec{\Omega} \quad (8.3.9)$$

This expression may be confirmed by solving $\det(\vec{\Omega} \cdot \vec{\sigma} - \lambda I) = 0$ and is analogous to the case for pure PMD where $\vec{\tau} \cdot \vec{\tau} = \tau^2$.

In the presence of PDL, the PSP's are not orthogonal (except, possibly, in transient or pathological cases). This is due to the complex value of the

operator $\vec{\Omega}$. The overlap γ^2 of the two eigenvectors can be computed in Stokes space from the dot-product $\hat{\tilde{p}}_+ \cdot \hat{\tilde{p}}_-$, see (2.5.65) on page 60. The calculation is simplified by rearranging (8.3.7) so that the operator has unit length and the eigenvalues are real:

$$\hat{\Omega} \cdot \vec{\sigma} |\tilde{p}_{\pm}\rangle = \pm |\tilde{p}_{\pm}\rangle$$

where $\hat{\Omega} = \vec{\Omega}/\lambda$, $\hat{\Omega} = \vec{w}_r + j\vec{w}_i$, and $\hat{\Omega} \cdot \hat{\Omega} = 1$. From this eigenvalue equation two auxiliary equations are computed, the first by multiplying the equation by $\langle \tilde{p}_{\pm} |$ and the second by $\langle \tilde{p}_{\pm} | \vec{\sigma}$:

$$\begin{aligned} \hat{\Omega} \cdot \langle \tilde{p}_{\pm} | \vec{\sigma} | \tilde{p}_{\pm} \rangle &= \pm \langle \tilde{p}_{\pm} | \tilde{p}_{\pm} \rangle \\ \langle \tilde{p}_{\pm} | \vec{\sigma} (\hat{\Omega} \cdot \vec{\sigma}) | \tilde{p}_{\pm} \rangle &= \pm \langle \tilde{p}_{\pm} | \vec{\sigma} | \tilde{p}_{\pm} \rangle \end{aligned}$$

Conversion to Stokes space gives

$$\hat{\Omega} \cdot \hat{p}_{\pm} = \pm 1, \text{ and } \hat{\Omega} + j\hat{\Omega} \times \hat{p}_{\pm} = \pm \hat{p}_{\pm} \quad (8.3.10)$$

where the tilde has been removed for brevity. Now, since $\hat{\Omega} \cdot \hat{\Omega} = 1$, the imaginary part of the dot product must vanish: this requires $\vec{w}_r \cdot \vec{w}_i = 0$. Therefore an orthogonal group of three (unnormalized) axes may be constructed, $(\vec{w}_r, \vec{w}_i, \vec{w}_r \times \vec{w}_i)$, and $\hat{p}_{\pm}$ can be projected onto this basis:

$$\hat{p}_{\pm} = c_r \vec{w}_r + c_i \vec{w}_i + c_{\times} (\vec{w}_r \times \vec{w}_i) \quad (8.3.11)$$

The real-valued coefficients are isolated through the dot products

$$\begin{aligned} c_r \vec{w}_r \cdot \vec{w}_r &= \vec{w}_r \cdot \hat{p}_{\pm} \\ c_i \vec{w}_i \cdot \vec{w}_i &= \vec{w}_i \cdot \hat{p}_{\pm} \\ c_{\times} (\vec{w}_r \times \vec{w}_i) \cdot (\vec{w}_r \times \vec{w}_i) &= (\vec{w}_r \times \vec{w}_i) \cdot \hat{p}_{\pm} \end{aligned}$$

Given that $\hat{\Omega} \cdot \hat{p}_{\pm} = \pm 1$, the first two coefficients are $c_r = 1/(\vec{w}_r \cdot \vec{w}_r)$ and $c_i = 0$. The third coefficient is evaluated from the dot-product and the second auxiliary equation (8.3.11)

$$c_{\times} = \frac{(\vec{w}_r \times \vec{w}_i) \cdot \hat{p}_{\pm}}{w_r^2 w_i^2} \longrightarrow c_{\times} = \frac{1}{w_r^2}$$

The normalized eigenvectors in Stokes space are then [31]

$$\hat{p}_{\pm} = \frac{\pm \vec{w}_r + \vec{w}_r \times \vec{w}_i}{\vec{w}_r \cdot \vec{w}_r} \quad (8.3.12)$$

Using the fact that $\hat{\Omega} \cdot \hat{\Omega}^* = w_r^2 + w_i^2$ and $\vec{\Omega} \cdot \vec{\Omega}^* = |\lambda|^2$, the overlap of the eigenvectors is computed as

$$\begin{aligned} \gamma^2 &= \frac{1 + \hat{p}_+ \cdot \hat{p}_-}{2} = \frac{w_r^2 + w_i^2 - 1}{w_r^2 + w_i^2 + 1} \\ &= \frac{|\vec{\Omega}|^2 - |\lambda|^2}{|\vec{\Omega}|^2 + |\lambda|^2} \end{aligned} \quad (8.3.13)$$

Clearly when $\vec{\Omega}$ is real, the case for pure PMD, the overlap integral vanishes. However, addition of any PDL at all pulls the two PSP's away from an orthogonal orientation.

8.3.3 PMD and PDL Evolution Equations

There are two evolution equations to derive, both being extensions of the pure PMD and pure PDL case. First, the evolution of the complex operator $\vec{\Omega}$ as a function of length is derived; the analogue to this equation is (8.2.39) on page 339, although here a different derivation is employed. Second, the evolution of the cumulative PDL $\vec{\Gamma}$ is derived; the analogue is (8.1.27) on page 310. In both cases, the combined effects of birefringence and PDL are accounted for.

Earlier, the evolution equation for $\vec{\tau}$ was derived by combining the partial derivatives of the output state $\vec{t}$ with respect to both length and frequency. This is the Poole method. The present situation is more difficult because both the vector direction and length vary with length and frequency. Instead, the method of Gisin *et al.* is used [18]. Their method is similar to that used in §8.1.4 except that sections are taken as discrete rather than in the continuum limit.

Given a transformation T such that $|t\rangle = T|s\rangle$, the transformation is partitioned into N homogeneous birefringent and lossy sections: $T = A_N T_N$, where A_N is the common loss and T_N represents the product of transformation matrices through N sections, $T_N = T_n T_{n-1} \dots T_1$. Capital subscripts denote section products and lower-case subscripts denote particular sections. The terms A_N and T_N are the discrete analogue to the continuous expression (8.1.22) on page 309.

To account for birefringence and loss, the spin-vector operator for each section is written as

$$T_n = \exp \left(\frac{(-j\omega\vec{\tau}_n + \vec{\alpha}_n) \cdot \vec{\sigma}}{2} \right) \quad (8.3.14)$$

This definition of T_n is not totally general because the vector direction of the loss and birefringence are aligned; this is a reasonable model because the origin of differential loss and birefringence (in the perturbation regime) is likely due to the same disturbance. A shorthand variable $\vec{g}$ is defined as $\vec{g}_n = -j\omega\vec{\tau}_n + \vec{\alpha}_n$, and $\vec{g} = g\hat{g}$.

The last element of the concatenation is separated from the remaining by writing $T_N = T_n T_{N-1}$. Given $T_{\omega,N} = T_{\omega,n} T_{N-1} + T_n T_{\omega,N-1}$, the operator $T_{\omega,N} T_N^{-1}$ may be written in incremental form as

$$T_{\omega,N} T_N^{-1} = \frac{-j(\vec{\tau}_n \cdot \vec{\sigma})}{2} + \exp \left(\frac{\vec{g}_n \cdot \vec{\sigma}}{2} \right) T_{\omega,N-1} T_{N-1}^{-1} \exp \left(\frac{-\vec{g}_n \cdot \vec{\sigma}}{2} \right)$$

With the identification $T_{\omega,k} T_k^{-1} = -j/2 (\vec{\Omega}_k \cdot \vec{\sigma})$, the operator is rewritten as

Table 8.3. Comparison of Pure PMD and Entangled PMD + PDL

PMD	PMD + PDL
$jU_\omega U^\dagger p\rangle = \pm\tau/2 p\rangle$	$jT_\omega T^{-1} \tilde{p}\rangle = \pm\lambda/2 \tilde{p}\rangle + a_0/2 \tilde{p}\rangle$
$jU_\omega U^\dagger = \frac{1}{2} (\vec{\tau} \cdot \vec{\sigma})$	$jT_\omega T^{-1} = \frac{1}{2} (\vec{\Omega} \cdot \vec{\sigma}) + a_0/2$
$(jU_\omega U^\dagger)^\dagger = jU_\omega U^\dagger$	$(jT_\omega T^{-1})^\dagger \neq jT_\omega T^{-1}$
$\vec{\tau}, \tau \rightarrow \text{real}$	$\vec{\Omega}, \lambda \rightarrow \text{complex}$
$\tau \rightarrow \text{DGD}$	$\lambda = \tau + j\eta \rightarrow \text{DGD} + \text{DAS}$
$\tau^2 = \vec{\tau} \cdot \vec{\tau}$	$\lambda^2 = \vec{\Omega} \cdot \vec{\Omega}$
$\hat{p} = \langle p \vec{\sigma} p \rangle \rightarrow \text{PSP}$	$\hat{p} = \langle p \vec{\sigma} p \rangle \rightarrow \text{PSP}$
$(1 + \hat{p}_+ \cdot \hat{p}_-)/2 = 0$	$(1 + \hat{p}_+ \cdot \hat{p}_-)/2 \geq 0$
$\partial\vec{\tau}/\partial z = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau}$	$\partial\vec{\Omega}/\partial z = \vec{\beta}_\omega + (\vec{\beta} + j\vec{\alpha}) \times \vec{\tau}$

$$\vec{\Omega}_N \cdot \vec{\sigma} = \vec{\tau}_n \cdot \vec{\sigma} + \exp\left(\frac{\vec{g}_n \cdot \vec{\sigma}}{2}\right) \left(\vec{\Omega}_{N-1} \cdot \vec{\sigma}\right) \exp\left(\frac{-\vec{g}_n \cdot \vec{\sigma}}{2}\right)$$

Use of the complex spin-vector operator expansion (2.5.8) on page 63, the relevant spin-vector identities, and identification of the embedded equation gives

$$\begin{aligned} \vec{\Omega}_N = & \vec{\tau}_n + \left(\vec{\Omega}_{N-1} \cdot \hat{g}_n\right) \hat{g}_n + \\ & \cosh g_n \left(\vec{\Omega}_{N-1} - \left(\vec{\Omega}_{N-1} \cdot \hat{g}_n\right) \hat{g}_n\right) - j \sinh g_n \left(\vec{\Omega}_{N-1} \times \hat{g}_n\right) \end{aligned}$$

To make the differential operator $\partial_z \vec{\Omega}_N$, a small increment of length is characterized by a small magnitude g_n , and the birefringence and PDL are written in per-length form as $\vec{\tau}(z)$ and $\vec{\alpha}(z)$. Moreover, recognize that $\tau(z) = \Delta n/c = \beta_\omega$. The resultant equation of motion for $\vec{\Omega}$ is

$$\frac{\partial \vec{\Omega}}{\partial z} = \vec{\beta}_\omega + \left(\vec{\beta} + j\vec{\alpha}(z)\right) \times \vec{\Omega} \quad (8.3.15)$$

In comparison with the pure PMD evolution equation (8.2.39), PDL adds to the curl of $\vec{\Omega}$ and drives the vector to a complex quantity. Li and Yariv have worked out the analytic solutions (8.3.15) in [43].

Regarding the cumulative PDL vector $\vec{\Gamma}$, the equation of motion is derived in the same way as that shown in §8.1.4 but with the transformation operator (8.3.14) substituted for that in (8.1.22). The equations of motion for $\vec{\Gamma}$ and the transmission of depolarized light are

$$\frac{d\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - (\vec{\alpha} \cdot \vec{\Gamma}) \vec{\Gamma} \quad (8.3.16a)$$

$$\frac{dT_{\text{depol}}}{dz} = (\vec{\alpha} \cdot \vec{\Gamma} - \alpha) T_{\text{depol}} \quad (8.3.16b)$$

While the depolarized transmission equation is the same once PMD is included, the cumulative PDL equation has a new term: the $\vec{\beta} \times \vec{\Gamma}$ generates a rotation of $\vec{\Gamma}$ about the local birefringence vector $\vec{\beta}$. This rotation is to be expected since linear birefringence always generates precessional motion in Stokes space.

8.3.4 Separation of PMD and PDL

In 1941 R.C. Jones showed that most any Jones matrix generated by any number of retarders and partial polarizers can always be reconstructed with two retarders and one partial polarizer such that

$$\mathbf{J}(\omega) = U(\omega)P(\omega)V(\omega) \quad (8.3.17)$$

where P represents a partial polarizer and U and V are unitary matrices [36]. In general each matrix is a function of optical frequency. The partial polarizer is a Hermitian matrix; accordingly it has real eigenvalues and perpendicular eigenvectors. Such a matrix can be decomposed in $H = SAS^\dagger$, where S is a matrix whose columns are the eigenvectors of H and Λ is a diagonal matrix whose entries are the corresponding eigenvalues.

The unitary operator to the left or right of P can be absorbed in the following way: decompose P and absorb one of its neighbors into a unitary matrix:

$$\begin{aligned} \mathbf{J}(\omega) &= USAS^\dagger V \\ &= UV(S^\dagger V)^\dagger \Lambda (S^\dagger V) \\ &= \tilde{U}(\omega) \tilde{P}(\omega) \end{aligned} \quad (8.3.18)$$

where $\tilde{U} = UV$ and $\tilde{P} = (S^\dagger V)^\dagger \Lambda (S^\dagger V)$. Likewise, $\mathbf{J} = \tilde{P}'\tilde{V}'$.

Consider a link composed of PMD and PDL. The transformation matrix can be written $T(\omega) = P(\omega)U(\omega)$. At any particular frequency all of the differential attenuation is concentrated in matrix P : the eigenvectors of P determine the Stokes direction of $\vec{\Gamma}$ and its eigenvalues are the maximum and minimum transmission. The PDL component can be isolated from T by taking advantage of the Hermitian properties of P :

$$TT^\dagger = P^2 = S\Lambda^2 S^\dagger \quad (8.3.19)$$

The eigenvectors of $TT^\dagger$ point in the direction of the cumulative PDL. Also, since the eigenvalues of P are real, the entries in Λ^2 must be positive. Note

that the magnitude and direction of the PDL vector is in general frequency dependent; the dependence is governed by the birefringence of the link which is concentrated in U . This shows that even with the decomposition PU , the PDL and PMD remain entangled.

The unitary matrix U is found once the PDL matrix P is calculated from $TT^\dagger$:

$$U(\omega) = P^{-1}(\omega)T(\omega) \quad (8.3.20)$$

Given U it is tempting to calculate the PMD properties from $jU_\omega U^\dagger$. For small PDL this form of U provides a correction to T for an the investigator who wants to isolate the PMD effects. Both Shtengel and Karlsson have reported using this correction [33, 39]. Although suitable to remove perturbations, one should keep in mind that τ generated from $jU_\omega U^\dagger$ is not the same as τ generated from $T_\omega T^{-1}$. Huttner *et al.* define an effective PMD τ_{eff} for $jU_\omega U^\dagger$ to highlight the fact that τ and τ_{eff} are two different quantities [31].

Table 8.4. Table of Important SOP, PDL, and PMD Relations**Evolution Equations**

SOP:	$\frac{d\hat{s}}{dz} = \vec{\beta} \times \hat{s}$
PDL:	$\frac{d\vec{\Gamma}}{dz} = \vec{\alpha} - (\vec{\alpha} \cdot \vec{\Gamma}) \vec{\Gamma}$
	$\frac{dT_{\text{depol}}}{dz} = (\vec{\alpha} \cdot \vec{\Gamma} - \alpha) T_{\text{depol}}$
PMD:	$\frac{d\vec{\tau}}{dz} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau}$
	$\frac{d\vec{\tau}_\omega}{dz} = \vec{\beta}_{\omega\omega} + \vec{\beta} \times \vec{\tau}_\omega + \vec{\beta}_\omega \times \vec{\tau}$
	$\frac{d\hat{s}}{d\omega} = \vec{\tau} \times \hat{s}$
PMD+PDL:	$\frac{d\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - (\vec{\alpha} \cdot \vec{\Gamma}) \vec{\Gamma}$
	$\frac{dT_{\text{depol}}}{dz} = (\vec{\alpha} \cdot \vec{\Gamma} - \alpha) T_{\text{depol}}$
	$\frac{d\hat{s}}{d\omega} = \vec{\Omega}_r \times \hat{s} - (\vec{\Omega}_i \times \hat{s}) \times \hat{s}$

Defining Expressions

SOP:	$ t\rangle = U s\rangle$
PDL:	$T_p = \langle t t \rangle = \langle s P^\dagger P s \rangle$
	$\rho_{\text{dB}} \equiv 10 \log_{10} \left(\frac{T_{\text{max}}}{T_{\text{min}}} \right)$
PMD:	$jU_\omega U^\dagger p_\pm\rangle = \pm \tau/2 p_\pm\rangle$
	$jU_\omega U^\dagger = \frac{1}{2} (\vec{\tau} \cdot \vec{\sigma})$
	$\vec{\tau} \times = R_\omega R^\dagger$
PMD+PDL:	$\vec{\Omega} \cdot \vec{\sigma} \tilde{p}_\pm\rangle = \pm (\tau + j\eta) \tilde{p}_\pm\rangle$

PMD Concatenation

$$\vec{\tau} = \sum_{k=1}^n R(n, k+1) \vec{\tau}_n$$

$$\vec{\tau}_\omega = \sum_{k=1}^n R(n, k+1) (\vec{\tau}_{n\omega} + \vec{\tau}_n \times \vec{\tau}(n))$$

$$R(n, k) = R_n R_{n-1} \cdots R_k$$

References

1. D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurement of the group-delay difference between the principal states of polarization on a low-birefringent terrestrial fiber cable," *Optics Letters*, vol. 12, no. 10, pp. 844–846, 1987.
2. A. J. Barlow, "Birefringence and polarization mode dispersion in spun single mode fibers," *Applied Optics*, vol. 20, no. 17, p. 2962, 1981.
3. P. Ciprut, B. Gisin, N. Gisin, R. Passy, J. Weid, F. Prieto, and C. W. Zimmer, "Second-order polarization mode dispersion: Impact on analog and digital transmissions," *Journal of Lightwave Technology*, vol. 16, no. 5, pp. 757–771, May 1998.
4. F. Curti, B. Daino, Q. Mao, F. Matera, and C. G. Someda, "Concatenation of polarization dispersion in single-mode fibres," *Electronics Letters*, vol. 14, no. 4, pp. 290–291, 1989.
5. J. N. Damask, "Methods to construct programmable PMD sources, Part I: Technology and theory," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 997–1005, Apr. 2004.
6. J. N. Damask, P. R. Myers, A. Boschi, and G. J. Simer, "Demonstration of a coherent PMD source," *IEEE Photonics Technology Letters*, vol. 15, no. 11, pp. 1612–1614, Nov. 2003.
7. E. Desurvire, *Erbium-Doped Fiber Amplifiers, Principles and Applications*. Hoboken, New Jersey: Wiley-Interscience, 2002.
8. A. Eyal, D. Kuperman, O. Dimenstein, and M. Tur, "Polarization dependence of the intensity modulation transfer function of an optical system with PMD and PDL," *IEEE Photonics Technology Letters*, vol. 14, no. 11, pp. 1515–1517, Nov. 2002.
9. A. Eyal, W. K. Marshall, M. Tur, and A. Yariv, "Representation of second-order polarization mode dispersion," *Electronics Letters*, vol. 35, no. 19, pp. 1658–1659, 1999.
10. A. Eyal and M. Tur, "A modified poincare sphere technique for the determination of polarization-mode dispersion in the presence of differential gain/loss," in *Tech. Dig., Optical Fiber Communications Conference (OFC'98)*, San Jose, CA, Feb. 1998, paper ThR1, p. 340.
11. R. Feced, S. J. Savory, and A. Hadjifotiou, "Interaction between polarization mode dispersion and polarization-dependent losses in optical communication links," *Journal of the Optical Society of America B*, vol. 20, no. 3, pp. 424–433, Mar. 2003.
12. E. Forestieri and L. Vincetti, "Exact evaluation of the Jones matrix of a fiber in the presence of polarization mode dispersion of any order," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1898–1909, 2001.
13. C. Francia, F. Bruyère, D. Penninckx, and M. Chbat, "PMD second-order effects on pulse propagation in single-mode optical fibers," *IEEE Photonics Technology Letters*, vol. 10, no. 12, pp. 1739–1741, Dec. 1998.
14. C. Francia and D. Penninckx, "Polarization mode dispersion in single-mode optical fibers: Time impulse response," *IEEE International Conference on Communications*, vol. 3, no. 6-10, pp. 1731–1735, June 1999.
15. N. Frigo, private communication, 2003.
16. —, "A generalized geometric representation of coupled mode theory," *IEEE Journal of Quantum Electronics*, vol. QE-22, no. 11, pp. 2131–2140, 1986.

17. N. Gisin, "Statistics of polarization dependent loss," *Optics Communications*, vol. 114, pp. 399–405, Feb. 1995.
18. N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Optics Communications*, vol. 142, pp. 119–125, Oct. 1997.
19. N. Gisin and J. P. Pellaux, "Polarization mode dispersion: Time versus frequency domains," *Optics Communications*, vol. 89, pp. 316–323, May 1992.
20. J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: <http://www.pnas.org>
21. —, "PMD fundamentals: Polarization mode dispersion in optical fibers; Appendix B: Relation between PMD vectors t and w ," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000, supplemental Appendix. [Online]. Available: <http://www.pnas.org>
22. H. A. Haus, *Waves and Fields in Optoelectronics*. Englewood Cliffs, New Jersey: Prentice-Hall, 1984.
23. B. L. Heffner, "Single-mode propagation of mutual temporal coherence: Equivalence of time and frequency measurements of polarization-mode dispersion," *Optics Letters*, vol. 19, no. 15, pp. 1104–1106, Aug. 1994.
24. —, "Optical pulse distortion measurement limitations in linear time invariant systems, and applications to polarization mode dispersion," *Optics Communications*, vol. 115, pp. 45–51, Mar. 1995.
25. —, "Influence of optical source characteristics on the measurement of polarization-mode dispersion of highly mode-coupled fibers," *Optics Letters*, vol. 21, no. 2, pp. 113–115, Jan. 1996.
26. F. Heismann, "Accurate Jones matrix expansion for all orders of polarization mode dispersion," *Optics Letters*, vol. 28, no. 11, p. 20132015, Nov. 2003.
27. F. Heismann and M. S. Whalen, "Fast automatic polarization control system," *IEEE Photonics Technology Letters*, vol. 4, no. 5, pp. 503–505, May 1992.
28. F. Heismann, "Analysis of a reset-free polarization controller for fast automatic polarization stabilization in fiber-optic transmission systems," *Journal of Lightwave Technology*, vol. 12, no. 4, pp. 690–699, Apr. 1994.
29. F. Heismann, D. A. Fishman, and D. L. Wilson, "Automatic compensation of first order polarization mode dispersion in a 10 gb/s transmission system," in *European Conference on Optical Communication (ECOC'98)*, vol. 1, Sept. 1998, pp. 529–530.
30. B. Huttner, C. D. Barros, B. Gisin, and N. Gisin, "Polarization-induced pulse spreading in birefringent optical fibers with zero differential group delay," *Optics Letters*, vol. 24, no. 6, pp. 370–372, Mar. 1999.
31. B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 2, pp. 317–329, Mar. 2000.
32. B. Huttner and N. Gisin, "Anomalous pulse spreading in birefringent optical fibers with polarization-dependent loss," *Optics Letters*, vol. 22, pp. 504–507, Apr. 1997.
33. E. Ibragimov, G. Shtengel, and S. Suh, "Statistical correlation between first and second order PMD," *Journal of Lightwave Technology*, vol. 20, no. 4, pp. 586–590, 2002.

34. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-12: Examinations and measurements - Polarization dependence of attenuation of a single-mode fibre optic component: Matrix calculation method*, International Electrotechnical Commission Std. IEC 61300-3-12, 1997. [Online]. Available: <https://www.iec.ch/>
35. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-2: Examinations and measurements - Polarization dependence of attenuation in a single-mode fibre optic device*, International Electrotechnical Commission Std. IEC 61300-3-2, 1999. [Online]. Available: <https://www.iec.ch/>
36. R. Jones, "A new calculus for the treatment of optical systems, Part II. proof of three general equivalence theorems," *Journal of the Optical Society of America*, vol. 31, no. 7, pp. 493–499, July 1941.
37. I. P. Kaminow, "Polarization in optical fibers," *IEEE Journal of Quantum Electronics*, vol. QE-17, no. 1, pp. 15–22, 1981.
38. M. Karlsson, "Polarization mode dispersion-induced pulse broadening in optical fibers," *Optics Letters*, vol. 23, no. 9, pp. 688–690, May 1998.
39. M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *Journal of Lightwave Technology*, vol. 18, no. 7, pp. 941–951, 2000.
40. H. Kogelnik, L. E. Nelson, and J. P. Gordon, "Emulation and inversion of polarization-mode dispersion," *Journal of Lightwave Technology*, vol. 21, no. 2, pp. 482–495, 2003.
41. H. Kogelnik, L. Nelson, J. P. Gordon, and R. Jopson, "Jones matrix for second-order polarization mode dispersion," *Optics Letters*, vol. 25, no. 1, pp. 19–21, 2000.
42. D. Kuperman, A. Eyal, O. Mor, S. Traister, and M. Tur, "Measurement of the input states of polarization that maximize and minimize the eye opening in the presence of PMD and PDL," *IEEE Photonics Technology Letters*, vol. 15, no. 10, pp. 1425–1427, Oct. 2003.
43. Y. Li and A. Yariv, "Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses," *Journal of the Optical Society of America B*, vol. 17, no. 11, pp. 1821–1827, Nov. 2000.
44. A. Mecozzi and M. Shtaif, "Signal to noise ratio degradation caused by polarization dependent loss and the effect of dynamic gain equalization," *Journal of Lightwave Technology*, 2004, accepted for publication.
45. C. Menyuk, D. Wang, and A. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," *IEEE Photonics Technology Letters*, vol. 9, no. 9, pp. 1247–1249, Sept. 1997.
46. S. M. R. M. Nezam, J. E. McGeehan, and A. E. Willner, "Theoretical and experimental analysis of the dependence of a signals degree of polarization on the optical data spectrum," *Journal of Lightwave Technology*, vol. 22, no. 3, pp. 763–772, Mar. 2004.
47. A. Orlandini and L. Vincetti, "A simple and useful model for Jones matrix to evaluate higher order polarization-mode dispersion effects," *IEEE Photonics Technology Letters*, vol. 13, no. 11, pp. 1176–1178, 2001.
48. —, "Comparison of the Jones matrix analytical models applied to optical system affected by high-order PMD," *Journal of Lightwave Technology*, vol. 21, no. 6, pp. 1456–1464, 2003.

49. D. Penninckx and V. Morenas, "Jones matrix of polarization mode dispersion," *Optics Letters*, vol. 24, no. 13, pp. 875–877, July 1999.
50. D. L. Peterson, B. C. Ward, K. B. Rochford, P. J. Leo, and G. Simer, "Polarization mode dispersion compensator field trial and field fiber characterization," *Optics Express*, vol. 10, no. 14, pp. 614–621, July 2002. [Online]. Available: <http://www.opticsexpress.org/>
51. C. D. Poole and C. R. Giles, "Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber," *Optics Letters*, vol. 13, no. 2, pp. 155–157, 1988.
52. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization mode dispersion in long single-mode fibers," *Electronics Letters*, vol. 22, no. 19, pp. 1029–1030, 1986.
53. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," *Optics Letters*, vol. 16, no. 6, pp. 372–374, 1991.
54. C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, 1994.
55. S. C. Rashleigh, "Origins and control and polarization effects in single-mode fiber," *Journal of Lightwave Technology*, vol. LT-1, no. 2, pp. 312–331, 1983.
56. S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single mode fibers," *Optics Letters*, vol. 3, no. 2, pp. 60–62, 1978.
57. W. Shieh, "Principal states of polarization for an optical pulse," *IEEE Photonics Technology Letters*, vol. 11, no. 6, pp. 677–679, June 1999.
58. W. Shieh and H. Kogelnik, "Dynamic eigenstates of polarization," *IEEE Photonics Technology Letters*, vol. 13, pp. 40–42, 2001.
59. M. Shtaif and A. Mecozzi, "Polarization-dependent loss and its effect on the signal-to-noise ratio in fiber-optic systems," *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 671–673, Feb. 2004.
60. *Measurement of Polarization Dependent Loss (PDL) of Single-Mode Fiber Optic Components*, Telecommunications Industry Association Std. TIA/EIA-455-157, 2000. [Online]. Available: <http://www.tiaonline.org/standards/>
61. S.-C. Wang, private communication, 2002, Insensitivity of Fourier phase to first and last section birefringent phase was first identified by Dr. Wang.
62. C. Xie and L. F. Mollenauer, "Performance degradation induced by polarization-dependent loss in optical fiber transmission systems with and without polarization-mode dispersion," *Journal of Lightwave Technology*, vol. 21, no. 9, pp. 1953–1957, Sept. 2003.
63. C. Xie, L. F. Mollenauer, and L. Moller, "Pulse distortion induced by polarization-mode dispersion and polarization-dependent loss in lightwave transmission systems," *IEEE Photonics Technology Letters*, vol. 15, no. 8, pp. 1073–1075, Aug. 2003.
64. M. Yoshida-Dierolf and V. Dierolf, "Analytical form of frequency dependence of DGD in concatenated single-mode fiber systems," *Journal of Lightwave Technology*, vol. 21, no. 10, pp. 2217–2223, Oct. 2003.

Statistical Properties of Polarization in Fiber

The topic of this chapter is the statistics of polarization, polarization-mode dispersion, and polarization-dependent loss, all in relation to behavior in single-mode fibers. The origin of these statistical properties is the birefringence within the mode-field diameter of the fiber. If a perfectly isotropic fiber existed, the polarization state at the output would match that at the input for all time, frequency, temperature, and length. This, however, is not the case. As illustrated in Fig. 9.1, there are several causes of fiber birefringence, both intrinsic and extrinsic. While an isotropic fiber is stress-free and has a perfectly concentric core (a), perturbations such as core ovality (b), stress-induced index gradient (c), and micro-bubbles (d) introduce birefringence at any cross-section of the fiber. The existence and study of fiber birefringence was well-known by the 1970s [29, 53, 54]. The length-scale of the birefringence is called the birefringent-beat length: $L_B = \lambda_0/\Delta n$, where λ_0 is the free-space wavelength and Δn is the birefringence. Typical birefringence values of early 1990s fiber are $\Delta n/n \sim 10^{-7}$, which corresponds to $L_B \simeq 20$ m at 1550 nm.

Birefringence, however, is only one contributing factor to the statistical behavior. A second factor is the fiber autocorrelation length L_C , which is the characteristic length over which the birefringent axis of the fiber changes. While the models used in the following are cast in the continuous limit, at the discrete level one can think of a fiber of length L segmented into N parts L_C long, where $N = L/L_C$. Within a segment the birefringence and its orientation is fixed, and at the junctions between adjacent segments the birefringent axis changes abruptly. In the continuous limit, L_C represents the characteristic length over which the randomly evolving birefringence loses memory of its preceding orientations. A typical value of L_C is 100 m.

The relationship between the fiber autocorrelation length L_C and the birefringent beat length L_B , and between L_C and the fiber length L , determines the regime in which the polarization ensembles behave. The principal regimes are illustrated in Fig. 9.2. In the first limit $L_C \gg L_B$, Fig. 9.2(a), the slow evolution of the birefringent axis along the fiber allows the optical field of the signal to follow. In the second limit $L_C \ll L_B$, Fig. 9.2(b), the birefringent

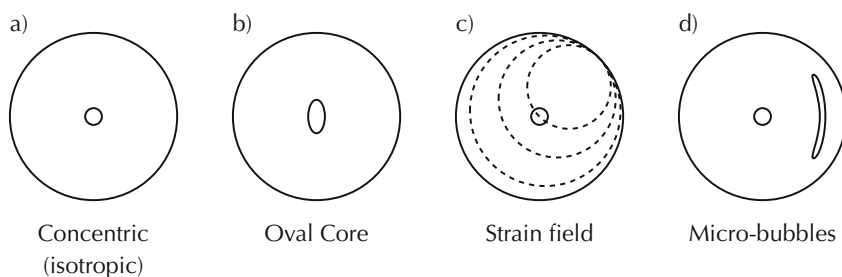


Fig. 9.1. Sources of birefringence within a cross-section of single-mode fiber. a) Perfect, isotropic fiber. b) Ovality of the core. c) Stress-induced index gradient. d) Physical defects like micro-bubbles or impurity concentrations.

axis changes too quickly for the optical field to follow, the result is a long-range average over the range of orientations. This effect is exploited to manufacture ultra-low PMD fiber: the fiber preform is spun during the drawing process at a rate designed to ensure $L_C \ll L_B$ [43]. For instance, Chen *et al.* [6] report a spin period of 1 m and a beat length of 10 m in their fiber. Higher resolution measurements are reported by Pietralunga *et al.* [46] and Galtarossa *et al.* [20]. To be sure, this is not a perfect cure as one must still include some length-scale for variation of the spin profile – this additional factor is illustrated in Fig. 9.2(c) – but the effective birefringence of the fiber is reduced by an order of magnitude.

The relationship between the autocorrelation length L_C and the total fiber length determines how the polarization-mode dispersion behaves. There are two extrema regimes, that of “low” (or “weak”) mode coupling and that of “high” (or “strong”) mode coupling (Fig. 9.2(d)). In the low mode-coupling regime, variation of birefringence orientation is low so the mean PMD increases linearly with length. In the high mode-coupling regime, the birefringence orientation is random beyond a correlation length so the mean PMD increases as the square-root of the length. As a practical matter, since square-root growth is slower than linear, one would like to reach this regime as quickly as possible. As shown in the following, the ratio L/L_C determines the regime; a low autocorrelation length L_C pushes a fiber toward high mode-coupling and root-length growth of the mean PMD.

An historic anecdote conveys the importance of the fiber autocorrelation length. Early fiber-transmission and characterization studies were done in the research lab where fiber is held on spools. When C. D. Poole went to measure for the first time a fiber spooled and then unspooled, he found that the PMD increased five-fold. This is an instance where the correlation length L_C is small on the spool, due to inhomogeneities of bending strain, and large unspooled. Indeed, de Lignie *et al.* report spooled and cabled measurements circa 1994 where they showed $L_C \sim 5$ m on the spool and $L_C \sim 500$ m cabled [9]. Their measurements from fiber to fiber show a wide range of values.

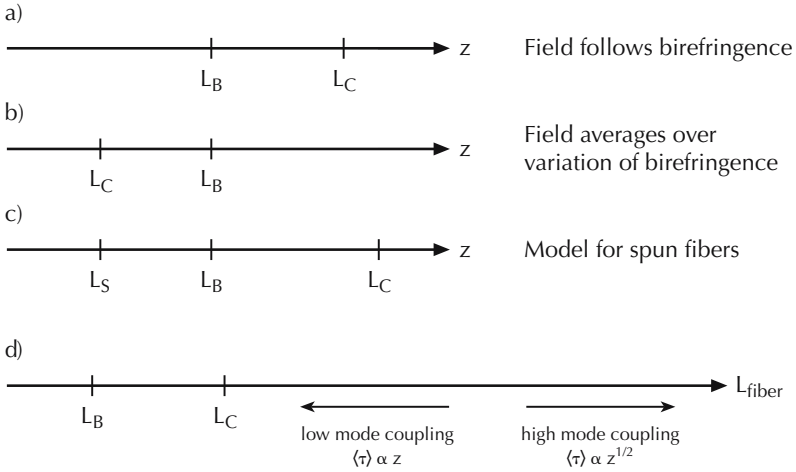


Fig. 9.2. Relationship between length scales within a single-mode fiber. a) Adiabatic regime $L_C \gg L_B$: the field follows the changing birefringence. b) Field-average regime $L_C \ll L_B$: the field cannot follow the changing birefringence vector and instead averages over the variation. c) Model for spun fibers where a third length scale L_S , the range over which the spin profile changes, is added. d) Low- and high-mode coupled PMD regimes: L_C in comparison with the total fiber length L .

For practical systems and design, there are three dimensions along which one would like to derive polarization-related statistics: propagation length, optical frequency, and time. In each case a statistical process must be defined to characterize the evolution on a microscopic level. The PMD evolution equation over length is well defined and the probability density converges in the limit of large ensembles. The ergodic nature of PMD lets “length” be replaced by “optical frequency” in the density functions and “long length” is replaced by “wide bandwidth.” The PMD autocorrelation function connects the length and frequency regimes. There is, however, no definite process for the time evolution. Submarine cable changes at a slow rate while aerial fiber changes in the millisecond range. Moreover, there is likely no spatial homogeneity to the temporal changes – for instance, a train may cross a cable at a particular location – so one cannot expect a neat answer. When the temporal changes are spatial homogeneous, P. J. Leo *et al.* have developed a Rayleigh-distribution model of SOP change that is useful to define what “speed” of change means [32].

Statistics for polarization, PMD, and PDL are derived in the following using diffusion processes. The physics of a diffusion process is first captured by a stochastic differential equation (SDE) and then translated to its partial-differential equation (PDE) analogue. Exposition of these mathematical tools is beyond the scope of this text and the reader is referred to Arnold and Oksendal for SDEs [1, 44], and Risken for PDEs [55]. Finally, Davenport is an invaluable reference on applied probability is [8].

9.1 Polarization Evolution Model

An optical mode confined within a fiber propagates only along one dimension; denote this direction z . The longitudinal electric field will propagate according to the time-harmonic factor $\exp(-jzk_0\sqrt{\varepsilon_r})$, where k_0 is the free-space wavenumber and ε_r is the relative permittivity of the fiber. The Helmholtz equation for the evolution of the electric field $\mathbf{E}$ in the plane perpendicular to z is therefore

$$\left(\frac{d^2}{dz^2} + k_0^2 \vec{\varepsilon}_r\right) \mathbf{E} = 0$$

where $\vec{\varepsilon}_r$ is written in tensor form in anticipation of the following. The common permittivity $\bar{\varepsilon}_r$ may be separated from the differential part such that

$$\vec{\varepsilon}_r = \bar{\varepsilon}_r + \frac{1}{2} \Delta \vec{\varepsilon}_r \cdot \vec{\sigma} \quad (9.1.1)$$

When the propagation is lossless, the common permittivity is the square of the refractive index: $\bar{\varepsilon}_r = n^2$. The common phase is removed from $\mathbf{E}$ to isolate the polarization state evolution using the factorization

$$\mathbf{E} = \exp\left(-j \int_0^z k_0 n(z') dz'\right) |s\rangle$$

where the integral simply accounts for the cumulative change of common index over the path; if the common index is fixed, the integral reduces to the more customary exponential phase factor. Substitution of this factorization and (9.1.1) into the wave equation, and dropping terms that are second-order in $|s\rangle$, makes

$$\left(\frac{d}{dz} + jk_0^2 \frac{\Delta \vec{\varepsilon}_r \cdot \vec{\sigma}}{4n}\right) |s\rangle = 0$$

In the absence of polarization-dependent loss, $\Delta \vec{\varepsilon}_r$ is real and its magnitude to first-order in Δn is

$$\Delta \varepsilon_r = n_+^2 - n_-^2 = (n + \Delta n/2)^2 - (n - \Delta n/2)^2 \simeq 2n\Delta n$$

As a matter of notation, the magnitude of the birefringent vector $\vec{\beta}$ is defined as

$$\beta = |\vec{\beta}| = k_0 \Delta n = \frac{\omega}{c} \Delta n \quad (9.1.2)$$

where the Δ on $\Delta\beta$ has been dropped for convenience. Moreover, the birefringent beat length $L_B = \lambda_o/\Delta n$ is related to the birefringence β as

$$L_B = 2\pi/\beta \quad (9.1.3)$$

With these definitions in hand, the polarization state evolves in the fiber according to [25]

$$\left(\frac{d}{dz} + \frac{j}{2} \vec{\beta} \cdot \vec{\sigma} \right) |s\rangle = 0 \quad (9.1.4)$$

The birefringence vector $\vec{\beta}$ is the local birefringence at any position along the fiber. Equation (9.1.4) describes the response of the polarization state due to the local birefringence. Converting to Stokes space, the differential equation of motion is

$$\frac{d\hat{s}}{dz} = \vec{\beta} \times \hat{s} \quad (9.1.5)$$

As expected, the polarization state precesses about the local birefringent axis at a rate governed by the strength of the birefringence.

Wai and Menyuk propose two models of how the local birefringence varies along an unspun fiber [40, 60, 62]. In both models the fiber exhibits no chirality:

$$\text{No circular birefringence: } \beta_3 = 0$$

This assertion has been experimentally verified for such fibers [21], while spun fibers show evidence of residual chirality [27]. The calculations that follow use the no-chirality assumption, while models for spun fibers can be found in [47]. Without a chiral factor, the birefringent matrix is

$$\vec{\beta} \cdot \vec{\sigma} = \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & -\beta_1 \end{pmatrix} \quad (9.1.6)$$

In their first model, the birefringence magnitude is fixed and the angle θ on the Poincaré equator randomly varies. In their second model, the cartesian birefringent components (β_1, β_2) are independent random variables. Both models give the correct evolution of the mean-square DGD, but the latter model, while a bit more involved, generates aperiodic PMD spectra.

9.1.1 Random Birefringent Orientation

For this first model, the birefringent matrix is

$$\vec{\beta} \cdot \vec{\sigma} = \beta \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \quad (9.1.7)$$

where θ is in Stokes space. The angle is modelled as a Brownian motion on $\mathbb{R}^1$ according to

$$\frac{d\theta}{dz} = g_\theta(z) \longrightarrow \langle g_\theta(z) \rangle = 0, \quad \langle g_\theta(z)g_\theta(z') \rangle = \sigma_\theta^2 \delta(z - z') \quad (9.1.8)$$

where g_θ is a white-noise stochastic process. Physically, the angle θ is subject to impulsive “kicks” as z increases, which in turn drives a random walk in angle. The impulsive nature of the kicks means that there is no memory or correlation from one kick to the next; each is random in its own right.

The Brownian probability density of θ subject to this motion is

$$\rho_\theta(\theta, z) = \frac{1}{\sqrt{2\pi\sigma_\theta^2 z}} \exp\left(-\frac{\theta^2}{2\sigma_\theta^2 z}\right)$$

where, as characteristic with this process, the variance increases linearly with length: $\text{var}(\theta) = \sigma_\theta^2 z$. The “strength” of the white-noise g_θ , σ_θ^2 , is now apparent: the stronger the noise the shorter the fiber length is necessary to reach a nearly uniform angular distribution between $[-\pi/2, \pi/2]$.

As with any random walk, eventually there is complete loss of correlation between some earlier position and the present. In this case, the fiber autocorrelation length L_C is defined as the length over which the angle θ loses correlation. The autocorrelation of θ is calculated by the expectation value of $\cos \theta(z)$:

$$E[\cos \theta(z)] = \int_{\mathbb{R}} \cos \theta \rho_\theta(\theta) d\theta = \exp\left(-\frac{\sigma_\theta^2 z}{2}\right) \quad (9.1.9)$$

The autocorrelation length L_C is the length at which the autocorrelation falls to e^{-1} ; therefore,

$$\sigma_\theta^2 = \frac{2}{L_C} \quad (9.1.10)$$

With this identification, the evolution of the probability density can be written in terms of L_C :

$$\rho_\theta(\theta, z) = \frac{1}{\sqrt{4\pi z/L_C}} \exp\left(-\frac{\theta^2}{4z/L_C}\right) \quad (9.1.11)$$

This distribution represents a diffusion of the angle θ with fiber length (Fig. 9.3). At an initial position the angle is known with certainty; propagating away from this point increases the uncertainty of the angle.

As the fiber length increases, the diffusion of θ approaches a uniform distribution. In fact, only the angle θ modulo π matters. Consider a starting angle of $\theta = 0$; at some distance z/L_C the density at $\theta = \pi/2$ will be within $1 - \varepsilon$ of uniform. Accounting for the wrap-around of the distribution modulo π , the length needed to fall within this error is $z/L_C = \pi^2/(8\varepsilon)$. For instance, it takes about $3L_C$ to realize a 5% deviation from a uniform distribution. Once the distribution is uniform there is no memory of the initial state.

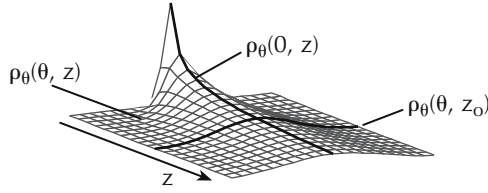


Fig. 9.3. Spatial evolution of the birefringent-angle θ probability density. At any position z_0 the density is a gaussian. On a length scale $z > L_C$ the density converges to a uniform distribution on $[-\pi/2, \pi/2]$.

9.1.2 Random Component Birefringence

For the second model, the entries of the birefringence matrix (9.1.6) are treated as independent Langevin processes:

$$\frac{d\beta_1}{dz} = -L_C^{-1}\beta_1 + g_1(z) \quad (9.1.12a)$$

$$\frac{d\beta_2}{dz} = -L_C^{-1}\beta_2 + g_2(z) \quad (9.1.12b)$$

The characteristics of the noise sources are

$$\langle g_1(z) \rangle = \langle g_2(z) \rangle = 0 \quad \langle g_1(z)g_2(z) \rangle = 0 \quad (9.1.13a)$$

$$\langle g_1(z)g_1(z') \rangle = \langle g_2(z)g_2(z') \rangle = \langle \beta^2 \rangle / L_C \delta(z - z') \quad (9.1.13b)$$

where $\langle \beta^2 \rangle$ is the rms value of the birefringence.

A Langevin equation describes a mean-reverting process driven by noise. Consider (9.1.12a) in the absence of $g_1(z)$: any initial condition decays exponentially over characteristic length L_C . Turning the noise source on disrupts this deterministic motion. The minus sign of the deterministic coefficient $-L_C^{-1}$ acts as a linear spring constant, pulling the solution back toward the mean with a strength proportional to the deviation. In the steady state, this spring force balances the noise source, resulting in a stationary gaussian distribution.

The solution to (9.1.12) is

$$\beta_i(z) = \beta_i(0) e^{-z/L_C} + \int_0^z e^{-(z-z')/L_C} g_i(z') dz'$$

The initial condition $\beta_i(0)$ decays exponentially on a scale given by the fiber autocorrelation length L_C . In the regime $z \gg L_C$ there is no memory of the initial state and a stationary distribution is reached. A two-dimensional sample-path of the birefringence is illustrated in Fig. 9.4. The steady-state density of β_i is readily determined by solution of the associated Fokker-Planck equation, and is

Realization of a birefringence-vector path

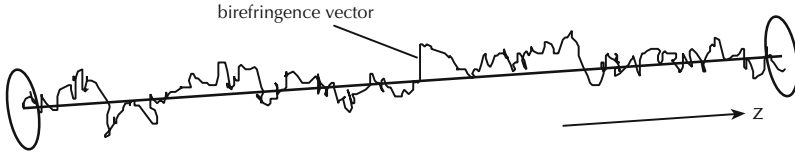


Fig. 9.4. Sample path of the birefringence vector in the steady-state. This path was calculated using a Karhunen-Loeve expansion of a Wiener process and a numerical integration of the Langevin equation (9.1.12). In the steady-state $\beta_{1,2}$ converge in distribution to i.i.d. stationary gaussian processes.

$$\rho_{\beta_i}(v) = \frac{1}{\sqrt{\pi \langle \beta^2 \rangle}} \exp \left(-\frac{v^2}{\langle \beta^2 \rangle} \right) \quad (9.1.14)$$

The variance of each component is $\text{var}(\beta_i) = \langle \beta^2 \rangle / 2$, which is independent of z . Moreover, as detailed in Appendix D, the radial and angular distributions of the local birefringence vector $\vec{\beta} = \hat{x}\beta_1 + \hat{y}\beta_2$ are Rayleigh and uniform distributions, respectively. Finally, the second moment of the Rayleigh distribution is $\langle \beta^2 \rangle$, which is what is expected on physical grounds. Therefore one writes

$$\text{var}(\beta_i) = \frac{1}{2} \text{var}(|\vec{\beta}|) \quad (9.1.15)$$

This and the preceding section have detailed physically reasonable forms of the local fiber birefringence vector $\vec{\beta}$ that drives the evolution of the polarization state and, consequently, the PMD evolution. A significant further study by Marcuse *et al.* details how these models are used to analyze pulse propagation, and particularly non-linear propagation, in fibers [36].

9.2 Polarization Diffusion in Single-Mode Fiber

The optical field evolves along a single-mode fiber subject to local birefringence perturbations. The characteristic length L_E is the length-scale over which the electric field loses memory and is called the polarization decorrelation length [60]. This is an additional length scale to those characteristic of the fiber, namely, the birefringent beat length L_B and the fiber autocorrelation length L_C .

One naturally expects the polarization decorrelation length to be related to the birefringent properties of the fiber. At one extreme, where $L_C \ll L_B$, the optical field averages over the rapidly varying birefringence, thereby changing slowly with respect to its initial state. At the other extreme where $L_B \ll L_C$, the field tends to follow the local birefringence and will, accordingly, change slowly with respect to the local state. Based on this physical picture, there are two choices for the polarization decorrelation length: a fixed definition $L_{E,\text{fixed}}$

that measures the local field with respect to the initial birefringence, and a local definition $L_{E,\text{local}}$ that measures the local birefringence.

To compute the polarization decorrelation length, the Stokes picture of polarization diffusion (9.1.5) is used. Recalling that the fiber model assumes no chirality, the component form of the precession equation reads

$$\frac{d}{dz} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} & \beta_2 & \\ -\beta_2 & & \beta_1 \\ & -\beta_1 & \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} \beta_2 S_3 \\ -\beta_1 S_3 \\ -\beta_2 S_1 + \beta_1 S_2 \end{pmatrix}$$

where capital S_k denotes a random variable. In order to simplify this expression prior to writing the diffusion generator, Wai and Menyuk introduce a rotation operator $R(z)$ to rotate the local birefringence to point along s_1 [62]. This operator is simple because $\beta_3 = 0$:

$$R(z) = \begin{pmatrix} \cos \theta(z) & -\sin \theta(z) & 0 \\ \sin \theta(z) & \cos \theta(z) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (9.2.1)$$

By defining the local Stokes coordinates such that $\tilde{s} = R(z)\hat{s}$, the precession equation (9.1.5) is transformed to

$$\frac{d}{dz} \tilde{s} = \left(R\vec{\beta} \times R^{-1} \right) \tilde{s} - R R_z^{-1} \tilde{s} \quad (9.2.2)$$

where R_z is the derivative of $R(z)$ with respect to z . This precession equation is called the local evolution equation, to distinguish it from (9.1.5) which describes fixed-reference evolution.

There are two derivations that can follow from (9.2.2) – the first using the fixed-birefringence fiber model, and the second using Langevin fiber model. The results of the two calculations are not qualitatively different; so the first, and simpler, model is detailed below.

For the first fiber model, the birefringent variation (9.1.7) and its noise source (9.1.8) is substituted into the local evolution equation. This gives

$$\frac{d}{dz} \begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \tilde{S}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\beta \tilde{S}_3 \\ \beta \tilde{S}_2 \end{pmatrix} + \begin{pmatrix} \tilde{S}_2 \\ -\tilde{S}_1 \\ 0 \end{pmatrix} g_\theta \quad (9.2.3)$$

This is a stochastic differential equation (SDE) in the Itô form. While it is quite beyond the scope of this text, the discerning reader should be aware that this equation is not quite correct. As Foschini shows [17], the Itô interpretation leads to an immediate departure of the polarization state from the unit sphere once (9.2.3) is integrated, while the associated Stratonovich interpretation does not. One has two choices on how to treat the analytics. Either the infinitesimal probability generator G accounts for the Stratonovich interpretation by adding a correction term or the SDE is translated from Stratonovich

to Itô form and the Itô generator is used. Examples of both treatments are given below.

The infinitesimal probability generator governs the diffusion of the probability density associated with the stochastic differential equation

$$dX_{i,z} = b(z, X_{i,z})dz + \sigma(z, X_{i,z})dB_{i,z} \quad (9.2.4)$$

where b is the column-vector coefficient of the drift and σ is the column vector coefficient of the Brownian motion – either are functions of length and the random variable. Brownian motion is related to white noise by $dB_z = g_z dz$. The expectation of a sufficiently smooth functional ψ on coordinates X_i evolves according to

$$\frac{d\langle\psi\rangle}{dz} = \langle G\psi\rangle \quad (9.2.5)$$

This is Kolmogorov's backward equation (KBE). The generator G is the probability generator of the Itô diffusion (see Foschini [17], Menyuk [62], Ok-sendal [44] (esp. Theorem 7.3.3, and (6.1.3)), and Risken [55] for more details). The importance of the probability generator is that it transforms a stochastic differential equation into a partial differential equation (PDE). Many powerful analytic and numeric tools are available to solve PDEs, making problems cast in this form more tractable. The polarization and PMD diffusions that follow are prime examples of physical processes developed first in SDE form to capture the differential behavior of the process and then solved in PDE form to determine the global behavior subject to the boundary conditions.

The Itô-sense diffusion generator has two components: $G_I = G_d + G_s$. These components account for, respectively, the deterministic drift of the system and the stochastic fluctuation. The Itô generator for (9.2.4) is

$$G_I = \sum_i b_i(x_i) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_i \sum_j (\sigma \sigma^T)_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} \quad (9.2.6)$$

The Stratonovich-sense G_S generator makes a correction for the drift and, while more complicated to express in general, for the present case it may be written as

$$G_S = G_I - \frac{1}{2} \sigma^2 \sum_i x_i \frac{\partial}{\partial x_i} \quad (9.2.7)$$

The calculations in this section use the Stratonovich-sense generator.

While the derivations of the above equations are advanced, its application is straightforward. The generator for (9.2.3) is

$$\begin{aligned} G = & -\beta \tilde{S}_3 \frac{\partial}{\partial \tilde{S}_2} + \beta \tilde{S}_2 \frac{\partial}{\partial \tilde{S}_3} \\ & + \frac{\sigma_\theta^2}{2} \left(\tilde{S}_2^2 \frac{\partial^2}{\partial \tilde{S}_1^2} + \tilde{S}_1^2 \frac{\partial^2}{\partial \tilde{S}_2^2} - 2\tilde{S}_1 \tilde{S}_2 \frac{\partial^2}{\partial \tilde{S}_1 \partial \tilde{S}_2} - \tilde{S}_1 \frac{\partial}{\partial \tilde{S}_1} - \tilde{S}_2 \frac{\partial}{\partial \tilde{S}_2} \right) \end{aligned} \quad (9.2.8)$$

where the noise strength σ_θ^2 is related to the fiber autocorrelation length via (9.1.10).

The evolution of the moments of $\tilde{S}$ are calculated using this generator and the KBE. For instance, the functionals $\psi(\tilde{S}) = \tilde{S}$ and $\psi(\tilde{S}) = \tilde{S}^2$ give the evolution of the first- and second-moments of $\tilde{S}$. These results are used to associate the polarization decorrelation length L_E with the fiber parameters.

For the evolution of the mean values, the functional ψ is $\psi(\tilde{S}_i) = \tilde{S}_i$. Calculation of $G\psi$ generates the following system of equations:

$$\frac{d}{dz} \langle \tilde{S}_1 \rangle = -\frac{1}{L_C} \langle \tilde{S}_1 \rangle \quad (9.2.9a)$$

$$\frac{d}{dz} \langle \tilde{S}_2 \rangle = -\frac{2\pi}{L_B} \langle \tilde{S}_3 \rangle - \frac{1}{L_C} \langle \tilde{S}_2 \rangle \quad (9.2.9b)$$

$$\frac{d}{dz} \langle \tilde{S}_3 \rangle = \frac{2\pi}{L_B} \langle \tilde{S}_2 \rangle \quad (9.2.9c)$$

For a non-zero initial condition, $\langle \tilde{S}_1(z) \rangle$ monotonically decays to zero and the remaining mean values undergo a damped oscillation to zero. The long-range values of the polarimetric means are all zero; the polarization state with respect to the local birefringence ultimately becomes completely uncorrelated. In the particular case when the initial state of the system is $\tilde{S}_1 = 1$, that is, the launch polarization is aligned to the local birefringent axis, the mean polarimetric values of the remaining two coordinates are $\langle \tilde{S}_2(z) \rangle = \langle \tilde{S}_3(z) \rangle = 0$, and the mean along the initial axis decays as

$$\langle \tilde{S}_1(z) \rangle = \exp(-z/L_C) \quad (9.2.10)$$

This simple case is all that is necessary to associate the polarization decorrelation length $L_{E,\text{local}}$ with the fiber autocorrelation length. Since the characteristic length over which the mean polarization is preserved is by definition the polarization decorrelation length, in light of (9.2.10) one makes the association

$$L_{E,\text{local}} = L_C \quad (9.2.11)$$

In the local reference frame the two characteristic lengths are equal.

Wai and Menyuk detail the transformation to the fixed reference frame from the local frame [62]. While their work may be consulted for the details, the results are

$$\langle S_1(z) \rangle = \begin{cases} S_1(0) \exp\left(-z/L'_{E,\text{fixed}}\right) & L_C \ll L_B \\ S_1(0) \exp\left(-z/L''_{E,\text{fixed}}\right) & L_C \gg L_B \end{cases} \quad (9.2.12)$$

where

$$L'_{E,\text{fixed}} = \frac{L_C}{2\pi^2 (L_C/L_B)^2} \quad (9.2.13a)$$

$$L''_{E,\text{fixed}} = L_C/2 \quad (9.2.13b)$$

These equations reenforce the physical understanding developed in the introduction of this chapter. With respect to the launched polarization state, when $L_C \ll L_B$ the polarization decorrelation length (9.2.13a) is much longer than the fiber autocorrelation length. This is because the birefringence changes too rapidly for the field to follow, which in turn makes the propagated field correlate with the launched field over a longer distance. Conversely, when $L_C \gg L_B$, the field follows the birefringence more faithfully, so the polarization state diffuses on a length scale more closely linked to the fiber autocorrelation length. In particular, notice that $L''_{E,\text{fixed}} = L_{E,\text{local}}/2$, which makes sense because the field follows the local birefringence, so the local-frame characteristic length should indeed be longer than the fixed reference frame.

These associations between the fiber autocorrelation and polarization decorrelation lengths are useful in a practical sense. While L_C is central to the statistical description of the optical field, the polarization decorrelation length is the measurable quantity. Equation (9.2.12) provides a means in which to determine L_C through the measurement of $L_{E,\text{fixed}}$.

For the evolution of the polarimetric second-moments, the functional ψ is set to $\psi(\tilde{S}_i) = \tilde{S}_i^2$ and the generator (9.2.8) remains the same. Calculation of $G\psi$ generates the following system of equations:

$$\begin{aligned} \frac{d}{dz} \langle \tilde{S}_1^2 \rangle &= -\frac{2}{L_C} \left(\langle \tilde{S}_1^2 \rangle - \langle \tilde{S}_2^2 \rangle \right) \\ \frac{d}{dz} \langle \tilde{S}_2^2 \rangle &= \frac{2}{L_C} \left(\langle \tilde{S}_1^2 \rangle - \langle \tilde{S}_2^2 \rangle \right) - \frac{4\pi}{L_B} \langle \tilde{S}_2 \tilde{S}_3 \rangle \\ \frac{d}{dz} \langle \tilde{S}_3^2 \rangle &= \frac{4\pi}{L_B} \langle \tilde{S}_2 \tilde{S}_3 \rangle \\ \frac{d}{dz} \langle \tilde{S}_2 \tilde{S}_3 \rangle &= \frac{2\pi}{L_B} \left(\langle \tilde{S}_2^2 \rangle - \langle \tilde{S}_3^2 \rangle \right) - \frac{2}{L_C} \langle \tilde{S}_2 \tilde{S}_3 \rangle \end{aligned} \quad (9.2.14)$$

As before there are local- and fixed-reference frame solutions, and the solutions differ for the two limits of L_C/L_B . The general form for the fixed-frame solution is

$$\langle S_{1,2}^2 \rangle \simeq \frac{1}{3} \left(1 \pm \frac{3}{2} \exp(-z/L_{E,1}) + \frac{1}{2} \exp(-z/L_{E,2}) \right) \quad (9.2.15a)$$

$$\langle S_3^2 \rangle \simeq \frac{1}{3} \left(1 - \exp(-z/L_{E,3}) \right) \quad (9.2.15b)$$

where in the first equation the “+” and “−” signs refer to $\tilde{S}_1$ and $\tilde{S}_2$, respectively. In the $L_C \ll L_B$ regime, the length scales are

$$L_{E,1} = 2L'_{E,\text{fixed}}, \quad \text{and} \quad L_{E,(2,3)} = 6L'_{E,\text{fixed}} \quad (9.2.16)$$

Conversely, in the $L_C \gg L_B$ regime, the length scales are

$$L_{E,1} = \frac{2}{7}L''_{E,\text{fixed}}, \quad \text{and} \quad L_{E,(2,3)} = \frac{2}{3}L''_{E,\text{fixed}} \quad (9.2.17)$$

In either length regime, the stationary variances of the diffusions converge to $\langle \tilde{S}_k^2 \rangle = 1/3$. The convergence rate for all three variances is roughly the same, but the small difference was studied by Wai and Menyuk in [61], where they showed that the absence of chirality in the fiber imparts a short-range anisotropy to the diffusions.

To summarize, polarization decorrelation happens with its own characteristic length-scale L_E in a single-mode fiber. That length scale is related to the fiber birefringence parameters in either a local or fixed frame of reference. In the local frame, the polarization decorrelation and fiber autocorrelation lengths are equal. In the fixed frame, the relationship depends on the regime in which the fiber is characterized: for $L_C \ll L_B$ the diffusion occurs at a rate related to L_B ; for $L_C \gg L_B$ the diffusion occurs at a rate related to L_C . The three polarimetric values all reach a mean of zero and a variance of $1/3$ beyond the diffusion limit. The diffusions detailed in this section are for the fixed-birefringence fiber model, but the results do not qualitatively change for the Rayleigh-distributed birefringence model.

9.3 RMS Differential-Group Delay Evolution

The equation of motion for polarization evolution (9.1.5) was recast in the preceding section into a stochastic differential equation whose solutions showed the behavior of the statistical moments of the polarization state. In parallel with this procedure, the equation of motion for polarization-mode dispersion evolution is studied. Recall from (8.2.39) on page 339 that the differential equation of motion for the PMD vector $\vec{\tau}$ is

$$\frac{\partial \vec{\tau}}{\partial z} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau} \quad (9.3.1)$$

where $\vec{\beta}$ is the local birefringence vector and $\vec{\beta}_\omega$ is its frequency derivative.

The solution to (9.3.1) for the mean-square magnitude of $\vec{\tau}$ in the diffusion limit is

$$\langle \tau^2(z) \rangle = 2 \langle \tau_c^2 \rangle \left(e^{-z/L_C} + z/L_C - 1 \right) \quad (9.3.2)$$

where $\langle \tau_c^2 \rangle$ is the mean-square DGD for a segment L_C long. The mean-square solution is written at this point in the discussion because it is apparently independent of any reasonable derivation. Poole [48, 49] first derived this equation, shortly followed by Curti [7], Foschini [17], and Gisin [23, 24], and later by Wai and Menyuk [60]. Gisin [24] showed that (9.3.2) is the mean-square deviation of the probability density that solves the Telegrapher's equation (a

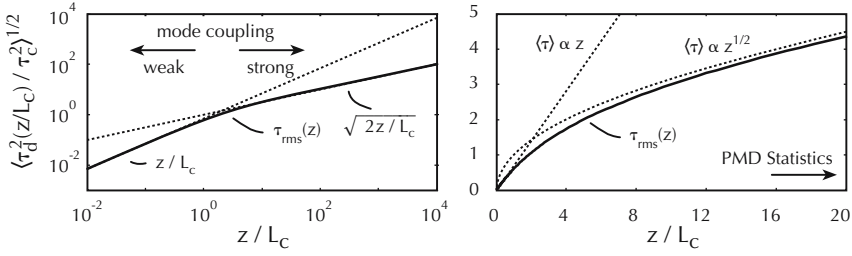


Fig. 9.5. The rms growth of $\langle \tau \rangle$ with length and its asymptotic limits. a) Log-log scale shows long-range behavior. For $z \ll L_C$ the rms growth is linear with length, while for $z \gg L_C$ the rms growth goes as root-length. The crossover is in the range $z \sim L_C$. b) Linear scale of the same. PMD fiber statistics are derived in the strong mode-coupling regime.

second-order parabolic PDE). While the details depend on which report is consulted, all of these researchers apply a gaussian-correlated noise to the fiber birefringence.

The root-mean-square growth of the DGD, defined by $\langle \tau \rangle_{\text{rms}} = \sqrt{\langle \tau^2(z) \rangle}$, shows two asymptotic limits:

$$\langle \tau \rangle_{\text{rms}}(z) = \begin{cases} (z/L_C) \langle \tau_c \rangle_{\text{rms}} & \text{weak coupling: } z \ll L_C \\ \sqrt{2z/L_C} \langle \tau_c \rangle_{\text{rms}} & \text{strong coupling: } z \gg L_C \end{cases} \quad (9.3.3)$$

where $\langle \tau_c \rangle_{\text{rms}} = \sqrt{\langle \tau_c^2 \rangle}$. The cross-over between these two limits is $z \sim L_C$. The two limits of the rms evolution function are shown in Fig. 9.5.

At one limit, the so-called “weak” mode-coupling limit, $\langle \tau \rangle_{\text{rms}}$ grows linearly with length z . In this regime the local birefringence is nearly aligned along the fiber, making the cumulative birefringence additive. At the other limit, the “strong” mode-coupling limit, $\langle \tau \rangle_{\text{rms}}$ grows as root-length: $z^{1/2}$. In this regime the fiber length is well beyond the fiber autocorrelation length, so the cumulative birefringence grows statistically. Given that the birefringence of fiber segments L_C long are uncorrelated, the birefringence variances across segments add, resulting in a standard deviation that grows as root-length. This behavior was already seen in the probability density (9.1.11) for the angular distribution of birefringence.

The analytic calculation of the PMD probability densities are made in the strong-coupling regime. The weak-coupling regime poses few novel problems because the statistics will closely match that of the local birefringence. The intermediate region that connects the weak- and strong-coupling regimes can be treated one of two ways. For the case of fiber the PMD statistics evolve through the intermediate region toward their stationary end-points. Tan *et al.* have studied these transient statics and show that the Maxwellian distribution is achieved by approximately $z \sim 30L_C$, but that the distribution tails take longer to mature [58, 63]. Another case is for PMD emulators, where a fixed

number (3–20) of sections is used. These statistics have also been investigated and shown to exhibit deviation from the stationary forms [30, 34].

Equation (9.3.2) is derived here following the diffusion formalism. Use of the fixed-birefringence model of §9.1.1 makes for a simpler calculation; the result is easily extended to Rayleigh-distributed birefringence. As with the polarization calculations, the PMD diffusion equation is simpler when converted to a local reference frame. Define $\tilde{\tau} = R(z)\vec{\tau}$, where $R(z)$ is as in (9.2.1). The resulting stochastic differential equation for (9.3.1) in component form is (cf. (9.2.3))

$$\frac{d}{dz} \begin{pmatrix} \tilde{\tau}_1 \\ \tilde{\tau}_2 \\ \tilde{\tau}_3 \end{pmatrix} = \begin{pmatrix} \beta_\omega \\ -\beta\tilde{\tau}_3 \\ \beta\tilde{\tau}_2 \end{pmatrix} + \begin{pmatrix} \tilde{\tau}_2 \\ -\tilde{\tau}_1 \\ 0 \end{pmatrix} g_\theta \quad (9.3.4)$$

The infinitesimal probability generator is

$$G = \beta_\omega \frac{\partial}{\partial \tilde{\tau}_1} - \beta\tilde{\tau}_3 \frac{\partial}{\partial \tilde{\tau}_2} + \beta\tilde{\tau}_2 \frac{\partial}{\partial \tilde{\tau}_3} + \frac{\sigma_\theta^2}{2} \left(\tilde{\tau}_2^2 \frac{\partial^2}{\partial \tilde{\tau}_1^2} + \tilde{\tau}_1^2 \frac{\partial^2}{\partial \tilde{\tau}_2^2} - 2\tilde{\tau}_1\tilde{\tau}_2 \frac{\partial^2}{\partial \tilde{\tau}_1 \partial \tilde{\tau}_2} - \tilde{\tau}_1 \frac{\partial}{\partial \tilde{\tau}_1} - \tilde{\tau}_2 \frac{\partial}{\partial \tilde{\tau}_2} \right) \quad (9.3.5)$$

Finally, the functional of interest is $\psi = \tilde{\tau}^2 = \tilde{\tau}_1^2 + \tilde{\tau}_2^2 + \tilde{\tau}_3^2$. A requisite auxiliary functional $\psi' = \tilde{\tau}_1$ is also needed. Calculating $G\psi$ and $G\psi'$, and combining the results, yields the equation of motion

$$\left(\frac{d^2}{dz^2} + \frac{1}{L_C} \frac{d}{dz} \right) \langle \tilde{\tau}^2 \rangle = 2\beta_\omega^2 \quad (9.3.6)$$

The solution to this equation is (9.3.1). A few details need clarification. The length of the PMD vector is invariant under rotation, so $|\tilde{\tau}| = |R(z)\vec{\tau}|$. The product $\beta_\omega L_C$ is the characteristic DGD per segment L_C long. To see this, simply expand the terms: $\tau_c = \beta_\omega L_C = \Delta n_g L_C / c$. Finally, the replacement of τ_c^2 in the solution of (9.3.6) with $\langle \tau_c^2 \rangle$ as reported in (9.3.2) comes with the Rayleigh-birefringence derivation detailed by Wai and Menyuk [62].

9.4 PMD Statistics

The PMD statistics that yield to analytic study are those related to the first- and second-order PMD vectors ($\vec{\tau}, \vec{\tau}_\omega$) and the autocorrelation of the PMD vector $\vec{\tau}$ with frequency. The statistics that have been analytically solved are for $\vec{\tau}$ and its components τ_i ; $\vec{\tau}_\omega$ and its components $\tau_{\omega,i}$; and the perpendicular and parallel components of the second-order vector: depolarization $|\vec{\tau}_{\omega,\perp}|$ and polarization-dependent chromatic dispersion $|\vec{\tau}|_\omega$, respectively. Additionally, the relation between depolarization and PDCD conditional on the DGD has been determined. The joint density of the magnitudes (τ, τ_ω) , however, is not

analytic but has been solved using importance sampling (IS) and, separately, special numerical techniques.

A remarkable property of PMD statistics is that they scale with a single scaling factor: the mean fiber DGD $\bar{\tau}$. The mean fiber DGD is itself related to the fiber length, fiber autocorrelation length, and birefringence variance. That association is clarified below, but once made the mean fiber DGD becomes its own “unit.” The mean fiber DGD is so ubiquitous in the statistics and measurement of PMD that a corruption of terms has entered the literature where “PMD” is *defined* as the mean fiber DGD. There should be no confusion, however, between PMD as a vector and mean DGD as a statistical unit.

The expression for the rms DGD evolution in the strong coupling limit connects the microscopic scale of the birefringence variation with the macroscopic properties of PMD statistics. This expression is therefore the gateway between micro- and macroscopic views of the same process. While this expression was derived above, the derivation of the PMD probability densities requires analytic tools that are well beyond the scope of this text. Instead, the results, principally of Foschini, will be quoted and the ambitious reader is referred to the cited papers.

The mean fiber DGD is connected to the stochastic model of PMD evolution in the following way. In the $z \gg L_C$ limit, the mean-square DGD grows as

$$\langle \tau^2(z) \rangle = 2 \langle \tau_c^2 \rangle z / L_C \quad (9.4.1)$$

As discussed in the following, the cartesian components of the PMD vector are i.i.d. gaussian random variables. The probability density of the length of the PMD vector is therefore Maxwellian. The first and second moments of this density are related (see Table D.1 on page 507), so (9.4.1) may be rewritten for the mean DGD:

$$\bar{\tau} \equiv \langle \tau(z) \rangle = \sqrt{\frac{8}{3\pi}} \frac{2 \langle \tau_c^2 \rangle z}{L_C} \quad (9.4.2)$$

There is some variation in the literature on the interpretation of $2 \langle \tau_c^2 \rangle / L_C$. Curti, for instance, writes $\bar{\tau} = \sqrt{8z/\pi L_C} d\tau$ [7]. Since (9.3.2) was derived using a Rayleigh statistic for the birefringence, the mean segment DGD is related to its second moment by $\langle \tau_c^2 \rangle = 4 \langle \tau_c \rangle^2 / \pi$. Association with Curti gives $d\tau = \sqrt{8/3\pi} \langle \tau_c \rangle$, which is an 8.5% difference. Separately, Poole and Favin [52] write $\bar{\tau} = \sqrt{8N/3\pi} \Delta\tau_p$, where $\Delta\tau_p$ is the fixed birefringence of a retardation plate and N is the number of plates. The connection to (9.4.2) requires $N = 2z/L_C$. This interpretation is used below to develop a discrete waveplate model. Other researchers reproduce (9.4.2) in the continuous limit [17, 23, 51, 62].

Armed with the definition of the mean fiber DGD $\bar{\tau}$, the statistics of PMD are presented below. The principal contributors to this field are Curti [7], who first derived the Maxwellian DGD distribution; Foschini [14–17], who derived the remaining PMD distributions; Karlsson [31], and Shtaif and

Mecozzi [56, 57], who derived the PMD autocorrelation functions; Ibragimov and Shtengel [28], who derived a conditional expression; and Fogal, Biondini, and Kath, who developed the IS methods [2, 11, 12].

9.4.1 Probability Densities

The expressions for the probability densities of the first- and second-order PMD vector are tabulated in Table 9.1. The vectors $\vec{\tau}$ and $\vec{\tau}_\omega$ are statistically dependent on one another. Plots of these densities are shown in Fig. 9.6 on linear scale, to emphasize the distribution about the mean, and semi-log scale, to emphasize the fall-off of the distribution tails.

The probability density for the DGD τ , which is the magnitude of the PMD vector $\tau = |\vec{\tau}|$, is Maxwellian. The origin of the Maxwellian distribution comes from the distribution of the radius of a sphere, where the cartesian coordinates of the sphere are i.i.d. gaussian random variables (see Appendix D). An important relation between the mean and mean-square for a Maxwellian distribution is

$$\langle \bar{\tau}^2 \rangle = \left(\frac{8}{3\pi} \right) \langle \bar{\tau} \rangle^2 \quad (9.4.3)$$

The $8/3\pi$ factor appears frequently in the discussion of PMD statistics. The Maxwellian and gaussian distributions scale linearly in $\bar{\tau}$. That is, the measured DGD values from any fiber can be normalized by the mean fiber DGD ($\tau/\bar{\tau}$) to produce a unit-scaled statistic.

The probability density for the magnitude SOPMD $\tau_\omega = |\vec{\tau}_\omega|$ is sech-tanh in form. The origin of the sech-tanh distribution comes from the distribution of the radius of a sphere, where the cartesian coordinates of the sphere are i.i.d. hyperbolic secant (sech) random variables. In contrast to the Maxwellian and gaussian distributions, the sech-tanh and sech distributions scale quadratically with $\bar{\tau}^2$: measurements can be normalized to a unit statistic by $\tau_\omega/\bar{\tau}^2$.

The second moments of the τ and τ_ω magnitudes are related by

$$\langle \tau_\omega^2 \rangle = \frac{1}{3} \langle \bar{\tau}^2 \rangle^2 \quad (9.4.4)$$

Moreover, a comparison of the tails of the Maxwellian and sech-tanh distributions (Fig. 9.6), shows that eventually the Maxwellian falls off faster. This is because the underlying gaussian distribution falls off quadratically (on a log scale) while that of the sech fall off linearly.

The cartesian components of the SOPMD vector are i.i.d. random variables, so their variance is one-third that of the vector length. Yet, the cartesian components are not directly or easily related to the distortion PMD imparts on a signal. However, the projections of the SOPMD vector onto the first-order vector are directly related, so one asks about the conditional dependence of these projections.

Table 9.1. Statistical Relations of PMD: $\vec{\tau} = \tau \hat{p}$, $\vec{\tau}_\omega = \tau_\omega \hat{p} + \tau \hat{p}_\omega$

Statistic	Symbol	$\langle \tau \rangle$	$\langle \tau^2 \rangle$	Density	Domain
DGD ^(a--d)	$ \vec{\tau} $	$\bar{\tau}$	$\langle \bar{\tau}^2 \rangle^{(*)}$	$\frac{32}{\pi^2} \frac{\tau^2}{\bar{\tau}^3} \exp\left(-\frac{1}{2} \frac{8\tau^2}{\pi \bar{\tau}^2}\right)$	$\tau \in [0, \infty)$
SOPMD ^(d)	$ \vec{\tau}_\omega $	$\left(\frac{2G}{\pi}\right) \bar{\tau}^2$	$\frac{1}{3} \langle \bar{\tau}^2 \rangle^2$	$\frac{8}{\pi \bar{\tau}^2} \frac{4\tau}{\bar{\tau}^2} \tanh\left(\frac{4\tau}{\bar{\tau}^2}\right) \operatorname{sech}\left(\frac{4\tau}{\bar{\tau}^2}\right)$	$\tau \in [0, \infty)$
$\vec{\tau}$ component ^(a--c)	τ_i	0	$\frac{1}{3} \langle \bar{\tau}^2 \rangle^2$	$\frac{2}{\pi \bar{\tau}} \exp\left(-\frac{1}{2} \frac{8\tau^2}{\pi \bar{\tau}^2}\right)$	$\tau \in (-\infty, \infty)$
$\vec{\tau}_\omega$ component ^(d)	$\tau_{\omega,i}$	0	$\frac{1}{9} \langle \bar{\tau}^2 \rangle^2$	$\frac{4}{\pi \bar{\tau}^2} \operatorname{sech}\left(\frac{4\tau}{\bar{\tau}^2}\right)$	$\tau \in (-\infty, \infty)$
PDGD ^(e,f)	$ \vec{\tau}_{\omega,\parallel} = \vec{\tau} _\omega$	0	$\frac{1}{27} \langle \bar{\tau}^2 \rangle^2$	$\frac{2}{\bar{\tau}^2} \operatorname{sech}^2\left(\frac{4\tau}{\bar{\tau}^2}\right)$	$\tau \in (-\infty, \infty)$
Depolarization ^(g)	$ \vec{\tau}_{\omega,\perp} $		$\frac{8}{27} \langle \bar{\tau}^2 \rangle^2$	$u^2 \tau \int_0^\infty J_0(u\tau\alpha) \operatorname{sech}\alpha (\alpha \tanh\alpha)^{1/2} d\alpha, \quad u = \frac{8}{\pi \bar{\tau}^2}$	$\tau \in [0, \infty)$
Depolarization vector ^(g)	$ \hat{p}_\omega $		$\frac{4}{9} \langle \bar{\tau}^2 \rangle^2$	$3u\tau \int_0^\infty \frac{\sinh^{3/2}\beta}{\sqrt{\beta} \cosh^{5/2}\beta} {}_1F_1\left(\frac{5}{2}, 1; -\frac{1}{2} u\tau^2 \beta \tanh\beta\right) d\beta$	$\tau \in [0, \infty)$
^(*) $\langle \bar{\tau}^2 \rangle = \left(\frac{3\pi}{8}\right) \bar{\tau}^2, \quad {}_1F_1\left(\frac{5}{2}, 1; v\right) = \frac{e^{v/2}}{3} \left((3+2v(3+v))I_0\left(\frac{v}{2}\right) + 2v(2+v)I_1\left(\frac{v}{2}\right)\right), \quad G = 0.915965\dots, \text{ Catalan's constant.}$					
^(a) Curti <i>et al.</i> [7], ^(b) Poole <i>et al.</i> [51], ^(c) Foschini [17], ^(d) Gisin [23], ^(e) Foschini [15], ^(f) Foschini [14], ^(g) Foschini [16].					

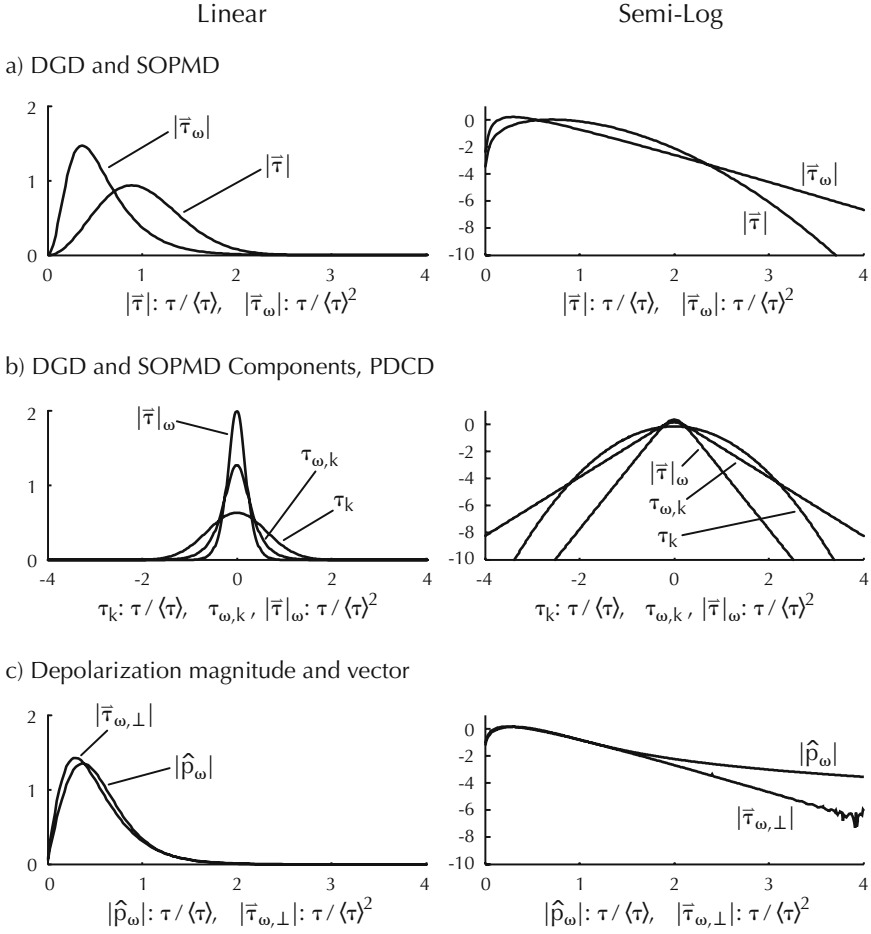


Fig. 9.6. Probability densities of first- and second-order PMD statistics, linear and semi-log scales. The log scale is in $\log_{10}$.

The second-order PMD vector is projected onto the direction of the first-order vector ($\vec{\tau}_\omega \cdot \hat{p}$) to produce parallel and perpendicular components. This action conditions the SOPMD vector to $\hat{p}$. Expanding $\vec{\tau}_\omega$ from $\vec{\tau} = \tau \hat{p}$ makes

$$\vec{\tau}_\omega = \tau_\omega \hat{p} + \tau \hat{p}_\omega \quad (9.4.5a)$$

$$= \vec{\tau}_{\omega,\parallel} + \vec{\tau}_{\omega,\perp} \quad (9.4.5b)$$

The parallel component is the polarization-dependent chromatic dispersion. This component changes the chromatic dispersion of the fiber from D to $D_{\text{eff}} = D \pm \tau_{\omega,\parallel}$ and, accordingly, induces pulse compression or expansion [15, 50]. Nelson gives the wavelength dependence of this component [42]. The PDCD magnitude has two synonymous notations: $|\vec{\tau}|_\omega = \tau_{\omega,\parallel}$. The perpen-

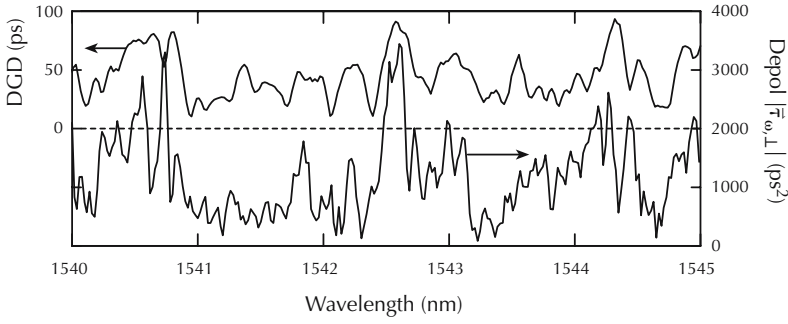


Fig. 9.7. Measurements of DGD τ and depolarization component $|\tilde{\tau}_{\omega,\perp}|$ of SOPMD from high-PMD fiber spool. Courtesy G. Shtengel [28].

dicular component is the depolarization, which is the tendency for the PMD vector to change direction. The effects of this component were treated in §8.2.

There is a strong tendency for $\tilde{\tau}_{\omega}$ to point away from $\tilde{\tau}$. The mean-square value of the PDCD and depolarization components in relation to that of the second-order vector are

$$\langle \tau_{\omega,\parallel}^2 \rangle = \frac{1}{9} \langle \tau_{\omega}^2 \rangle, \text{ and } \langle \tau_{\omega,\perp}^2 \rangle = \frac{8}{9} \langle \tau_{\omega}^2 \rangle \quad (9.4.6)$$

(Note that the rms of both components scale as $\tilde{\tau}^2$, as does the full SOPMD vector.) Even in comparison to a cartesian component, the PDCD component is diminished:

$$\langle \tau_{\omega,\parallel}^2 \rangle = \frac{1}{3} \langle \tau_{\omega,i}^2 \rangle \quad (9.4.7)$$

Depolarization is clearly the dominant component of SOPMD and is therefore the dominant impairment on an optical signal.

One may ask how do the SOPMD-projected components vary with a particular sample value of DGD for a fixed mean DGD. This question comes about when testing PMD compensators: when the DGD value is high, what form of SOPMD in a fiber can one expect? The answer is that the higher the DGD, the higher the expected depolarization. Ibragimov and Shtengel [28] show that the conditional expectations scale as

$$\langle \tau_{\omega,\parallel}^2 | \tau \rangle = \frac{1}{9} \langle \tau_{\omega}^2 \rangle \quad (9.4.8a)$$

$$\langle \tau_{\omega,\perp}^2 | \tau \rangle = \frac{2}{9} \left(\tau^2 \sqrt{3 \langle \tau_{\omega}^2 \rangle} + \langle \tau_{\omega}^2 \rangle \right) \quad (9.4.8b)$$

When the sample value τ^2 equals its mean-square value (9.4.4), the conditional expression (9.4.8b) reduces to (9.4.6). These expressions show that the rms PDCD magnitude is determined solely by the mean fiber DGD, while the rms depolarization magnitude scales with sample DGD. These relations are borne out in experiment. Figure 9.7 illustrates DGD and magnitude-depolarization

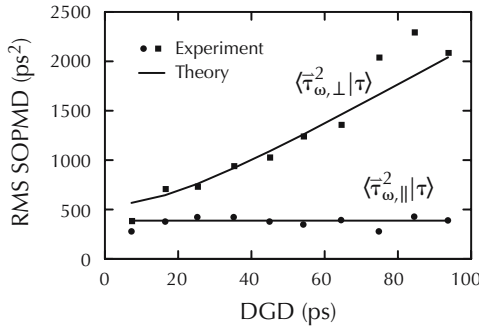


Fig. 9.8. Measurements of mean-square PDCD and depolarization as a function of DGD. Data taken on a single fiber with fixed mean DGD. Courtesy G. Shtengel [28].

measurements taken on a high-PMD fiber; the correlation is evident. Figure 9.8 shows an analysis of the data superimposed with the conditional distributions.

Finally, the joint probability distribution function (JPDF) of (τ, τ_ω) is a distribution central to the characterization and validation of receiver performance against PMD. The JPDF is not analytic, but can be calculated via brute force, special numerical techniques, or using importance sampling methods. A brute-force example of the JPDF is shown in Fig. 9.9. The lines indicate contours of constant probability. This calculation by P. J. Leo took a week running on a distributed computer network; the code formalism used the Stokes-based PMD concatenation rules for $\vec{\tau}$ and $\vec{\tau}_\omega$, thereby saving an order-of-magnitude in time compared to the equivalent Jones-matrix calculation. For comparison, a JPDF from measured field data is shown in Fig. 9.10.

The importance-sampling methods developed by Fogal, Biondini, and Kath [2, 3, 11, 12] also calculate the PMD vector in Stokes space, but with the innovation that a bias is added to the scattering of the polarization state from section to section. The bias can be made to accumulate preferentially large DGD values, large SOPMD values, or both. Importance sampling provides the formalism to rescale the resulting statistics by the probability of the underlying bias. Contours that extend to 10^{-20} have been demonstrated. Moreover, IS methods have been adopted to the expedient estimation of system outage probabilities due to PMD [33].

Finally, Forestieri has developed special numerical techniques to evaluate the PDCD component density and, subsequently, the joint first- and second-order density [13]. The author reports an efficient algorithm for fast evaluation of the JPDF, and shows that in comparison to the IS method the tails of the distribution fall more slowly.

The JPDF of (τ, τ_ω) is universal because of the scaling rules for DGD and SOPMD. In particular, the DGD and SOPMD axes scale as $\bar{\tau}$ and $\bar{\tau}^2$, respectively. Therefore, once the JPDF is calculated for some $\bar{\tau}$, it can be

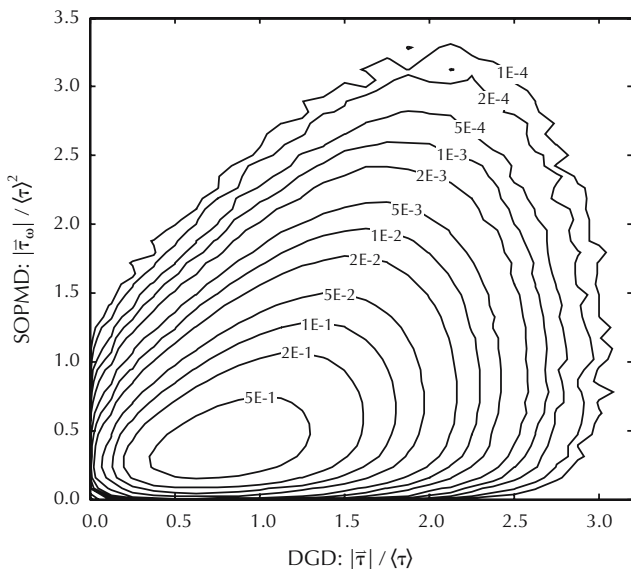


Fig. 9.9. Joint first- and second-order PMD density function. This is a universal distribution, scaled on the abscissa by $\langle\tau\rangle$ and on the ordinate by $\langle\tau\rangle^2$. Calculated from 10^9 realizations of a 2000-section fiber, using Stokes-based concatenation rules. Courtesy P. J. Leo.

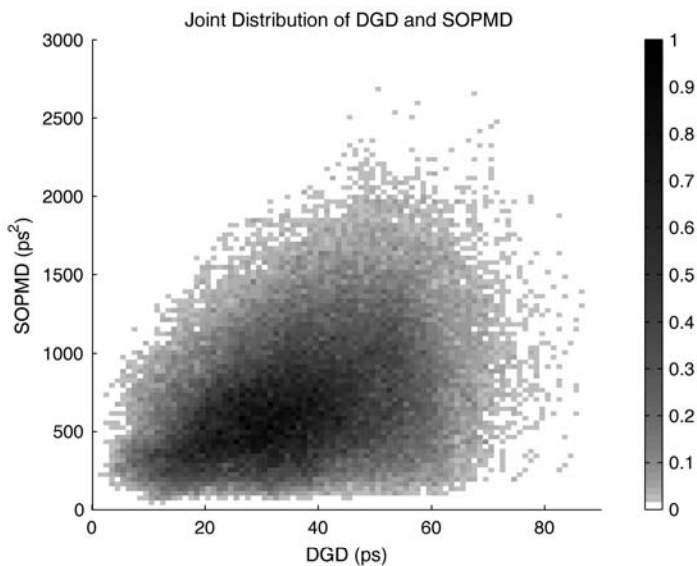


Fig. 9.10. Measured joint first- and second-order PMD magnitudes from fiber described in 9.11.

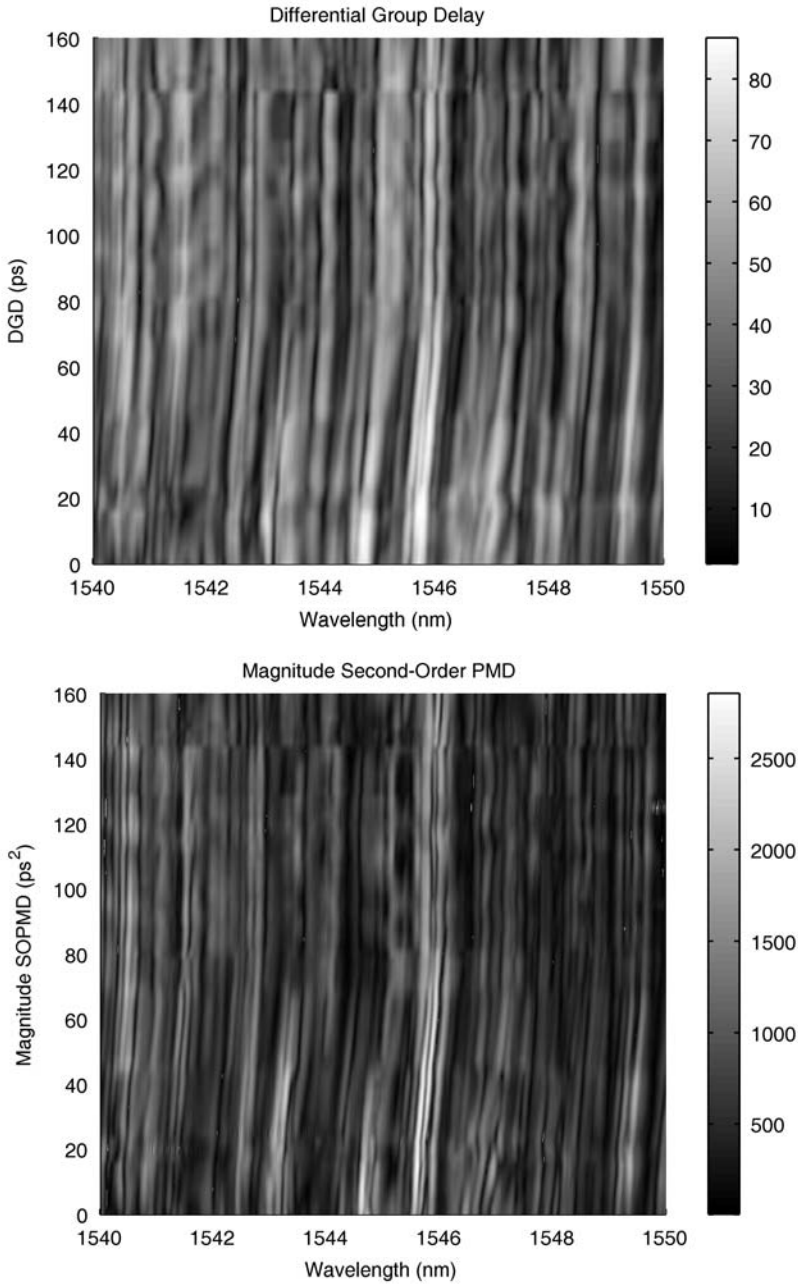


Fig. 9.11. Field measurements of differential-group delay and magnitude second-order PMD over 10 nm and 160 hrs. Notice the general adiabatic evolution is disrupted at about 145 hr. Courtesy D. Peterson, MCI [45].

arbitrarily rescaled to associate with any fiber. Such scaling, especially with the range available using IS, can be exploited to make solid predictions of the outage probability due to PMD in lightwave systems.

The nature of the JPDF also reveals the fallacy of the “PMD is DGD” concept in testing receivers and PMD compensators, whether optical or electrical. At almost any level of DGD there is nearly zero probability that SOPMD is zero, or even small. Yet most receiver testing at the time of this writing is done by introducing only DGD. Such results can significantly underestimate the receiver operation in real-world environments (but they glorify product performance to the parties that fund the work).

An example of rather adiabatic temporal behavior of first- and second-order PMD is shown in Fig. 9.11. The data was taken over 160 hrs in half-hour increments on installed fiber from a particularly old link [45]. At any time the wavelength variation is high, but the temporal evolution for this buried cable is slow. An exception exists at about 145 hrs where the fiber was disturbed (unintentionally). The disturbance must have been small because the spectra before and after the event are well correlated. A large disturbance will erase all memory of the past and set the fiber in a new state. This data also shows that for channels that fall on high first- and/or second-order states, the high PMD levels may persist for a long time before there is a change. A detailed study of the temperature dependence of PMD in installed cables is reported by Brodsky *et al.* [5].

9.4.2 Autocorrelation Functions

The autocorrelation function of the PMD vector is an essential descriptor for the characterization of PMD behavior. There are four properties that the autocorrelation function reveal:

- PMD is a wide-sense stationary process in frequency. The autocorrelation function is the only demonstrated method to connect ensemble averages with frequency averages.
- The PMD-vector correlation bandwidth depends only on the mean DGD $\bar{\tau}$.
- The variance of an estimated value of $\bar{\tau}$ derived from measurement depends inversely on the total measurement bandwidth.
- All moments of the PMD vector depend only on $\bar{\tau}$ and the moment order.

The autocorrelation function tells how large a frequency separation is necessary for two PMD vectors $\vec{\tau}(\omega)$ and $\vec{\tau}(\omega')$ to become statistically independent. Physically, PMD is generated by the fiber birefringence $\beta(z, \omega)$. The longitudinal evolution of the birefringence $\beta(z)$ is modelled as white noise or a concatenation of waveplates, while the frequency dependence $\beta(\omega)$ is linear: $\beta(\omega) = w\Delta n/c$. Recall from the PMD concatenation rules §8.2.4 that component PMD vectors precess about one another with frequency. For small frequency variation the output PMD vector retains its length and pointing

Table 9.2. Autocorrelation Functions for PMD Vector $\vec{\tau}$ and DGD Squared τ^2

Parameter	Form	Expression
PMD vector ^(a,b)	$\langle \vec{\tau}_{\omega'} \cdot \vec{\tau}_{\omega} \rangle$	$R_{\vec{\tau}}(\mathcal{W}) = \text{sinhc}(\mathcal{W}/6) \exp(-\mathcal{W}/6)$
DGD squared ^(c)	$\langle \tau_{\omega'}^2 \tau_{\omega}^2 \rangle$	$R_{\tau^2}(\mathcal{W}) = \frac{3}{5} \left(1 + \frac{2}{3} \frac{(1 - R_{\vec{\tau}}(\mathcal{W}))}{\mathcal{W}/6} \right)$
Estimator uncertainty ^(a,b,d,e)		$\bar{\tau}_{\text{est}}(B_f) \simeq \bar{\tau} \left(1 \pm \frac{0.9}{\sqrt{B_f \bar{\tau}}} \right)$
Bias error of statistic moments ^(f)		$\bar{\tau}(B_f) = \sqrt{\frac{8}{3\pi} \langle \bar{\tau}^2 \rangle} \pm \frac{8}{9\sqrt{2}} \frac{1}{2\pi B_f}$

$$\mathcal{W} = \Delta\omega^2 \langle \bar{\tau}^2 \rangle, \quad \Delta\omega^2 \langle \bar{\tau}^2 \rangle = \frac{3\pi^3}{2} (\Delta f \bar{\tau})^2, \quad R_{\tau} = \sqrt{R_{\tau^2}}, \quad R_{\hat{\tau}} = R_{\vec{\tau}}/R_{\tau}$$

$$\Delta\omega_{\text{PMD-ACF}} \sqrt{\langle \bar{\tau}^2 \rangle} \simeq 4.4, \quad \Delta f_{\text{PMD-ACF}} \bar{\tau} \simeq \frac{2}{\pi}$$

Note: Terms such as τ_{ω} denote $\tau(\omega)$, not a frequency derivative.

^(a,c) Shtaif and Mecozzi [39, 56, 57], ^(b) Karlsson and Brentel [31],

^(d) Gisin *et al.* [22], ^(e) The Poole and Favin coefficient was 0.66 rather than 0.9 [52],

^(f) Boroditsky *et al.* [4].

direction. As the frequency is further detuned, the PMD vector takes on a different length and pointing direction. The PMD autocorrelation function tells how large, on average, a frequency separation is necessary for two such PMD vectors to become statistically uncorrelated.

The normalized autocorrelation functions for the PMD vector $R_{\vec{\tau}}(\mathcal{W})$ and the DGD squared $R_{\tau^2}(\mathcal{W})$ are given in Table 9.2¹. Both functions are even functions of the argument, continuous, dependent only on frequency difference $\Delta\omega$ and not absolute frequency, and have a maximum at $\Delta\omega = 0$. These properties are characteristic of a wide-sense stationary process in frequency.

The PMD vector autocorrelation function originates with the dot-product of the PMD vector at two different frequencies, averaged over frequency:

$$\langle \vec{\tau}(\omega') \cdot \vec{\tau}(\omega) \rangle = \langle \bar{\tau}^2 \rangle R_{\vec{\tau}}(\mathcal{W}) \quad (9.4.9)$$

Figure 9.12(a) plots $R_{\vec{\tau}}(\mathcal{W})$ as a function of cyclic frequency $\Delta f \bar{\tau}$. The product $\Delta\omega^2 \langle \bar{\tau}^2 \rangle$ (or $\Delta f \bar{\tau}$) alone determines the form of the ACF. In an extensive study Lin and Agrawal have connected this and higher-order ACFs to pulse broadening and distortion [35].

¹ The function sinhc is a $\sinh$ function divided by its argument. In general a trigonometric function followed by the letter “c” is to be divided by its argument. At the origin, $\text{sinhc}(0) = 1$.

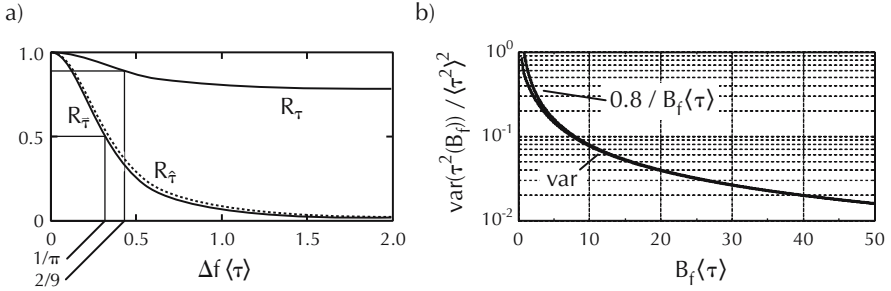


Fig. 9.12. Autocorrelation function and log variance. a) Autocorrelation function $R_{\vec{\tau}}$ of the PMD vector as function of $\Delta f \bar{\tau}$. FWHM is $\Delta f \bar{\tau} \simeq 2/\pi$. The PMD-vector ACF is dominated by depolarization, $R_{\hat{\tau}}$ (dashed curve). The root-mean-square DGD ACF has a secondary effect. b) Normalized $\log_{10}$ variance of estimated value $\langle \tau^2(B_f) \rangle$ as function of measurement bandwidth (in Hertz) B_f . Variance is well approximated by $16\sqrt{2}/(9\pi B_f \bar{\tau})$ for $B_f \bar{\tau} > 5$.

Especially important is the autocorrelation function bandwidth, defined as the full-width half-maximum (FWHM) of the function. In cyclic frequency the FWHM for the PMD vector is

$$\Delta f_{\text{PMD-ACF}} \bar{\tau} \simeq \frac{2}{\pi} \quad (9.4.10)$$

This is a useful relation to know. For instance, with $\bar{\tau} = 30$ ps, the correlation bandwidth is $\Delta f \simeq 21$ GHz, or $\Delta\lambda \simeq 0.16$ nm. Measurements of high-PMD fiber should at least have a 20 GHz resolution to capture the variation, but any higher resolution creates correlated measurements that oversample the PMD. Statistical estimates should be made only after the sample points of a high-resolution measurement are decimated (or are treated with other signal-processing methods) down to the autocorrelation bandwidth.

Another example is the failure of estimating mean DGD on installed fiber when the measurements are taken through a narrowband optical filter such as one port of a multiplexer. A typical channel bandwidth for such a filter is 40 GHz. The mean fiber DGD must be 32 ps or greater to have more than one statistically independent measurement within the filter passband.

As $R_{\vec{\tau}}$ is the autocorrelation of the PMD vector, the question is which component of that vector, the DGD or the pointing direction, dominates the decorrelation. In the preceding section it was determined that depolarization dominates PDCD, which is representative of the strong tendency for the PMD vector to change direction. The autocorrelation function also reflects this behavior.

The autocorrelation function of the mean-square DGD was derived by Shtaif and Mecozzi [57] to answer just this question. As the DGD-squared is the dot-product of the PMD vector with itself $\tau^2(\omega) = \vec{\tau}(\omega) \cdot \vec{\tau}(\omega)$, the correlation between $\tau^2(\omega')$ and $\tau^2(\omega)$ is

$$\langle \tau^2(\omega') \tau^2(\omega) \rangle = \frac{5}{3} \langle \bar{\tau}^2 \rangle^2 R_{\tau^2}(\mathcal{W}) \quad (9.4.11)$$

where R_{τ^2} is listed in Table 9.2. The autocorrelation functions R_τ for the DGD and $R_{\hat{\tau}}$ for the pointing-direction of the PMD are then extracted as shown in the table. These two functions are also plotted in Fig. 9.12(a). The correlation bandwidths are all about the same (with $\Delta f_{\text{ACF-DGD}} \bar{\tau} \simeq 4/9$) but the DGD ACF falls from $R_\tau(0) = 1$ to $R_\tau(\infty) = \sqrt{3/5}$, or about 22%. The PMD-vector ACF is dominated by the unit-vector ACF $R_{\hat{\tau}}$, consistent with the tendency of the PMD vector to change direction, or depolarize. The autocorrelation function for the PSP has been recently reported by Bao *et al.* [26].

The PMD-vector autocorrelation function can be modified to determine the weights of all moments of the PMD vector relative to the mean-square DGD [56]. Unlike the preceding ACFs, the moment relations apply to a single frequency. The two moment relations are

$$\langle \bar{\tau}^{(n-k)} \cdot \bar{\tau}^{(n+k+1)} \rangle = 0 \quad (9.4.12a)$$

$$\langle \bar{\tau}^{(n-k)} \cdot \bar{\tau}^{(n+k)} \rangle = (-1)^k \frac{(2n)!}{3^n (n+1)!} \langle \tau^2 \rangle^{n+1} \quad (9.4.12b)$$

where $\bar{\tau}^{(n)}$ refers to the $(n+1)$ order of $\bar{\tau}$. While even/odd moments vanish, like moments ($k = 0$) grow quickly with n . This is another reflection of the disorder and complicated structure of the DGD spectrum. Recall from the Fourier analysis of the DGD spectrum that an increase in the number of elementary PMD segments in a concatenation increases the number of Fourier components in the spectrum. Since $\langle \tau^2 \rangle \propto z$, higher moments grow increasingly quickly as the fiber length increases, consistent with the Fourier picture.

The factorial-function coefficient to $\langle \tau^2 \rangle$ in (9.4.12b) grows very quickly with n . The origin of this coefficient is the white noise that underlies the model. The moments of a Brownian motion are Hermite polynomials evaluated at the origin. The resulting Hermite coefficients grow as $(2n)!/n!$; this growth is reflected in the moments of the PMD vector since it is a derived process from Brownian motion of the birefringent vector. The relative growth of the coefficient for successive moments is

$$\frac{\langle \bar{\tau}^{(n)} \cdot \bar{\tau}^{(n)} \rangle}{\langle \bar{\tau}^{(n-1)} \cdot \bar{\tau}^{(n-1)} \rangle} = \frac{2n(2n-1)}{3(n+1)} \simeq \frac{4}{3} n$$

For large n the growth is linear in n , again consistent with Hermite polynomial behavior.

Autocorrelation Function Derivations

The PMD-related ACFs can be derived exclusively using a stochastic-calculus treatment of the birefringence and PMD vector [37]. The PMD evolution equation (9.3.1) on page 397 may be rewritten in SDE form as

$$d\vec{\tau} = d\vec{B}_z + \omega d\vec{B}_z \times \vec{\tau} \quad (9.4.13)$$

where

$$d\vec{B}_z = \vec{\beta}_\omega dz, \quad d\vec{B}_z \cdot d\vec{B}_z = \gamma^2 dz, \quad \text{and} \quad \langle dB_{z,j} dB_{z,k} \rangle = 0 \quad j \neq k \quad (9.4.14)$$

The differential $d\vec{B}_z$ is a three-entry column vector of i.i.d. Brownian motions that are the driving force in the evolution of $d\vec{\tau}$. The Brownian motion represents the local birefringence vector. The subscript z denotes that this motion is longitudinal in z and defines the domain over which averages will be taken. Brownian motion grows as $\sqrt{z}$, so according to the rules of stochastic calculus terms up to order z are included, thus $d\vec{B}_z \cdot d\vec{B}_z = \gamma^2 dz$, where γ^2 is the strength of the motion. (White noise is the formal derivative of Brownian motion: $dB_z = g_z dz$.)

The SDE (9.4.13) is in the Stratonovich form and must be translated into the Itô form. That translation generates an additional term, which when included makes

$$d\vec{\tau} = d\vec{B}_z + \omega d\vec{B}_z \times \vec{\tau} - \frac{\gamma^2 \omega^2}{3} \vec{\tau} dz \quad (9.4.15)$$

This correction term shifts the drift of $\vec{\tau}$ but not the random behavior.

The first calculation is the mean-square of the DGD, or τ^2 . Treating $\vec{\tau}$ as a stochastic variable, the differential of τ^2 using Itô's chain rule is

$$d(\tau^2) = 2\vec{\tau} \cdot d\vec{\tau} + d\vec{\tau} \cdot d\vec{\tau} \quad (9.4.16)$$

To clarify the following calculations, the Stratonovich form of $d\vec{\tau}$ (9.4.13) is first substituted into (9.4.16). Keeping only terms that survive an average, this partial solution gives

$$\begin{aligned} d(\tau^2) &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \omega^2 (d\vec{B}_z \times \vec{\tau})^2 \\ &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \omega^2 \left(\gamma^2 \tau^2 dz - (d\vec{B}_z \cdot \vec{\tau})^2 \right) \\ &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \frac{2}{3} \omega^2 \gamma^2 \tau^2 dz \end{aligned}$$

where

$$(d\vec{B}_z \cdot \vec{\tau})^2 = \sum_{k=1}^3 dB_{z,k} \cdot dB_{z,k} \tau_k^2 + \text{cross-terms}$$

and $dB_{z,k} \cdot dB_{z,k} = \gamma^2 dz/3$. Now, adding the Itô drift correction from (9.4.16) back into $d\vec{\tau}$ and keeping terms only up to order dz gives

$$d(\tau^2) = 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz \quad (9.4.17)$$

Averaging this expression over z eliminates terms of order $d\vec{B}_z$. Subsequent integration over length results in the mean-square evolution of the DGD

$$\langle \tau^2(z) \rangle = \gamma^2 z \quad (9.4.18)$$

Identification with (9.3.2) on page 397 gives $\gamma^2 = 2 \langle \tau_c^2 \rangle / L_C$.

The ACF for the PMD vector can now be calculated. Denoting $\vec{\tau}_\omega = \vec{\tau}(\omega)$ (rather than the frequency derivative of $\vec{\tau}$), the Itô differential of the dot product $\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}$ is

$$d(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}) = (d\vec{\tau}_\omega) \cdot \vec{\tau}_{\omega'} + \vec{\tau}_\omega \cdot (d\vec{\tau}_{\omega'}) + d\vec{\tau}_\omega \cdot d\vec{\tau}_{\omega'}$$

Substitution of (9.4.15) and keeping terms only up to order dz gives

$$\begin{aligned} d(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}) &= d\vec{B}_z \cdot (\vec{\tau}_\omega + \vec{\tau}_{\omega'}) + \Delta\omega \vec{\tau}_{\omega'} \cdot (d\vec{B}_z \times \vec{\tau}_\omega) + \gamma^2 dz \\ &- \frac{1}{3} \gamma^2 (\omega^2 + \omega'^2) (\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}) dz + \omega\omega' \left(\gamma^2 (\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}) dz - (d\vec{B}_z \cdot \vec{\tau}_\omega) (d\vec{B}_z \cdot \vec{\tau}_{\omega'}) \right) \end{aligned}$$

Subsequent averaging over z eliminates many terms. The average over the product of inner products in particular gives

$$\left\langle (d\vec{B}_z \cdot \vec{\tau}_\omega) (d\vec{B}_z \cdot \vec{\tau}_{\omega'}) \right\rangle = \frac{1}{3} \gamma^2 (\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}) dz \quad (9.4.19)$$

Completing the average over all terms gives the differential form of the autocorrelation

$$d \langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \rangle = \left(1 - \frac{1}{3} \Delta\omega^2 \langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \rangle \right) \gamma^2 dz$$

Integration gives

$$\langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \rangle = \frac{3}{\Delta\omega^2} \left(1 - \exp \left(-\frac{\Delta\omega^2 \gamma^2 z}{3} \right) \right) \quad (9.4.20)$$

Replacement of $\langle \bar{\tau}^2 \rangle = \gamma^2 z$ and normalization by $\langle \bar{\tau}^2 \rangle$ results in the normalized autocorrelation function listed in Table 9.2:

$$R_{\vec{\tau}}(\Delta\omega^2 \langle \bar{\tau}^2 \rangle) = \text{sinhc} \left(\frac{\Delta\omega^2 \langle \bar{\tau}^2 \rangle}{6} \right) \exp \left(-\frac{\Delta\omega^2 \langle \bar{\tau}^2 \rangle}{6} \right) \quad (9.4.21)$$

The limits of the ACF are $R_{\vec{\tau}}(0) = 1$ and $R_{\vec{\tau}}(\infty) = 0$. It is remarkable that the only terms that enter the ACF are the frequency difference $\Delta\omega$ and the mean-square DGD $\langle \bar{\tau}^2 \rangle$. The mean-square DGD in turn is directly proportional to the square of the mean fiber DGD. Once again the mean fiber DGD is the “unit” by which a PMD-related statistical quantity is governed.

Lastly, the autocorrelation of the DGD squared is calculated. The kernel of the calculation is $\tau_\omega^2 \tau_{\omega'}$, where $\tau_\omega^2 = \vec{\tau}_\omega \cdot \vec{\tau}_\omega$. Treating τ_ω^2 as a stochastic variable, the differential is

$$d(\tau_\omega^2 \tau_{\omega'}) = (d\tau_\omega^2) \tau_{\omega'} + \tau_\omega^2 (d\tau_{\omega'}) + (d\tau_\omega^2) (d\tau_{\omega'}) \quad (9.4.22)$$

Substitution of (9.4.17) into (9.4.22) and keeping terms only up to order dz gives

$$d(\tau_\omega^2 \tau_{\omega'}^2) = 2\tau_{\omega'}^2 (d\vec{B}_z \cdot \vec{\tau}_\omega) + 2\tau_\omega^2 (d\vec{B}_z \cdot \vec{\tau}_{\omega'}) \\ + (\tau_\omega^2 + \tau_{\omega'}^2) \gamma^2 dz + 4 (d\vec{B}_z \cdot \vec{\tau}_\omega) (d\vec{B}_z \cdot \vec{\tau}_{\omega'})$$

Averaging over z and using (9.4.19) leaves

$$d\langle \tau_\omega^2 \tau_{\omega'}^2 \rangle = 2\gamma^4 z dz + \frac{4\gamma^2}{3} \langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \rangle dz \quad (9.4.23)$$

where an Itô isometry removes the random component in $\langle \tau_\omega^2 \rangle$:

$$\langle \tau_\omega^2 \rangle = \left\langle \int d(\tau_\omega^2) \right\rangle = \left\langle \int 2\vec{\tau}_\omega \cdot d\vec{B}_z + \gamma^2 dz \right\rangle \\ = \int \gamma^2 dz$$

This average is no longer a function of ω . Substitution of the PMD-vector ACF into (9.4.23) makes for a straightforward integration, resulting in

$$\langle \tau_\omega^2 \tau_{\omega'}^2 \rangle = \langle \bar{\tau}^2 \rangle^2 + \frac{4\langle \bar{\tau}^2 \rangle}{\Delta\omega^2} - \frac{12}{\Delta\omega^4} \left(1 - \exp\left(-\frac{\Delta\omega^2 \langle \bar{\tau}^2 \rangle}{3}\right) \right) \quad (9.4.24)$$

Subsequent normalization as shown in (9.4.11) and rearrangement of terms produces the normalized ACF for the DGD squared:

$$R_{\tau^2}(\Delta\omega^2 \langle \bar{\tau}^2 \rangle) = \frac{3}{5} \left(1 + \frac{2}{3} \frac{(1 - R_{\bar{\tau}}(\Delta\omega^2 \langle \bar{\tau}^2 \rangle))}{\Delta\omega^2 \langle \bar{\tau}^2 \rangle / 6} \right) \quad (9.4.25)$$

As with the PMD-vector ACF, the DGD-squared ACF depends only on the frequency difference and mean fiber DGD. The limits of this autocorrelation are $R_{\tau^2}(0) = 1$ and $R_{\tau^2}(\infty) = 3/5$.

9.4.3 Mean-DGD Measurement Uncertainty

The PMD ACF gives the minimum bandwidth over which two neighboring PMD vectors are statistically independent. The PMD ACF can also be used to determine the uncertainty of an estimator of the mean DGD of a fiber. This important application has been studied by Gisin *et al.* [22], Karlsson and Brentel [31], Shtaf and Mecozzi [57], and Boroditsky *et al.* [4].

There is a difference in framework between the first three reports and the most recent. In particular, the relation between the mean-square DGD and average DGD is $\bar{\tau}^2 = 8/3\pi \langle \bar{\tau}^2 \rangle$ is considered exact in the former reports while

Boroditsky *et al.* explain that equality holds only over infinite bandwidth (or ensemble averages). In fact, in the limit of zero bandwidth there is an 8% systematic error between mean-square and average DGD. In the broadband regime $B \langle \bar{\tau}^2 \rangle > 30$ (discussed below), the error between the DGD moments is

$$\bar{\tau} = \sqrt{\frac{8}{3\pi} \langle \bar{\tau}^2 \rangle} \pm \frac{8}{9\sqrt{2}} \frac{1}{B} \quad (9.4.26)$$

where B is the full measurement bandwidth in radians. Measurements of low-PMD fibers are susceptible to this error.

Putting aside this systematic error for the moment, there are two ways to estimate the mean DGD from a measurement: average the DGD values across frequency, or average the DGD-squared values across frequency and take the square-root. The former is a straight average, while the latter gives the rms value. The studies show that the rms average gives a slightly better estimate. The variance of the rms estimate is detailed here, and Shtaif and Mecozzi give a brief comparison.

Consider the estimate of the mean-square DGD over a radian bandwidth $B = \omega_2 - \omega_1$:

$$\bar{\tau}_{\text{est}}^2(B) = \frac{1}{B} \int_B \tau^2(\omega) d\omega \quad (9.4.27)$$

The variance of $\bar{\tau}_{\text{est}}^2(B)$ is, by definition,

$$\begin{aligned} \text{var}(\bar{\tau}_{\text{est}}^2(B)) &= E[\bar{\tau}_{\text{est}}^2(B) \bar{\tau}_{\text{est}}^2(B)] - E^2[\bar{\tau}_{\text{est}}^2(B)] \\ &= \frac{1}{B^2} E \left[\int_B d\omega \int_B d\omega' \tau^2(\omega) \tau^2(\omega') \right] - \langle \bar{\tau}^2 \rangle^2 \\ &= \frac{1}{B} \int_B d\omega' \langle \tau^2(\omega) \tau^2(\omega' - \omega) \rangle - \langle \bar{\tau}^2 \rangle^2 \end{aligned}$$

where the double integral reduces to a single integral since the integrand depends only on the frequency difference and not absolute value. The last integrand has already been calculated, see (9.4.24). Additionally, it is more relevant to look at the normalized variance so comparisons can be made. Thus, normalizing the variance by $\langle \bar{\tau}^2 \rangle^2$ and computing the integral gives

$$\begin{aligned} \frac{\text{var}(\bar{\tau}_{\text{est}}^2(B))}{\langle \bar{\tau}^2 \rangle^2} &= 16 \left(\frac{4 - B^2 \langle \bar{\tau}^2 \rangle}{B^4 \langle \bar{\tau}^2 \rangle^2} \right) + \frac{32}{3} \left(\frac{B^2 \langle \bar{\tau}^2 \rangle - 6}{B^4 \langle \bar{\tau}^2 \rangle^2} \right) e^{-B^2 \langle \bar{\tau}^2 \rangle / 12} \\ &\quad + \frac{16\sqrt{3\pi}}{9} \frac{1}{B\sqrt{\langle \bar{\tau}^2 \rangle}} \text{erf} \left(\frac{B\sqrt{\langle \bar{\tau}^2 \rangle}}{2\sqrt{3}} \right) \end{aligned}$$

This function is plotted in Fig. 9.12(b). The asymptotic limit for $B \langle \bar{\tau}^2 \rangle > 30$ is

$$\frac{\text{var}(\bar{\tau}_{\text{est}}^2(B))}{\langle \bar{\tau}^2 \rangle^2} \simeq \frac{16\sqrt{3\pi}}{9} \frac{1}{B\sqrt{\langle \bar{\tau}^2 \rangle}} \quad (9.4.28)$$

Translation to bandwidth in Hertz and average DGD gives

$$B\sqrt{\langle\bar{\tau}^2\rangle} = \sqrt{\frac{3\pi^3}{2}} (B_f\bar{\tau}) \quad (9.4.29)$$

where $B = 2\pi B_f$. Thus,

$$\frac{\text{var}(\bar{\tau}_{\text{est}}^2(B_f))}{\langle\bar{\tau}^2\rangle^2} \simeq \frac{16\sqrt{2}}{9\pi} \frac{1}{B_f\bar{\tau}}, \quad B_f\bar{\tau} > 5 \quad (9.4.30)$$

Comparison of this approximate variance formula is given in Fig. 9.12(b).

The expected error is estimated by removing the normalization in expression (9.4.28). Since the estimated quantity is the mean-square of the DGD, plus and minus one standard deviation from the “true” value is $\bar{\tau}_{\text{est}}^2 \simeq \langle\bar{\tau}^2\rangle \pm \sqrt{\text{var}(\bar{\tau}_{\text{est}}^2)}$. Substitution of the variance by (9.4.28), translation to cyclic frequency (9.4.29), and converting both sides to mean DGD gives the expression for the estimator uncertainty:

$$\bar{\tau}_{\text{est}}(B_f) \simeq \bar{\tau} \left(1 \pm \frac{0.9}{\sqrt{B_f\bar{\tau}}} \right) \quad (9.4.31)$$

where the coefficient in the numerator comes from $\sqrt{16\sqrt{2}/(9\pi)}$. This coefficient agrees with Gisin [22]. Moreover, the expression shows that reduction of the standard deviation of the estimated value of $\bar{\tau}$ is a slow function: $1/\sqrt{B_f}$.

For example, consider an uncertainty of $\pm 10\%$: $B_f\bar{\tau} \simeq 110$. For a mean DGD of 10 ps, the required measurement bandwidth is $\sim 11,000$ GHz, or ~ 90 nm. To halve the uncertainty the bandwidth must be quadrupled. It is an open question whether an estimator with a faster convergence can be found. The square-root form for the mean-square estimator suggests an estimator based on the fourth-power of the DGD spectrum. This requires a higher-order autocorrelation function. Another way to increase the certainty of the mean DGD is to take multiple uncorrelated measurements over time. That uncertainty goes as $1/\sqrt{N}$ with N measurements; again a slow function but useful nonetheless.

Returning to Boroditsky *et al.*, the authors show that average DGD estimated from the magnitude SOPMD spectrum gives both an unbiased estimator and reduces the measurement uncertainty by 30%. The reduction in measurement uncertainty is equivalent to effectively doubling the measurement bandwidth. They further show that average DGD estimated from the PDCD spectrum along yields a better estimate of average DGD compared to direct mean-square DGD spectrum analysis. However, the magnitude SOPMD spectrum fluctuates roughly twice as fast as the corresponding DGD spectrum, which in turn requires greater care in measurement. The vector MPS technique should produce sufficiently accurate measurements. Moreover, the width of the PDCD density is only 1/9 that of the magnitude SOPMD spectrum, so again, care must be used in obtaining a sufficiently accurate measurement to effectively employ these techniques.

9.4.4 Discrete Waveplate Model

The analytic developments of this chapter are derived from a Brownian motion model of the local birefringence vector. The powerful tools of stochastic calculus and partial-differential equations are then employed to derive statistical properties of polarization and PMD. However, the cascaded waveplate model is very often used instead. The waveplate model concentrates differential delay into homogeneous segments and then abruptly mode-mixes between adjacent segments. The waveplate model is suitable as a good approximation in certain regimes as long as it is correctly constructed. While there are several variations, the model below converges to the correct statistics.

In the regime $L \gg L_C \gg L_B$, where L is the fiber length, the waveplate model illustrated in Fig. 9.13(a) gives a reasonable approximation for the PMD. In particular, the rms DGD statistics follow (9.4.1). The model uses N equal-length waveplates where each plate is $L_C/2$ long and there are $N_c = 2L/L_C$ waveplates in total. The statistics track for $N_c \gg 30$. The physical waveplate orientation is uniformly distributed on $[-\pi/2, \pi/2]$ and zero chirality is asserted. The birefringence (magnitude) of each plate is a random variable selected from a Rayleigh distribution.

There are two aspects to be worked out. One relates to the frequency bandwidth and step size and the other to the gaussian distributions of the cartesian components of the birefringence. First the frequency grid. To derive a good statistic, uncorrelated DGD values over a sufficiently wide bandwidth must be calculated. At the discrete level, the total bandwidth B_f comes from N_f points of step size Δf : $B_f = N_f \Delta f$. The minimum uncorrelated bandwidth for an average DGD $\bar{\tau}$ is $\Delta f \bar{\tau} = 2/\pi$, so the bandwidth-mean-DGD product is $B_f \bar{\tau} = 2N_f/\pi$. Substitution into the mean-DGD estimate (9.4.31) gives

$$\bar{\tau}_{\text{est}}(N_f) \simeq \bar{\tau} \left(1 \pm \frac{1.13}{\sqrt{N_f}} \right) \quad (9.4.32)$$

This is the basis on which N_f is set. For instance, $N_f = 500$ gives a standard deviation of 5%.

Next the birefringent distribution is determined. The mean-square DGD as a function of length (9.4.1) is rewritten as

$$\langle \tau^2(N_c) \rangle = \langle \beta_\omega^2 \rangle L_C^2 N_c \quad (9.4.33)$$

where $N_c = 2z/L_C$ and $\tau_c = \beta_\omega L_C$. The PMD distribution is Maxwellian, so a $3\pi/8$ scale factor relates its first and second moments. The birefringent distribution is Rayleigh, so the second-moment $\langle \beta_\omega^2 \rangle$ is a factor of two larger than the variance of the underlying i.i.d. gaussian cartesian-component distributions. Putting these together gives the variance of the component gaussian:

$$\sigma_{\beta,k}^2 = \frac{3\pi\bar{\tau}^2}{16L_C^2 N_c} \quad (9.4.34)$$

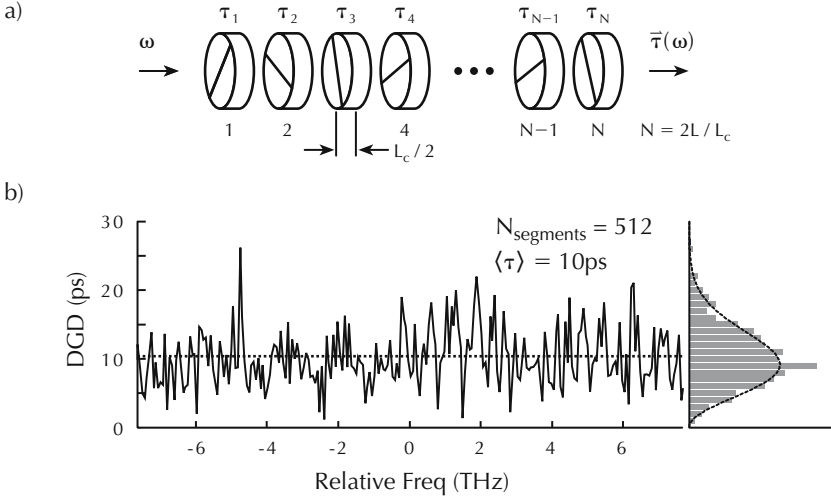


Fig. 9.13. Waveplate model of a fiber, good for the $L \gg L_C \gg L_B$ regime. a) Waveplates have uniformly distributed e -axis orientations and are each $L_C/2$ long. There are $N_c = 2L/L_C$ segments in total. The DGD per segment is determined from a Rayleigh distribution. b) Realization of a DGD spectrum for $N_c = 512$ and $N_f = 600$, given $\langle \tau \rangle = 10$ ps. The DGD distribution is shown to the right. The calculation uses large-enough frequency steps so that DGD values are statistically uncorrelated.

Each cartesian component of the birefringence is randomly picked from a normal distribution with density $\rho_{\beta,k} = N(0, \sigma_{\beta,k}^2)$. The segment birefringence magnitude is then $\beta_\omega = \sqrt{\beta_{\omega,1}^2 + \beta_{\omega,2}^2}$, with a corresponding segment DGD of $\tau_c = \beta_\omega L_C$.

Finally, the average Stokes-vector rotation per frequency step $\Delta\omega \bar{\tau}_c$ is estimated by combining (9.4.33) recast in terms of $\bar{\tau}$ and $\bar{\tau}_c$, and the uncorrelated frequency step $\Delta f \bar{\tau}$. The result is

$$\Delta\omega \bar{\tau}_c = \sqrt{\frac{3\pi^2}{2N_c}} \quad (9.4.35)$$

For instance, with $N_c = 512$, the average Stokes rotation per frequency step is $\Delta\omega \bar{\tau}_c \simeq 9.7^\circ$. This is a good check because a single step in excess of $\Delta\omega \bar{\tau}_c > \pi$ creates an ambiguity as to whether the PMD completed more than a half-revolution in one direction or less than half in the other direction.

A cropped spectral window of a DGD spectrum constructed in the manner outlined is illustrated in Fig. 9.13(b). The waveplate cascade was made with $N_c = 512$ waveplates calculated at $N_f = 600$ uncorrelated frequency points. The resulting distribution and its Maxwellian fit are plotted on the right. One instance of the concatenation using this number of waveplates and frequency

points is not sufficient to derive high-quality statistics, especially on the tail of the Maxwellian. But multiple runs of this cascade using different realizations of the birefringence vector will build up the statistics as a rate of $\sqrt{N}$. Note that varying the waveplate length as well as the aforementioned factors does not improve the statistics and works only to lower the convergence rate.

9.4.5 Karhunen-Loève Expansion of Brownian Motion

Other than the waveplate model above, the derivations in this chapter have relied on Brownian motion as the driving term for the evolution of various parameters. Brownian motion is often modelled on a microscopic, step-by-step level where the displacement for each step comes from choosing a random value from a gaussian density. This approach works but has no analytic expression.

A useful alternative is the Karhunen-Loeve (KL) expansion of Brownian motion [59]. The KL expansion gives a macroscopic view of the motion on an interval and guarantees the proper covariance. For Brownian motion the KL expansion on $[0, 1]$ is

$$B_z = \sum_{k=1}^{\infty} \frac{\sqrt{2}}{\pi \left(k + \frac{1}{2}\right)} \xi_k \sin \left(\pi \left(k + \frac{1}{2}\right) z\right) \quad (9.4.36)$$

where ξ_k are random variables with density $N(0, 1)$. Each term in the summation spans the entire interval. Higher values of k produce higher oscillations but with lower amplitudes. In practice the sum is taken large enough to fill in the necessary spatial resolution and is thereafter truncated. Figure 9.14(a) shows four sample paths generated by (9.4.36).

The KL expansion is the function-space analogue of a Markov process at the discrete level. On this level a Markov process is determined purely by its covariance matrix A . The eigenvectors and values are found from the equation $Ax = \lambda x$. The spectral theorem gives the entries in A in terms of its eigenvectors and values: $A = \sum_{k=1}^n \lambda_k v_k v_k^T$. On the continuous level, the eigenvalue equation is

$$\int_0^1 K(z, y) \varphi(y) dy = \lambda \varphi(z) \quad (9.4.37)$$

where the covariance of the process is $K(z, y) = z \wedge y$. The symbol $\wedge$ means the minimum of the two quantities. The spectral theorem says that the entries of the covariance function, entries which are now functions not vectors, are

$$K(z, y) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(y) \varphi_k(z) \quad (9.4.38)$$

where λ_k are the eigenvalues of (9.4.37) and $\varphi_k(z)$ are its eigenvectors.

The eigenvalue equation is solved by substituting in the covariance of Brownian motion. This gives

$$\int_0^1 (z \wedge y) \varphi(y) dy = \lambda \varphi(z)$$

The integral is separated into two pieces to make an integral equation:

$$\int_0^z y \varphi(y) dy + \int_z^1 z \varphi(y) dy = \lambda \varphi(z) \quad (9.4.39)$$

This is an integral equation which can be solved by taking successive derivatives. Looking ahead a couple of steps, the differential equation produced by the above integral equation is second order, and therefore requires two boundary conditions to be solved uniquely. One boundary condition is found directly from this integral equation: when $z = 0$ then $\varphi(0) = 0$.

Differentiating both sides of (9.4.39) with respect to z makes

$$z \varphi(z) + \int_z^1 \varphi(y) dy - z \varphi(z) = \lambda \varphi'(z)$$

where $\varphi'(z)$ denotes the first derivative with respect to z . After cancelling the two terms in the left, a second boundary condition is determined: for $z = 1$ $\varphi'(1) = 0$. Differentiating again gives

$$\varphi(z) = -\lambda \varphi''(z) \quad (9.4.40)$$

This ODE is solved subject to the boundary conditions $\varphi(0) = \varphi'(1) = 0$. The general solution is

$$\varphi(z) = A \sin(az) + B \cos(bz)$$

where the derivatives yield

$$\varphi'(z) = aA \cos(az) - bB \sin(bz)$$

$$\varphi''(z) = -a^2 A \sin(az) - b^2 B \cos(bz)$$

The boundary conditions restrict the four unknown coefficients in the following way:

$$\begin{aligned} \varphi(0) = 0 &\longrightarrow B = 0 \\ \varphi'(1) = 0 &\longrightarrow aA \cos(a) = 0 \end{aligned}$$

Substitution of $\varphi(z)$ into the differential equation (9.4.40) gives the definition for a : $a = 1/\sqrt{\lambda}$. Summarizing these restrictions, the solution thus far is

$$\varphi(z) = A \lambda^{-1/2} \sin \left(z \lambda^{-1/2} \right) \quad (9.4.41)$$

subject to the condition that

$$A \lambda^{-1/2} \cos \left(\lambda^{-1/2} \right) = 0$$

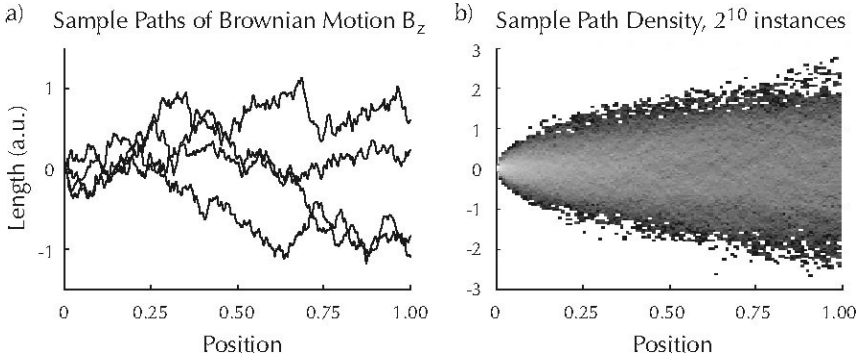


Fig. 9.14. Sample paths of Brownian motion created by the Karhunen-Loeve expansion. a) Four sample paths on the interval $[0, 1]$. b) Density of 2^{10} sample paths on the interval. As expected, the density width increases as square-root of the length.

This boundary condition is satisfied when

$$\frac{1}{\sqrt{\lambda_k}} = \pi \left(k + \frac{1}{2} \right), \quad k \in \mathbb{Z}$$

Rearranging, the eigenvalue is

$$\lambda_k = \left(\pi \left(k + \frac{1}{2} \right) \right)^{-2} \quad (9.4.42)$$

Finally, the coefficient A is determined from the orthogonality condition of $\varphi(z)$:

$$\int_0^1 \varphi_k^2(y) dy = 1$$

This is satisfied when $A = \sqrt{2}$. The eigenvector function is therefore

$$\varphi_k(z) = \sqrt{2} \sin \left(\pi \left(k + \frac{1}{2} \right) z \right), \quad z \in [0, 1], k \in \mathbb{Z} \quad (9.4.43)$$

The KL expansion for a general gaussian process G_z is

$$G_z = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k \varphi_k(z) \quad (9.4.44)$$

where the random variable ξ is characterized by $E[\xi_k] = 0$ and $E[\xi_k^2] = 1$. One can show that $E[G_y G_z]$ reproduces the correct covariance (9.4.38). Substitution of (9.4.43) into (9.4.44) makes the expression (9.4.36) stated at the beginning.

Figure 9.14(b) shows the density of B_z on the interval $[0, 1]$ for 2^{10} sample paths. As expected the density grows as the square-root of the length.

9.5 PDL Statistics

The statistics of polarization-dependent loss are derived in an analogous manner as those for polarization and PMD. The local differential loss is modelled as a white-noise process and the cumulative PDL is determined by the diffusion of the PDL vector $\vec{\Gamma}$ along the fiber. Concomitant with the assumption of no chiral birefringence, circular dichroism is excluded from the PDL model.

When immersed in a random birefringence medium, the axes of minimum and maximum transmission of a local differential loss element $\vec{\alpha}$ are scrambled from one point to another. Recall the evolution of the cumulative PDL vector in the presence of birefringence from (8.3.16) on page 377:

$$\frac{d\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - (\vec{\alpha} \cdot \vec{\Gamma}) \vec{\Gamma} \quad (9.5.1)$$

The cross-product term spins the cumulative PDL vector about the local birefringent axis $\vec{\beta}$, scrambling its orientation. Propagation through multiple randomly oriented birefringent elements drives the PDL vector toward isotropic coverage of the Poincaré sphere. The second term pulls the cumulative PDL vector toward the local element, while the last term governs the growth and decay of $\vec{\Gamma}$.

The statistics for PDL reported in the literature are based on PDL immersed in random birefringence [10, 19, 38, 64]. PMD statistics, by contrast, were derived in the absence of PDL. The reason PMD is included in PDL statistics is because a long concatenation of pure PDL is not likely in a telecommunications link. The consequence of PMD inclusion is that the local differential loss is treated as three-dimensional i.i.d. white noise in Stokes space. The correlations of the white-noise vector $\vec{\alpha}$ are

$$\langle \alpha_j(z) \rangle = 0, \quad \langle \alpha_j(z) \alpha_k(z') \rangle = \frac{\sigma_\alpha^2}{3} \delta_{j,k} (z - z') \quad (9.5.2)$$

where σ_α^2 is the strength of the disturbance. A differential Brownian vector is defined as $d\vec{B}_z = \vec{\alpha} dz$ such that $d\vec{B}_z \cdot d\vec{B}_z = \sigma_\alpha^2 dz$.

The evolution equation (9.5.1) with the cross-product removed (as its effect averages to zero in the isotropic PDL model) is rewritten in SDE form as

$$d\vec{\Gamma} = \left(I - \vec{\Gamma} \vec{\Gamma} \cdot \right) d\vec{B}_z$$

This diffusion equation is interpreted in the Stratonovich sense and must be translated to Itô form in order to use Itô calculus. The translation makes [38]

$$d\vec{\Gamma} = -\frac{\sigma_\alpha^2}{3} (2 - \Gamma^2) \vec{\Gamma} dz + \left(I - \vec{\Gamma} \vec{\Gamma} \cdot \right) d\vec{B}_z$$

The diffusion generator for this equation is

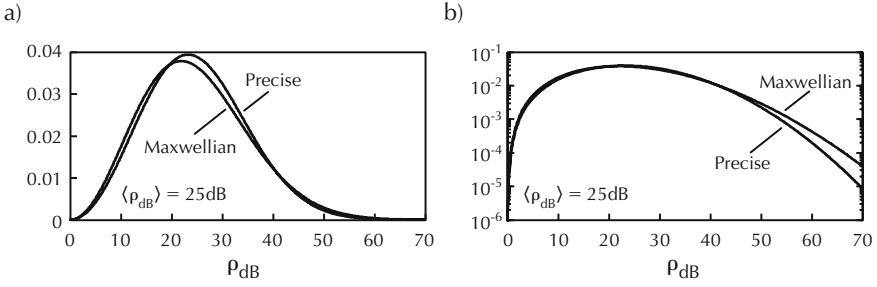


Fig. 9.15. PDL probability density and Maxwellian approximation versus decibel value, linear and semi-log scales. The log scale is in $\log_{10}$.

$$G = -\frac{\sigma_\alpha^2 (2 - \Gamma^2)}{3} \left(\sum_{i=1}^3 \Gamma_i \frac{\partial}{\partial \Gamma_i} + \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \Gamma_i \Gamma_j \frac{\partial^2}{\partial \Gamma_i \partial \Gamma_j} \right) + \frac{\sigma_\alpha^2}{6} \sum_{i=1}^3 \frac{\partial^2}{\partial \Gamma_i^2} \quad (9.5.3)$$

where $(\sigma \sigma^T) = (I - \vec{\Gamma} \vec{\Gamma}^T) (I - \vec{\Gamma} \vec{\Gamma}^T)^T = (I - (2 - \Gamma^2) \vec{\Gamma} \vec{\Gamma}^T)$. One can now calculate expectations of the diffusion using Kolmogorov's backward equation (9.2.5) on page 394.

It is an oddity of PDL that diffusions of Γ_k , Γ^n , $\vec{\Gamma}^n$ and the like are difficult to solve while those of the logarithm of $T_{\max}/T_{\min}$ make closed solutions. It would appear that the $(2 - \Gamma^2)$ coefficient is only cleanly removed when an logarithmic function is used. Fukada does, however, succeed in expressing the probability densities in linear terms [18]. For the present, the moments of the PDL magnitude expressed in decibels are used. Recall the definition:

$$\rho_{dB} = 10 \log_{10} \left(\frac{1 + \Gamma}{1 - \Gamma} \right) \quad (9.5.4)$$

It is particular to PDL that even though the cartesian components of the local PDL vectors are modelled as isotropic i.i.d. Gaussian random variables the cumulative distribution is not strictly Maxwellian. Shtaif and Mecozzi show that for low cumulative PDL (25 dB or less, although indeed this is extremely high for a lightwave system) the distribution is approximately Maxwellian [38]. Galtarossa and Palmieri use the diffusion generator to calculate the PDL distribution exactly and validate the Shtaif and Mecozzi approximation [19].

The details of the Galtarossa and Palmieri calculation are laborious but indeed elegant. Central to their calculations are the functionals $\psi = \mathbf{p}^{2k}$ and $\psi' = \mathbf{p}^{2k+1}/\Gamma$. From this they construct the characteristic function (CF) of the $\mathbf{p}^2$ density and, after proof of convergence, inverse-Fourier transform the CF to the density proper. Subsequent conversion to the density of $\mathbf{p}$ ($= \rho_{dB}$) makes

$$\rho_p(p, \tilde{z}) = \frac{2p^2}{\gamma^3 \sqrt{2\pi} \tilde{z}^3} \frac{\sinh(p/\gamma)}{(p/\gamma)} \exp\left(-\frac{p^2}{2\gamma^2 \tilde{z}}\right) \exp\left(-\frac{\tilde{z}}{2}\right) \quad (9.5.5)$$

for $p \geq 0$, where $\tilde{z} = z\sigma_\alpha^2/3$ and $\gamma = 20 \log 10e \simeq 8.868$. This density is plotted on linear and semi-log scale in Fig. 9.15. The first and second moments are

$$\langle \mathbf{p}(\tilde{z}) \rangle = \gamma \left(\sqrt{\frac{2\tilde{z}}{\pi}} e^{-\tilde{z}/2} + (1 + \tilde{z}) \operatorname{erf}\left(\frac{\tilde{z}}{2}\right) \right) \quad (9.5.6a)$$

$$\langle \mathbf{p}^2(\tilde{z}) \rangle = \gamma^2 (\tilde{z} + 3) \tilde{z} \quad (9.5.6b)$$

In the limit of large $\tilde{z}$ the cumulative PDL grows linearly with $\tilde{z}$. The Shtaif and Mecozzi second moment is, by comparison,

$$\langle \mathbf{p}^2(\tilde{z}) \rangle = \frac{9\gamma^2}{2} (e^{2\tilde{z}/3} - 1) \quad (9.5.7)$$

Both expressions are equal to second order in $\tilde{z}$.

The Maxwellian approximation to the PDL density function (9.5.5) written in terms of the second moment is

$$\rho_p(p, \tilde{z}) \simeq \frac{2p^2}{\sqrt{2\pi} (\langle \mathbf{p}^2(\tilde{z}) \rangle / 3)^3} \exp\left(-\frac{p^2}{2 (\langle \mathbf{p}^2(\tilde{z}) \rangle / 3)}\right), \quad p \geq 0 \quad (9.5.8)$$

Figure 9.15 shows a comparison between the Maxwellian approximation and the precise PDL density. The cumulative mean PDL is $\langle \mathbf{p} \rangle = 25$ dB; for lower mean PDL's the approximation improves. However, the extreme case plotted here indicates the divergence of the true distribution from the Maxwellian: the tails of the precise distribution fall faster for high PDL values. This means that the cartesian-component distributions of the logarithm PDL vector $\vec{\mathbf{p}}_{\text{dB}}$ have slightly shorter tails than the Gaussian distributions of $\vec{B}_z$.

Finally, PDL and PMD are complementary regarding partial polarization. PMD tends to depolarize a perfectly polarized input, while PDL tends to repolarize a perfectly unpolarized input. While not presented here, the statistics of PDL-induced repolarization are derived by Menyuk *et al.* and have an approximate Maxwellian form as well [41].

References

1. L. Arnold, *Stochastic Differential Equations: Theory and Applications*. Malabar, Florida: Krieger Publishing Company, 1992, reprinted from original 1974 edition.
2. G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, no. 2, pp. 310–312, Feb. 2002.
3. G. Biondini, W. Kath, and C. Menyuk, "Importance sampling for polarization-mode dispersion: techniques and applications," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.
4. M. Boroditsky, M. Brodsky, N. J. Frigo, P. Magill, and M. Shtaif, "Improving the accuracy of mean DGD estimates by analysis of second-order PMD statistics," *IEEE Photonics Technology Letters*, vol. 16, no. 3, pp. 792–794, Mar. 2004.
5. M. Brodsky, P. Magill, and N. J. Frigo, "Polarization-mode dispersion of installed recent vintage fiber as a parametric function of temperature," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 209–211, Jan. 2004.
6. X. Chen, M. Li, and D. A. Nolan, "Polarization mode dispersion of spun fibers: An analytical solution," *Optics Letters*, vol. 27, no. 5, pp. 294–296, Mar. 2002.
7. F. Curti, B. Daino, G. de Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," *Journal of Lightwave Technology*, vol. 8, no. 8, pp. 1162–1166, Aug. 1990.
8. W. B. Davenport, *Probability and Random Processes*. New York: McGraw-Hill, Inc., 1970.
9. M. C. de Lignie, H. Nagel, and M. van Deventer, "Large polarization mode dispersion in fiber optic cables," *Journal of Lightwave Technology*, vol. 12, no. 8, pp. 1325–1329, Aug. 1994.
10. A. El Amari, N. Gisin, B. Perny, H. Zbinden, and C. W. Zimmer, "Statistical prediction and experimental verification of concatenations of fiber optic components with polarization dependent loss," *Journal of Lightwave Technology*, vol. 16, no. 3, pp. 332–339, Mar. 1998.
11. S. L. Fogal, G. Biondini, and W. L. Kath, "Correction to: Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, pp. 1487–1489, 2002.
12. —, "Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, no. 9, pp. 1273–1275, Sept. 2002.
13. E. Forestieri, "A fast and accurate method for evaluating joint second-order PMD statistics," *Journal of Lightwave Technology*, vol. 21, no. 11, pp. 2942–2952, Nov. 2003.
14. G. J. Foschini, L. Nelson, R. Jopson, and H. Kogelnik, "Probability densities of the second order polarization mode dispersion including polarization dependent chromatic dispersion," *IEEE Photonics Technology Letters*, vol. 12, no. 3, pp. 293–295, Mar. 2000.
15. G. J. Foschini, R. M. Jopson, L. E. Nelson, and H. Kogelnik, "The statistics of PMD-induced chromatic fiber dispersion," *Journal of Lightwave Technology*, vol. 17, no. 9, pp. 1560–1565, Sept. 1999.
16. G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1882–1886, Dec. 1991.

17. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," *Journal of Lightwave Technology*, vol. 9, no. 11, p. 1439, Nov. 1991.
18. Y. Fukada, "Probability density function of polarization dependent loss (pdl) in optical transmission system composed of passive devices and connecting fibers," *Journal of Lightwave Technology*, vol. 20, no. 6, pp. 953–964, June 2002.
19. A. Galtarossa and L. Palmieri, "The exact statistics of polarization-dependent loss in fiber-optic links," *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 57–59, Jan. 2003.
20. A. Galtarossa, L. Palmieri, and D. Sarchi, "Measure of spin period in randomly birefringent low-PMD fibers," *IEEE Photonics Technology Letters*, vol. 16, no. 4, pp. 1131–1133, Apr. 2004.
21. A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, "Statistical characterization of fiber random birefringence," *Optics Letters*, vol. 25, no. 18, pp. 1322–1324, Sept. 2000.
22. N. Gisin, B. Gisin, J. Von der Weid, and R. Passy, "How accurately can one measure a statistical quantity like polarization-mode dispersion?" *IEEE Photonics Technology Letters*, vol. 8, no. 12, pp. 1671–1673, 1996.
23. N. Gisin, "Solutions of the dynamical equation for polarization dispersion," *Optics Communications*, vol. 86, pp. 371–373, 1991.
24. N. Gisin and J. P. Pellaux, "Polarization mode dispersion: Time versus frequency domains," *Optics Communications*, vol. 89, pp. 316–323, May 1992.
25. J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: <http://www.pnas.org>
26. S. Hadjifaradji, L. Chen, D. S. Waddy, and X. Bao, "Autocorrelation function of the principal state of polarization vector for systems having PMD," *IEEE Photonics Technology Letters*, vol. 16, no. 6, pp. 1489–1491, June 2004.
27. B. Huttner, J. Reecht, N. Gisin, R. Passy, and J. Weid, "Distributed beatlength measurement in single-mode fibers with optical frequency-domain reflectometry," *Journal of Lightwave Technology*, vol. 20, no. 5, pp. 828–835, May 2002.
28. E. Ibragimov, G. Shtengel, and S. Suh, "Statistical correlation between first and second order PMD," *Journal of Lightwave Technology*, vol. 20, no. 4, pp. 586–590, 2002.
29. I. P. Kaminow, "Polarization in optical fibers," *IEEE Journal of Quantum Electronics*, vol. QE-17, no. 1, pp. 15–22, 1981.
30. M. Karlsson, "Probability density functions of the differential group delay in optical fiber communication systems," *Journal of Lightwave Technology*, vol. 19, no. 3, pp. 324–331, Mar. 2000.
31. M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," *Optics Letters*, vol. 24, no. 14, pp. 939–941, July 1999.
32. P. J. Leo, G. R. Gray, G. J. Simer, and K. B. Rochford, "State of polarization changes: Classification and measurement," *Journal of Lightwave Technology*, vol. 21, no. 10, pp. 2189–2193, Oct. 2003.
33. I. T. Lima, A. O. Lima, J. Zweek, and C. R. Menyuk, "Efficient computation of outage probabilities due to polarization effects in a WDM system using a reduced stokes model and importance sampling," *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 45–47, Jan. 2003.

34. I. T. Lima, R. Khosravani, P. Ebrahimi, E. Ibragimov, C. R. Menyuk, and A. E. Willner, "Comparison of polarization mode dispersion emulators," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1872–1881, Dec. 2001.
35. Q. Lin and G. P. Agrawal, "Correlation theory of polarization mode dispersion in optical fibers," *Journal of the Optical Society of America B*, vol. 20, no. 2, pp. 292–301, Feb. 2003.
36. D. Marcuse, C. R. Menyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 15, no. 9, pp. 1735–1746, Sept. 1997.
37. A. Mecozzi, private communication, 2004.
38. A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 3, pp. 313–315, Mar. 2002.
39. —, "Study of the two-frequency moment generating function of the PMD vector," *IEEE Photonics Technology Letters*, vol. 15, no. 12, pp. 1713–1715, Dec. 2003.
40. C. R. Menyuk and P. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," *Journal of the Optical Society of America B*, vol. 11, no. 7, pp. 1288–1296, July 1994.
41. C. Menyuk, D. Wang, and A. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," *IEEE Photonics Technology Letters*, vol. 9, no. 9, pp. 1247–1249, Sept. 1997.
42. L. Nelson, R. Jopson, H. Kogelnik, and G. J. Foschini, "Measurement of depolarization and scaling associated with second-order polarization mode dispersion in optical fibers," *IEEE Photonics Technology Letters*, vol. 11, no. 12, pp. 1614–1617, Dec. 1999.
43. D. Nolan, X. Chen, and M.-J. Li, "Fibers with low polarization-mode dispersion," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1066–1077, Apr. 2004.
44. B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th ed. New York: Springer, 1998.
45. D. L. Peterson, B. C. Ward, K. B. Rochford, P. J. Leo, and G. Simer, "Polarization mode dispersion compensator field trial and field fiber characterization," *Optics Express*, vol. 10, no. 14, pp. 614–621, July 2002. [Online]. Available: <http://www.opticsexpress.org/>
46. S. M. Pietralunga, M. Ferrario, P. Martelli, and M. Martinelli, "Direct observation of local birefringence and axis rotation in spun fiber with centimetric resolution," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 212–214, Jan. 2004.
47. A. Pizzinat, L. Palmieri, B. S. Marks, C. R. Menyuk, and A. Galtarossa, "Analytical treatment of randomly birefringent periodically spun fibers," *Journal of Lightwave Technology*, vol. 21, no. 12, pp. 3355–3363, Dec. 2003.
48. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Optics Letters*, vol. 13, no. 8, pp. 687–689, Aug. 1988.
49. —, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," *Optics Letters*, vol. 14, no. 10, pp. 523–525, 1989.
50. C. D. Poole and C. R. Giles, "Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber," *Optics Letters*, vol. 13, no. 2, pp. 155–157, 1988.

51. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," *Optics Letters*, vol. 16, no. 6, pp. 372–374, 1991.
52. C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, 1994.
53. S. C. Rashleigh, "Origins and control and polarization effects in single-mode fiber," *Journal of Lightwave Technology*, vol. LT-1, no. 2, pp. 312–331, 1983.
54. S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single mode fibers," *Optics Letters*, vol. 3, no. 2, pp. 60–62, 1978.
55. H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. New York: Springer, 1989.
56. M. Shtaiif and A. Mecozzi, "Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states," *IEEE Photonics Technology Letters*, vol. 12, no. 1, pp. 53–55, Jan. 2000.
57. —, "Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion," *Optics Letters*, vol. 25, no. 10, pp. 707–709, May 2000.
58. Y. Tan, J. Yang, W. L. Kath, and C. R. Menyuk, "Transient evolution of the polarization-dispersion vector's probability distribution," *Journal of the Optical Society of America B*, vol. 19, no. 5, pp. 992–1000, May 2002.
59. E. Vanden-Eijnden, private communication, Courant Institute of Mathematical Sciences, New York University, N.Y., 2003.
60. P. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 19, no. 19, pp. 1517–1519, Oct. 1994.
61. —, "Anisotropic diffusion of the state of polarization in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 20, no. 24, pp. 2493–2495, Dec. 1995.
62. —, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 14, no. 2, pp. 148–157, Feb. 1995.
63. J. Yang, W. L. Kath, and C. R. Menyuk, "Polarization mode dispersion probability distribution for arbitrary distances," *Optics Letters*, vol. 26, no. 19, pp. 1472–1474, Oct. 2001.
64. M. Yu, C. Kan, M. Lewis, and A. Sizmann, "Statistics of polarization-dependent loss, insertion loss, and signal power in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 12, pp. 1695–1697, Dec. 2002.

Review of Polarization Test and Measurement

There are two aspects to test and measurement that are addressed in industry: the measurement and quantification of polarization effects such as SOP, PMD, and PDL; and the calibrated generation of these effects. Most measurement techniques use a polarimeter to measure the Stokes parameters of the light directly. Using predetermined and calibrated launch states of polarization at the input, the resulting Stokes parameters may be analyzed to determine SOP, PMD, and PDL. In order for such equipment to adhere to traceable standards, test artifacts for these effects have to be available. The National Institute for Standards and Technology in the United States fulfills this role, and the Telecommunication Industry Association (TIA), the International Telecommunications Union (ITU), and the International Electrotechnical Commission (IEC) develop standard test methodologies.

Polarization-mode dispersion, PDL, and sometimes SOP fluctuation generally cause impairments in an optical communications link. To quantify the impairment, it is necessary to have test instrumentation that programmatically generates these effects. To date there is no standard way to generate SOP fluctuation calibrated to natural speeds, such as the SOP change in aerial fiber or, on the other extreme, under-sea fiber. P. Leo *et al.* offer one proposal [72]. Artifacts for PMD and PDL are available as are instruments that make PMD and PDL in a calibrated manner. Since PMD and PDL can interact to make impairments worse than either effect alone, instruments that make PDL interspersed with PMD are necessary; some initial demonstrations have been reported [98].

This chapter gives an overview of the current state-of-the-art in polarization test and measurement. The latter half of the chapter is dedicated to programmable PMD generation, a topic that has not been covered as a whole before.

10.1 SOP Measurement

The starting point for any measurement of polarization state and its fluctuation, polarization-dependent loss and gain, and polarization-mode dispersion is the direct measurement of the Stokes parameters of the light.

A trade-off exists between ease of assembly and calibration versus speed of a polarimeter. The “rotating waveplate” polarimeter, recently analyzed by Williams for error sources [106], requires only a quarter-wave waveplate, a linear polarizer, and a detector. Such a simple construction leads to a precision polarimeter, yet the read-out speed is limited by the waveplate rotation rate. In contrast, the staring polarimeter literally stares at the incoming light through a sequence of waveplates and polarizers, and detects the conditioned light on a segmented detector. The read-out rate is limited by the detector speed. A staring polarimeter requires several waveplates, a polarizer, and at least four detectors. Balancing the detector responsivity and calibrating for waveplate misalignment is more arduous than calibrating the rotating waveplate type, which generally leads to a lower-precision polarimeter. Its solid-state construction and fast read-out, however, offer advantages in live-traffic applications and for field-portable instruments. Hague provides a review of various early polarimeters [43].

Collett may have been the first to build a staring polarimeter, which he used to measure the polarization of nanosecond optical pulses [9]. The conversion of a Jones vector to a Stokes vector was detailed in §1.4 and §2.5.1. Collett directly implements this scheme by making six simultaneous polarimetric measurements and inferring the seventh. His polarimeter is called a differential polarimeter because the intensity of orthogonal states pairs, e.g. (I_x, I_y) , is measured directly, as opposed to inferred. Differential polarimetry is sometimes used in astrophysics and biomedical applications because of its high sensitivity.

Siddiqui and Heffner independently improved on the Collett design by reducing from five to three the number of distinct polarizers and waveplates [56, 91]. The logical diagram of their designs is illustrated in Fig. 10.1. There are four intensities that are measured by a quad detector: $I_{0,1,2,3}$. I_0 is the direct beam intensity. I_1 and I_2 are measured through linear polarizers oriented at 0° and 45° , respectively. I_3 is measured after the light transits a quarter-wave plate and polarizer. The orientation of the waveplate and polarizer, as illustrated, is such that the waveplate transforms right-hand circular polarization into the low-loss aperture of the polarizer. The input beam originates from a fiber and expands to one or more millimeters in diameter before collimation. Siddiqui uses a single lens, while Heffner uses a segmented concave mirror to separate and individually focus the four beams on their respective paths. Through separate adjustment of the mirror segments the intensity of each path can be balanced.

The four Stokes parameters are derived from the measured intensities according to

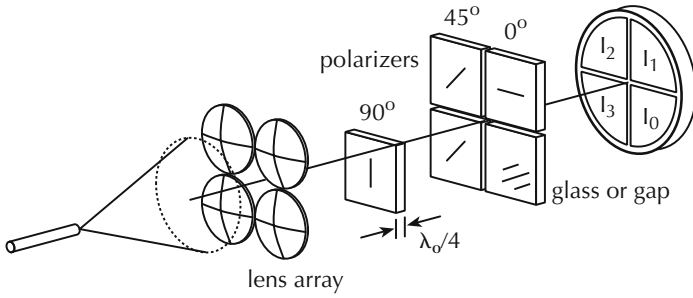


Fig. 10.1. Illustrative staring polarimeter for high-speed measurement of Stokes parameters. A four-quadrant detector measures the light transmitted along four parallel paths. The first path is all-pass while the next three condition the light by polarizing it along S_1 , S_2 , and S_3 . The bandwidth of the staring polarimeter depends only on the sensitivity of the photodetectors and back-end electronics. A very high-end speed could be in the GHz range, but most operate at MHz rates.

$$S_0 = I_0 \quad (10.1.1a)$$

$$S_k = 2I_k - I_0, \quad k = 1, 2, 3 \quad (10.1.1b)$$

The Mueller matrix for the path between the lens and detectors is

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} I_0 \\ I_1 \\ I_2 \\ I_3 \end{pmatrix} \quad (10.1.2)$$

The formal equivalence of these relations to the rigorous transformation from Jones to Stokes is verified using (1.4.14–1.4.17) on page 17, where $I_0 = I_x + I_y$. Moreover, to within a complex constant the associated Jones vector can be reconstructed as detailed in §1.4.1.

In practice, the Mueller matrix in (10.1.2) is only an idealization. While the matrix can always be constructed as shown up to the first two rows, the realistic form of the third row depends on the relative alignment of the 45° polarizer with respect to the one at 0° . Misalignment mixes in part of I_1 . The same holds true with the fourth row, but in addition the quarter-wave waveplate has a wavelength dependence. Away from center frequency the waveplate over or under rotates the polarization state, allowing a mixing with I_2 . The wavelength dependence requires a low- or zero-order true wave waveplate and a calibration table over wavelength. Also, the adiabatic expansion from fiber to collimator is sometimes replaced by a cascade of polarization beam splitters. The polarization-dependent loss of the PBS's imparts deleterious polarization dependence to the optical path which, ideally, can be calibrated out, but at higher cost and lower accuracy. Finally, the separate articulation of the segmented concave mirror in the Heffner design, or the four lenses as illustrated, provides for power balancing among the four paths during construction.

The staring polarimeter as illustrated is a bulk-optic component requiring several parts. An alternative in-fiber polarimeter is demonstrated by Westbrook *et al.* [30, 36, 102, 103]. His group imprints multiple blazed gratings into the fiber core to deflect light through the cladding into a detector array. While the printing cost may be higher than the cost of any particular part in the bulk-optic analogue, mass production may be inexpensive and the small form-factor and in-line nature have natural applications both in telecommunications as well as fiber sensors.

10.2 PDL Measurement

To assess the polarization-dependent loss of a component or transmission link, one must determine the minimum and maximum transmission through the device under test (DUT). A key difficulty is that the orientation of the PDL axis is unknown. A brute force approach is the so-called “all-states” method, which scans the input polarization over all states and looks for transmission extrema at the output. This method is obsolete because of the time required and the possible error between getting close to the extrema but not actually finding it precisely. Furthermore, the all-states method is prone to repeatability problems, which leads to each measurement being somewhat random.

Favin and Nyman pioneered the “four-states” method that determines the PDL precisely after measuring the transmission for only four input polarization states [35, 81, 82]. This method has been adopted through international standards bodies [65, 66, 94] and has been incorporated into most commercially available products. Moreover, Craig of the National Institute of Standards and Technology (NIST) has reported several details on the uncertainties present in the four-states method [12–14].

The four-states method has two parts. First there is a functional analysis of the transmission using the Mueller matrix. The result of this analysis yields expressions for the minimum and maximum transmission in terms of the Mueller-matrix entries alone. Second, estimates for the relevant matrix entries are determined by measurement of the DUT using only four input polarization states. This method works equally well for one wavelength as for a range of wavelengths. In the latter case, errors due to waveplate retardance change must be accounted for; such details can be found in the literature.

The functional analysis starts with the mapping of an arbitrary input Stokes vector to the output Stokes vector:

$$\begin{pmatrix} T_0 \\ T_1 \\ T_2 \\ T_3 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ - & - & - & - \\ - & - & - & - \\ - & - & - & - \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (10.2.1)$$

Since the transmission intensity T_0 is the only quantity of interest, the polarimetric values of T_1, T_2, T_3 are ignored, as are the respective entries in the

Mueller matrix. The ratio of transmitted to incident intensity is

$$T_0/S_0 = m_{11} + m_{12} \left(\frac{S_1}{S_0} \right) + m_{13} \left(\frac{S_2}{S_0} \right) + m_{14} \left(\frac{S_3}{S_0} \right) \quad (10.2.2)$$

The question is what values of $\vec{S}$ generate the minimum and maximum values of T_0 for fixed albeit unknown values of m_{1k} . The analytic problem can be solved using the Lagrange multiplier method [14]. The linear function f is to be maximized under the constraint g ,

$$\begin{aligned} f &= m_{11} + m_{12}s_1 + m_{13}s_2 + m_{14}s_3 \\ g &= s_1^2 + s_2^2 + s_3^2 - 1 = 0 \end{aligned}$$

where $s_k = S_k/S_0$. Under this particular constraint, the function f surely has extrema points. At any extremum, $df = 0$. Accordingly,

$$df = 0 = m_{12} \frac{\partial f}{\partial s_1} ds_1 + m_{13} \frac{\partial f}{\partial s_2} ds_2 + m_{14} \frac{\partial f}{\partial s_3} ds_3$$

The differential of g can also be taken, and the two expressions are added in linear superposition with the Lagrangian scale factor λ . Separation of like terms from the expression $dg + \lambda dg = 0$ yields the set of equations

$$\begin{aligned} \frac{\partial f}{\partial s_1} + \lambda \frac{\partial g}{\partial s_1} &= m_{12} + 2\lambda s_1 = 0 \\ \frac{\partial f}{\partial s_2} + \lambda \frac{\partial g}{\partial s_2} &= m_{13} + 2\lambda s_2 = 0 \\ \frac{\partial f}{\partial s_3} + \lambda \frac{\partial g}{\partial s_3} &= m_{14} + 2\lambda s_3 = 0 \end{aligned} \quad (10.2.3)$$

Extraction of s_k from each expression and substitution into g determines the value of the Lagrangian multiplier λ :

$$2\lambda = \pm \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2} \quad (10.2.4)$$

Substitution of (10.2.4) into (10.2.3) determines the extrema values for s_k in terms of the Mueller matrix entries. Substitution of the resultant s_k values into (10.2.2) yields the extreme values in transmission:

$$T_{\max} = m_{11} + \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2} \quad (10.2.5a)$$

$$T_{\min} = m_{11} - \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2} \quad (10.2.5b)$$

Therefore the entries in the first row of the Mueller matrix completely determines the minimum and maximum transmission ratios. Indeed the PDL is immediately given by [35]

$$\rho_{\text{dB}} = 10 \log_{10} \left(\frac{m_{11} + \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2}}{m_{11} - \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2}} \right) \quad (10.2.6)$$

Measurement is needed to formulate estimators for the values m_{1k} . The typical implementation of the four-states method is to probe the DUT with states $\{S_1, -S_1, S_2, S_3\}$, Fig. 10.2(a). There is nothing particular about this choice of states other than the obvious three orthogonal coordinates in Stokes space. However, it is clear that the best choice of states should provide the best estimators for the Mueller entries. The aforementioned states are generated via a polarizer and a removable or rotatable half-wave and quarter-wave waveplate. A first baseline experiment is performed with the DUT bypassed. Denote the transmitted intensities of the four experiments I_a, I_b, I_c, I_d . A second experiment is then performed with the DUT inline. The Mueller matrix of the device effects the input polarization states; the transmitted intensities, from (10.2.1), are

$$\begin{aligned} I_1 &= (m_{11} + m_{12})I_a, & I_2 &= (m_{11} - m_{12})I_b \\ I_3 &= (m_{11} + m_{13})I_c, & I_4 &= (m_{11} + m_{14})I_d \end{aligned}$$

These equations are rewritten in matrix form to relate the Mueller entries to the DUT output intensities:

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} = \begin{pmatrix} I_a & I_a & & \\ I_b & -I_b & & \\ I_c & & I_c & \\ I_d & & & I_d \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \end{pmatrix}$$

Inversion of the 4×4 matrix (which is not a Mueller matrix) yields expressions for m_{1k}

$$\begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \end{pmatrix} = \begin{pmatrix} (I_w + I_x) / 2 \\ (I_w - I_x) / 2 \\ I_y - (I_w + I_x) / 2 \\ I_z - (I_w + I_x) / 2 \end{pmatrix} \quad (10.2.7)$$

where $I_{w,x,y,z} = I_{1,2,3,4} / I_{a,b,c,d}$. Substitution of these estimates for the Mueller entries into (10.2.6) generates an estimate for the PDL of the DUT.

A critical aspect of the measurement is that the detector which detects all I has minimal PDL. Earlier detectors in fact exhibited PDL larger than 0.01 dB, where the requisite measurement accuracy was 0.001 dB. Nyman *et al.* were the first to add a depolarizer before the detector to eliminate the PDL from the test set [82]. In their work they added a 23 m length of unpumped erbium-doped fiber (EDF). The input light to the EDF is absorbed by the erbium ions and emitted at a longer wavelength. Repeated absorption and emission can completely depolarize the light. Although the conversion efficiency from polarized to depolarized light was 0.032%, use of this nonlinear diffuser was

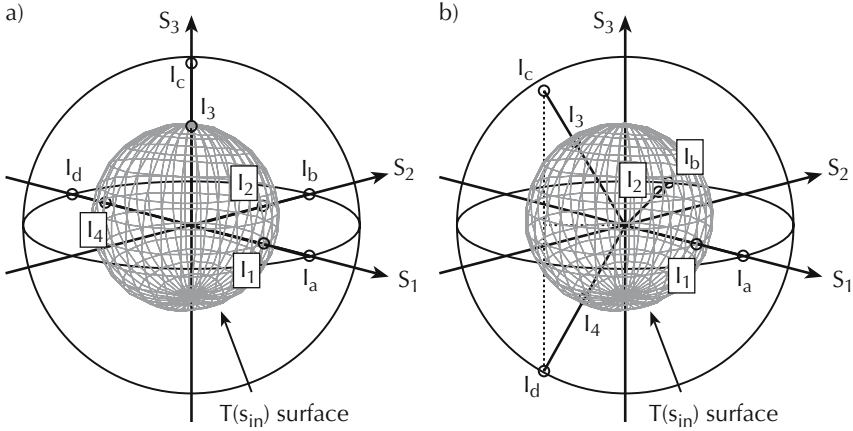


Fig. 10.2. Four-states measurement method for PDL. Calibration measurements are made without the DUT inline. Those measurement intensities are $I_{a,b,c,d}$. The DUT is subsequently spliced in and intensities $I_{1,2,3,4}$ are measured. a) Standard four-states $\{S_1, -S_1, S_2, S_3\}$ and T_p surface. b) Tetrahedral four-states with same T_p surface. The tetrahedral group has a 120° separate between all states, or maximum discrepancy.

the first to allow high-precision measurements. In measurement systems commercially available at the time of this writing, improved detectors or detectors preceded by a wedge depolarizer are used.

An alternative arrangement to the standard four-state method is proposed here. To achieve the best estimators for the Mueller entries, the probe polarization states should exhibit maximum discrepancy. That means that they should all be as far apart from one another as possible. For four points in a three-dimensional space this is a tetrahedral orientation, where each state is 120° away from the others (Fig. 10.2(b)). A tetrahedrally orientated set of four polarizations can be achieved with a polarizer and at most three waveplates. Repeating the preceding analysis for four such states, where two of the states lie on the equator, the Mueller entries are

$$\begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \end{pmatrix} = \begin{pmatrix} -I_x + I_y + I_z \\ I_w + I_x - I_y - I_z \\ (I_w + 5I_x - 3I_y - 3I_z) / \sqrt{3} \\ (I_y - I_z) / \sqrt{3} \end{pmatrix} \quad (10.2.8)$$

That the estimators for the Mueller entries all rely on more measurement information than the standard four-state method will reduce the overall error.

A means currently embraced by industry to improve the accuracy is to extend the four-state method to six states [12], where the probe states are $\{\pm S_1, \pm S_2, \pm S_3\}$. While application of the preceding analysis determines how the Mueller entries relate to the measured intensity ratios, note that the six-

state method, in addition to augmenting the measurements, uses maximum discrepancy between probe states. Six states of maximum discrepancy argue for increased sensitivity.

10.3 PMD Measurement

There are three principal PMD features that are of interest to measure, depending on application. One feature is the mean DGD $\langle\tau\rangle$ of a fiber or link. The preceding chapter detailed how $\langle\tau\rangle$ is the sole scaling parameter necessary to specify the statistics of all orders of PMD as well as its autocorrelation function. Another feature is the PMD vector as a function of frequency $\vec{\tau}(\omega)$. This vector information is necessary to characterize first- and higher-order PMD of a component or fiber directly, and is necessary to correlate receiver performance in the presence of PMD. The third feature is the direct measurement of fiber birefringence as a function of position. As birefringence is the origin of PMD, its characterization has led to important experimental information. For instance, measurement of fiber birefringence has validated the zero-chirality model of the birefringence for unspun fibers.

Table 10.1 classifies the demonstrated PMD measurement methods according to the principal parameter(s) they report. The wavelength scanning (WS) and interferometric (INT) methods are suitable to ascertain quickly the mean DGD of a fiber. These two methods are related via Fourier transform. The PMD vector as a function of wavelength can be measured using four related techniques, Jones Matrix Eigenanalysis (JME), Mueller Matrix Method (MMM), the Poincaré Sphere Analysis (PSA), and the Attractor-Precessor Method (APM); or a different technique here called the Vector Modulation Phase-Shift (V-MPS) method. Two basic differences between the first four vector methods and the latter are in the first instance a CW tunable laser and polarimeter is used while in the second instance an RF-modulated tunable laser and network analyzer are used instead. Finally, the local birefringence can be measured using polarization-dependent optical time-domain reflectometry (P-OTDR).

An alternative classification is adopted by Williams where measurement techniques are grouped according to the coherence time of the probe source in relation to the mean DGD of the device under test (DUT) [107]. Frequency-domain classification is for source coherence times τ_c much longer than mean DGD: $\tau_c \gg \langle\tau\rangle$. Time-domain classification is the opposite case, where $\tau_c \ll \langle\tau\rangle$. Frequency-domain techniques are the WS method, JME, MMM, PSA, and APM methods. Time-domain techniques are the INT and P-OTDR methods. The scalar and vector MPS methods are hybrids of the two, where the phase-shift measures the time-of-flight while the modulation is imparted on a high-coherence carrier that scans wavelength.

Tied in with most practical PMD measurements is the presence of PDL. As detailed in the preceding section, PDL can be measured, identified, and

Table 10.1. Classification of PMD Measurement Techniques

Abbev.	Method	Infer	Measurement	Num. Diff.	PDL Tol.
Mean DGD:					
WS ^(a)	Wavelength Scanning	$\langle \tau \rangle$	$T_p(\omega)$	no	yes
INT ^(b)	Interferometric	$\langle \tau \rangle$	$R(t - t_0)$	no	yes
Vector PMD:					
JME ^(c)	Jones Matrix Eigenanalysis	$\vec{\tau}(\omega), \langle \tau \rangle$	$S_{\text{out}}(\omega, S_{\text{in}})$	yes	yes
MMM ^(d)	Mueller Matrix Method	$\vec{\tau}(\omega), \langle \tau \rangle$	$S_{\text{out}}(\omega, S_{\text{in}})$	yes	no
PSA ^(e)	Poincaré Sphere Analysis	$\vec{\tau}(\omega), \langle \tau \rangle$	$S_{\text{out}}(\omega, S_{\text{in}})$	yes	no
APM ^(f)	Attractor-Precursor Method	$\vec{\tau}(\omega), \langle \tau \rangle, \eta$	$S_{\text{out}}(\omega, S_{\text{in}})$	yes	yes
S-MPS ^(g)	Scalar modulation phase-shift	$\tau(\omega), \langle \tau \rangle$	$\phi_{\text{RF}}(\omega, S_{\text{in}})$	no	yes
V-MPS ^(h)	Vector modulation phase-shift	$\vec{\tau}(\omega), \langle \tau \rangle$	$\phi_{\text{RF}}(\omega, S_{\text{in}})$	no	yes
Local birefringence:					
P-OTDR ⁽ⁱ⁾	Polarization Optical Time-Domain Reflectometry	$\beta(z)$		no	n/a
P-OFDR ^(j)	Polarization Optical Frequency-Domain Reflectometry	$\beta(z)$		no	n/a

^(a) Poole, Favin [86], ^(b) Gisin, Hefner [40, 52], ^(c) Hefner [47],
^(d) Jopson *et al.* [68], ^(e) Cyr [15], ^(f) Eyal *et al.* [34], ^(g) Williams [105],
^(h) Nelson *et al.* [80], ⁽ⁱ⁾ Galtarossa *et al.* [10], ^(j) Huttner *et al.* [61].

extracted in the presence of PMD. The PDL information can be used to extract the pure PMD effects (cf. §8.3.4). The three measurement methods adapted to this procedure are the JME, APM, and MPS methods. The JME method converts Stokes measurements into Jones transfer matrices, which are subsequently resolved into Hermitian and unity components. The unitary matrices are then analyzed for PMD and the Hermitian matrices for PDL. The MPS methods, both scalar and vector, use a four-states measurement at the input. The four-states method can combine PDL and PMD measurement into one overall system characterization. The PSA attempts to eliminate PDL effects before data analysis by driving the measured data into three orthogonal coordinates. The MMM method is the least equipped to handle PDL as it is recommended to measure only two polarization states (which can be fixed) and relies on an equation that is only approximate in the presence of PDL (which is fixed by the APM method). Finally, the WS method is reportedly tolerant to some degree of PDL; even though the measured data changes substantially with the addition of PDL, the number of mean crossings or extrema remains unchanged.

10.3.1 Mean DGD Measurement

Wavelength Scanning Method

Perhaps the simplest technique for mean-DGD measurement is the wavelength-scanning (WS) method developed by Poole and Favin [86]. The setup is illustrated in Fig. 10.3. The core of the measurement is the intensity response of a DUT, typically fiber, placed between crossed polarizers. The response can be measured either with a tunable laser swept through wavelength and detector, or with a broadband source resolved with an optical spectrum analyzer (OSA). In either case the transmission intensity, on a linear scale, is measured over frequency. The intensity is related to the analyzer polarization $\hat{p}$ and the SOP output from the fiber $\hat{s}(\omega)$ by

$$T_p(\omega) = \frac{1}{2} (1 + \hat{s}(\omega) \cdot \hat{p})$$

Since $\hat{s}(\omega)$ is uniformly distributed on the Poincaré sphere in the long-length regime one expects $\langle T_p(\omega) \rangle = 1/2$.

Once $T_p(\omega)$ is measured it can be analyzed either with a mean-level crossing analysis or a extrema-counting analysis. As indicated in the figure, mean-level crossing is the number of times the intensity crosses the mean-level of the spectrum. Extrema counting is also related to the mean DGD, but can be problematic in the presence of noise. Williams provides a decision algorithm to distinguish between noise and signal extrema [112]. The TIA standard for the WS method is available in [95].

Poole and Favin show that, in the long-length regime, $\langle \tau \rangle$ is related to N_m and N_e , the count of mean-level crossings and extrema, respectively, by

$$\langle \tau \rangle_m = 4 \frac{N_m}{B}, \quad \text{and} \quad \langle \tau \rangle_e \simeq 0.805\pi \frac{N_e}{B} \quad (10.3.1)$$

where B is the bandwidth in radial frequency: $B = \omega_2 - \omega_1$. The 0.805 coefficient is reported by Williams [112] and is a correction to the original Poole and Favin factor of 0.824. There are two more details to be addressed: the full measurement bandwidth and the measurement resolution. The uncertainty of mean DGD is inversely related to the measurement bandwidth and is governed according to the PMD autocorrelation function §9.4.3. Williams shows that the measurement resolution $\Delta\omega$ must meet or exceed $\langle \tau \rangle \Delta\omega \leq \pi/12$ in order to achieve the asymptotic crossing and extrema density necessary for a reliable measurement.

An interesting attribute of the WS method is that a measure of “long” or “short” regime is possible [86]. In the long-length regime the ratio of crossing to extrema is $N_m/N_e \sim 1.58$, while in the short-length regime the ratio is unity: $N_m/N_e = 1$. With a good measurement of the crossing and extrema count, the count ratio is a measure of the regime in which the DUT resides.

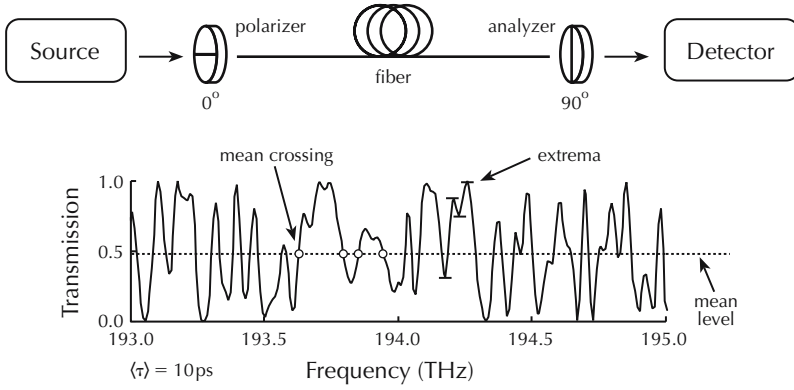


Fig. 10.3. Wavelength-scanning method to characterize $\langle\tau\rangle$ [86]. Either a broadband source and optical spectrum analyzer detector, or a tunable laser and broadband detector, are used to view the frequency-dependent intensity variation of the DUT between two crossed polarizers. The WS method associates the number of mean level crossing or intensity extrema to the mean DGD. The mean-DGD uncertainty is related to the measurement bandwidth.

Interferometric Method

The interferometric measurement method is closely related to the WS technique. The TIA standard for this method is available in [92]. An exemplar setup is illustrated in Fig. 10.4. As with the WS method, the DUT is placed between two crossed polarizers. The source is broadband and an interferometer is added in the optical path; as illustrated it is located at the output. The interferometer is polarization insensitive, the 50/50 beam splitter (BS) is a power splitter. Translation of one arm of the Michelson delays one path from the other to generate an interferogram at the detector, illustrated in the figure. The interferogram is recorded by the photocurrent I as a function of relative delay τ in the two arms. The variance of the interferogram σ_I is

$$\sigma_I^2 = \frac{\int \tau^2 I(\tau) d\tau}{\int I(\tau) d\tau} - \left(\frac{\int \tau I(\tau) d\tau}{\int I(\tau) d\tau} \right)^2 \quad (10.3.2)$$

Heffner relates the interferogram standard deviation to the mean DGD $\langle\tau\rangle$ according to the relation [52]

$$\langle\tau\rangle = \sqrt{\frac{2}{\pi}} \sigma_I \simeq 0.789 \sigma_I \quad (10.3.3)$$

Moreover, in the same paper he details the role of the source bandwidth. The 0.789 coefficient is the asymptotic limit for very wide source bandwidth. Lowering the source bandwidth first increases the coefficient value and then decreases it. Unlike the WS method, the result from an interferometric measurement depends centrally on the source characteristics. Finally, Heffner

points out an error in the Gisin paper [40] in that the interferometer above is an electric-field interferometer, not one that measures intensity. An intensity interferogram is studied in subsection “PMD Impulse Response” starting on page 358, where the second-moment of an intensity interferogram is equal to the RMS DGD value of the DUT (see Fig. 8.33).

Gisin and Heffner both show that the field interferogram and intensity spectrum generated by the WS method are Fourier transform pairs [40, 51–53]. Thus the analysis for the wavelength scanning method can be done by a moments calculation in the Fourier domain.

The central problem associated with the interferometric method is the separation or elimination of the source signature from the PMD-induced interferogram. Figure 10.5 illustrates the two demonstrated methods to make this separation. In the first method, independently proposed by Barlow, Gisin, and Cyr [1, 16, 17, 41], a known, fixed DGD element is concatenated with the DUT (Fig. 10.5(a)). The fixed DGD element, which can be a piece of polarization-maintaining fiber, serves to bias the DUT interferogram away from the zero-delay origin and the source-induced signal. In the second method, first proposed by Heffner [50] and later improved upon by Martin [78], looks to cancel the source signal within the interferometer directly. To do so, Martin adds a quarter-wave waveplate to one arm of the Michelson. Double-pass of the waveplate imparts a π phase shift in one arm with respect to the other. For every position of the translating arm of the interferometer the common delay τ_o is nominally cancelled, whereas the differential delays $\pm\tau/2$ due to PMD remain intact. The bandwidth of the quarter-wave plate plays a vital role in the source-cancellation method and must be considered.

10.3.2 PMD Vector Measurement

Direct measurement of the PMD vector gives high-resolution information on the state, frequency, and time evolution of PMD in a DUT. The average of the vector length, or DGD, over frequency can reproduce the mean DGD as measured by the WS or INT methods; but this is not the central purpose of the vector methods. Measurement of the PMD vector requires polarimetric stability of the DUT at least over the time it takes to launch the plurality of polarization states. Installed fiber plants often fluctuate quickly, making it difficult to use vector measurements.

Two sub-categories of measurement techniques are those that measure the output Stokes vectors as a function of frequency and input polarization state, and those that use a lock-in amplifier to measure output power as a function of frequency and input polarization state. In the first case analysis depends on the Stokes-vector change across frequency steps, while in the second case the vector is directly measured at each frequency because a narrowband modulation is imprinted on the probe signal.

Comparison across these varied techniques shows that the JME and APM methods carry the most advantages for the Stokes-based methods as effects

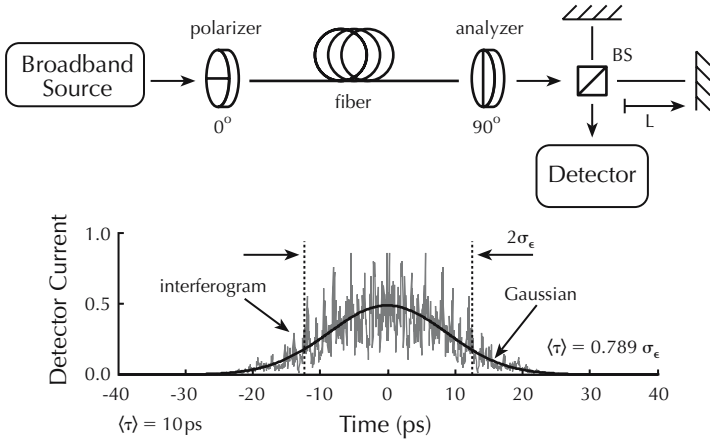


Fig. 10.4. Interferometric method to characterize $\langle\tau\rangle$. A broadband source is used to probe a DUT located between crossed polarizers, and the output is analyzed to produce an interferogram. The second moment of the interferogram is associated with the mean DGD of the DUT. The source bandwidth is directly entwined with the interferogram variance and must be accounted for.

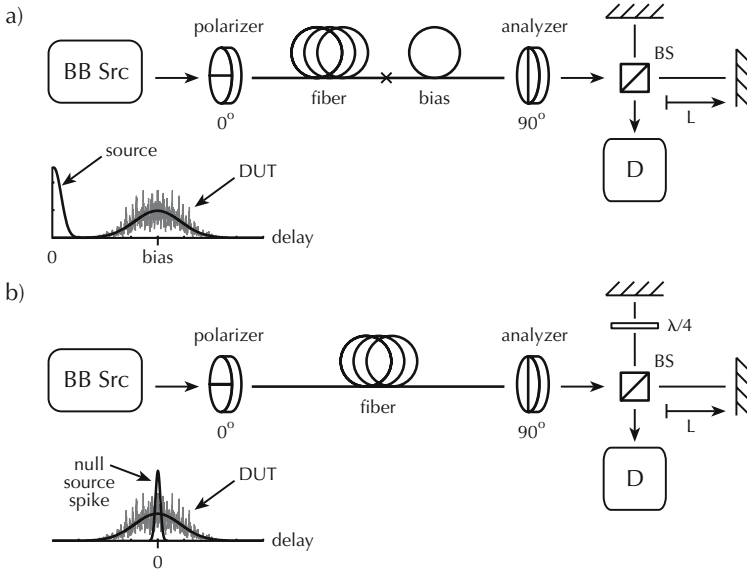


Fig. 10.5. The interferogram in the preceding figure is idealized: the source-induced peak at the origin was numerically removed. There are two ways to separate the source peak from the PMD-induced interferogram. a) A bias from a fixed, known DGD element is concatenated with the DUT [1, 41]. b) The source peak is cancelled by adding a quarter-wave waveplate to one arm of the interferometer [78]. Double-pass of the waveplate imparts a π phase shift in that arm for every position of the other arm.

of PDL can be stripped, and that the vector-MPS method is the most advantageous of all because it does not require frequency differencing and is largely immune to PDL.

Jones Matrix Eigenanalysis

Jones matrix eigenanalysis was developed in the early 1990's by Heffner [47, 55, 57]. Heffner discretized Poole and Wagner's PMD eigenvalue equation [85] to arrive at a solution using measurements of the Jones matrix. The measurement setup is illustrated in Fig. 10.6. A narrow-line tunable laser is the probe source. Wavelength accuracy is essential, so either the laser must have a built-in wavemeter or an external one must be added. At each frequency ω three polarizations are launched in sequence: P_a , P_b , and P_c . The light is transmitted through the DUT and is resolved by a polarimeter. The Stokes parameters $S_a(\omega)$, $S_b(\omega)$, $S_c(\omega)$ are in this way measured over frequency. Figure 10.6 illustrates the motion of $S_a(\omega)$, $S_b(\omega)$, $S_c(\omega)$ in Stokes space for three frequencies over a narrow band through an arbitrary DUT. Arcs are traced in frequency, a different arc for each input state. Over a wide frequency band an arc can have a complicated shape.

Once the Stokes vectors are measured the data is analyzed to determine the PMD vector. The Stokes vectors at each frequency are first converted to a Jones matrix at that frequency: $S_{a,b,c}(\omega) \rightarrow J(\omega)$. The conversion is detailed in §1.4.1 on page 17. Heffner's prescription at this point is to solve the PMD eigenvalue equation, but Karlsson and Shtengel introduce an intermediate step [64, 70]. In order to remove the effects of PDL on the data set, the Jones matrix is resolved into Hermitian and unitary components. The unitary component contains the PMD information and is fed into the remainder of the Heffner method. The details of this matrix decomposition are given in §8.3.4 on page 378. The decomposition converts the Jones matrices to unitary matrices: $J(\omega) \rightarrow U(\omega)$.

Recall from (8.2.10) on page 329 that the eigenvalue equation for PMD is

$$jU_\omega U^\dagger |p_\pm\rangle = \pm \tau/2 |p_\pm\rangle$$

The forward-difference equation analogue is

$$[U(\omega + \Delta\omega)U^\dagger(\omega) - (1 - j\tau_g\Delta\omega)] |p\rangle = 0 \quad (10.3.4)$$

Since $jU_\omega U^\dagger$ is traceless the eigenvalues of this equation are $\lambda_\pm = 1 \mp j\tau\Delta\omega/2$. Therefore the DGD at ω is

$$\tau(\omega) = j \frac{\lambda_+(\omega) - \lambda_-(\omega)}{\Delta\omega} \quad (10.3.5)$$

The PSP vectors are the eigenvectors $|p_\pm\rangle$ of (10.3.4). At the time of this writing, Shtengel has posted a LabView library to drive an Agilent 8509 instrument to measure the full PMD vector as a function of frequency [89].

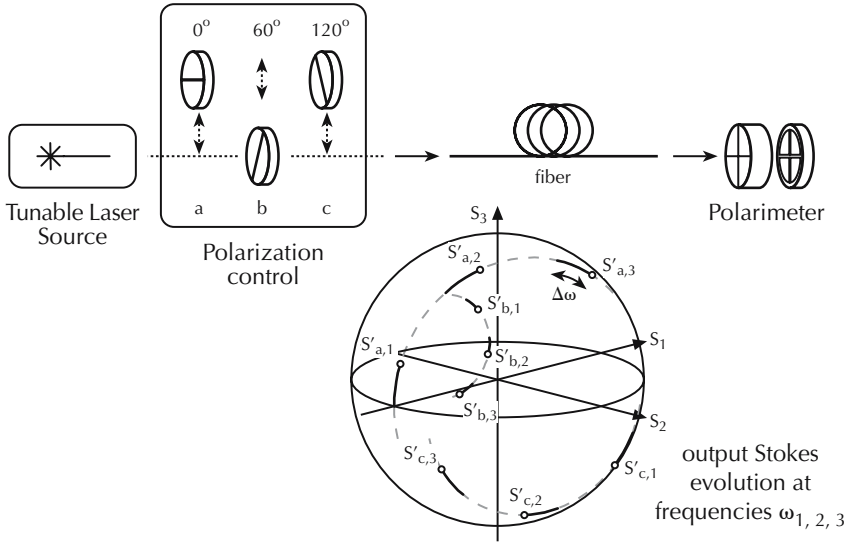


Fig. 10.6. Jones matrix eigenanalysis method to characterize $\vec{\tau}(\omega)$ [47]. Light from tunable laser with built-in wavemeter (or an external wavemeter) is serially polarized into three different states. The polarized light transits the DUT and is resolved by a polarimeter. At each frequency the Stokes vectors are measured for the three launch states; the Stokes vectors are then converted to a Jones matrix. Below shows a measurement fragment in Stokes space. Arcs are traced out on the sphere, one arc for each launch.

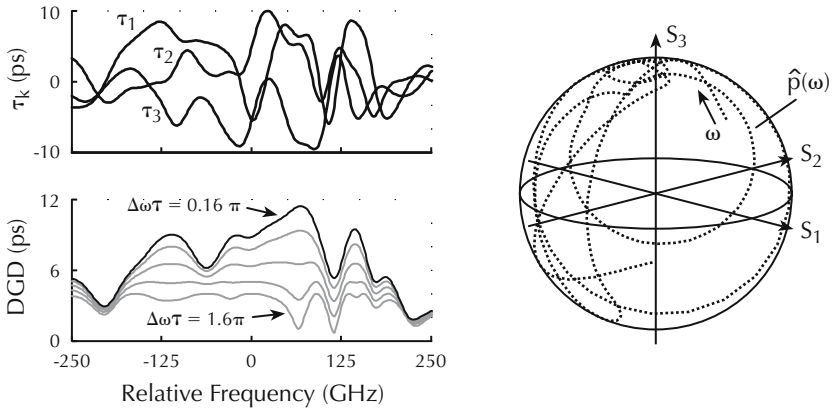


Fig. 10.7. Exemplar results of JME applied to a modelled fiber. The PMD vector is resolved into its cartesian components, plotted as PSP's in Stokes space, and plotted as DGD as a function of frequency. Data folding occurs if the frequency step size, local DGD product exceeds 180° .

Separately, Heffner reports validation this method in [48, 49, 54]. Williams gives a comparison between the WS, INT, and JME methods in [108].

Figure 10.7 illustrates a calculation of the JME measurement. Solution of the eigenvalues and vectors allows one to plot the three Stokes components of $\vec{\tau}(\omega)$ separately. The unit-vector $\hat{\tau}(\omega)$ maps the PSP spectrum of the DUT while the length $\tau(\omega)$ gives the DGD spectrum. Since full vector information of the PMD is available, second- and higher-order PMD can be estimated, although higher-order differences are required.

There are two practical issues regarding the JME method. First is that differences of eigenvalues and frequencies are used to calculate τ (10.3.5). Noisy data leads to noisy eigenvalues, which will upset the calculated values. Also, uncertainty of the true frequency difference $\Delta\omega$ will likewise lead to errors. Karlsson uses a multi-point estimator for the first derivative of the eigenvalues [70].

Second is the relation between the frequency step $\Delta\omega$ and the local DGD value. To first order, a frequency change generates precession of the output polarization about the PSP. Assuming a stationary PSP for the moment, the larger the frequency step the larger the precession angle. However, a step so large that $\tau\Delta\omega > 2\pi$ is ambiguous. Moreover, a step such that $\tau\Delta\omega > \pi$ is also ambiguous because the direction of the PMD vector cannot be determined uniquely (plus or minus). The step size is restricted to $\tau\Delta\omega \leq \pi$ to avoid ambiguity. For a fiber DUT, the step size is related to the mean fiber DGD via

$$\langle\tau\rangle\Delta\omega \leq \frac{\pi}{4} \quad (10.3.6)$$

to ensure almost no local DGD value is so great as to lead to a rotation greater than π . The effects of increasing step size $\Delta\omega$ on the data are illustrated in the DGD plot in Fig. 10.7. For $\tau\Delta\omega = 0.16\pi$ the calculated DGD values are close to the actual values. As the step size increases the values fall. At the location of the lower arrow, indicating $\tau\Delta\omega = 1.6\pi$, the curvature of the DGD spectrum actually inverts. This is called data folding [68] and leads to errant measurements.

Mueller Matrix Method

The Mueller Matrix Method was developed in that late 1990s by Jopson *et al.* [68, 69]. Contrary to the JME method, the measured Stokes vectors are analyzed directly rather than converted to equivalent Jones matrices. That the Stokes vectors are not converted to a Jones transfer matrix keeps the formalism concise but prevents the decomposition of the measured data into PMD and PDL components.

Consider a DUT with frequency-dependent transfer matrix $T(\omega)$. The output polarization state, as a Jones vector, is related to the input state as $|t\rangle = T(\omega)|s\rangle$. Under the assumption of zero PDL, $T(\omega) = U(\omega)$; the PMD vector is identified through $jU_\omega U^\dagger = (\vec{\tau} \cdot \vec{\sigma})/2$. The Stokes-space analogue for

this state transformation is $\hat{t} = R(\omega)\hat{s}$, where $R(\omega)$ is the rotation operator (2.6.22) on page 68:

$$R = \cos \omega \tau I + (1 - \cos \omega \tau)(\hat{r}\hat{r}\cdot) + \sin \omega \tau(\hat{r}\times)$$

where $\vec{\tau} = \tau\hat{r}$ and is a function of frequency. To identify the PMD vector in Stokes space, the derivative of the transformation is taken, $\hat{t}_\omega = R_\omega\hat{s}$, and the expression is rearranged in terms of the output polarization only: $\hat{t}_\omega = R_\omega R^\dagger \hat{t}$. With no PDL, the output state precesses about $\vec{\tau}$ according to $\hat{t}_\omega = \vec{\tau} \times \hat{t}$. The PMD vector is therefore $\vec{\tau} \times = R_\omega R^\dagger$. The finite-difference equivalent to $R_\omega R^\dagger$ is itself a rotation:

$$\begin{aligned} R_\Delta(\Delta\omega; \omega) &= R(\omega + \Delta\omega)R^\dagger(\omega) \\ &= \cos \Delta\omega \tau I + (1 - \cos \Delta\omega \tau)(\hat{r}\hat{r}\cdot) + \sin \Delta\omega \tau(\hat{r}\times) \end{aligned}$$

where both τ and $\hat{r}$ are evaluated at ω and it is assumed that $\Delta\omega$ is sufficiently small so that $\hat{r}_\omega = \hat{r}_{\omega+\Delta\omega}$ to first order. Combining the information that $\text{Tr}(\hat{r}\hat{r}\cdot) = 1$, that $\hat{r}\times$ is resolved according to (2.6.29) on page 70, and denoting $\Delta\varphi = \Delta\omega\tau$, the PMD vector can be reconstructed from the matrix elements R_{ij} of R_Δ :

$$\cos \Delta\varphi = \frac{1}{2} (\text{Tr}(R_\Delta) - 1) \quad (10.3.7)$$

and

$$\begin{aligned} r_1 \sin \Delta\varphi &= \frac{1}{2} (R_{23} - R_{32}) \\ r_2 \sin \Delta\varphi &= \frac{1}{2} (R_{31} - R_{13}) \\ r_3 \sin \Delta\varphi &= \frac{1}{2} (R_{12} - R_{21}) \end{aligned} \quad (10.3.8)$$

The DGD is calculated by

$$\tau(\omega) = \frac{\cos^{-1}((\text{Tr}(R_\Delta) - 1)/2)}{\Delta\omega} \quad (10.3.9)$$

and the PMD vector as $\vec{\tau}(\omega) = \tau(\omega)\hat{r}(\omega)$.

To implement the MMM method the rotation matrix $R(\omega)$ must be constructed. Consider the three launch states $s_a = (1, 0, 0)^T$, $s_b = (0, 1, 0)^T$, and $s_c = (0, 0, 1)^T$. Transmission of state s_a through R makes $t_a = R s_a$; the vector t_a is the first column of matrix R . The second and third columns of R are similarly determined. Thus a basic prescription for the measurement of R is given: measure the DUT with input launches as 0° , 45° , and 90° in physical angles. The MMM method can use the same experimental setup as the JME (Fig. 10.6) with different polarizers.

Jopson shows that only two independent polarization launches are necessary for the MMM method. The rotation matrix R has only three independent parameters: a precession angle, and azimuth and declination angle of the vector. A first launch alone determines two of the R parameters. A second launch

is the minimum necessary to determine the third parameter. Denote the measured result of two launches as $\hat{t}_1$ and $\hat{t}_2$, the constructed columns of R are

$$t_3 = \frac{\tilde{t}_1 \times \tilde{t}_2}{|\tilde{t}_1 \times \tilde{t}_2|}, \quad t_2 = t_3 \times \tilde{t}_1, \quad \text{and} \quad t_1 = \tilde{t}_1 \quad (10.3.10)$$

These constructed vectors form an orthonormal basis.

In comparison to the JME method, MMM offers only a simplified calculation. Errors from frequency differencing are of the same order (compare (10.3.5) and (10.3.9)) and the step size $\Delta\omega$ for MMM carries the same restriction as JME (10.3.6), see [39]. In the presence of PDL, JME and MMM differ significantly. JME can directly decompose the data into PMD and PDL components. MMM cannot and relies on an inexact equation of motion. As derived in §8.3.1, the equation of motion for the output polarization state in the presence of PMD and PDL is the non-rigid precession

$$\frac{d\hat{t}}{d\omega} = \vec{\Omega}_r \times \hat{t} - \vec{\Omega}_i \times \hat{t} \times \hat{t} \quad (10.3.11)$$

where $\vec{\Omega}_r$ and $\vec{\Omega}_i$ are the real and imaginary components of $jT_\omega T^\dagger$. There is no simple identification of $\vec{\Omega}_{r,i}$ to PMD and PDL, other than the limiting condition that $\vec{\Omega}_i \rightarrow 0$ and $\vec{\Omega}_r \rightarrow \vec{\tau}$ for zero PDL.

The cross-products of (10.3.10) attempt to “straighten out” the data, but measurements of only two states provides no ability to cross-check. Shtengel has shown that in the presence of PDL the two-state MMM gives spurious results as compared with JME [90]. Surely there is a boundary below which PDL does not affect the MMM measurements considerably; this boundary as a function of $\langle\tau\rangle$ has yet to be explored.

The Poincaré sphere analysis (PSA) method is equivalent to the MMM method and also relies on two or three launch states. The reader is referred to the work of Cyr of Exfo, Corp., for more information [15, 96].

The Attractor-Precessor Method

The Attractor-Precessor Method (APM) relies on the output-state equation of motion (10.3.11). This equation is called attractor-precessor because local PDL pulls the polarization state toward it and the local birefringence induces precession. APM, which lies between the JME and MMM methods, has been demonstrated by Eyal and Tur [33, 34] and is simplified here using spin-vector formalism. In [34], the authors write the Frigo equation of motion, (8.3.4) on page 373, for the output unit vector; the result is (10.3.11). The vectors $\vec{\Omega}_{r,i}$ have a total of six unknowns: length, and azimuth and declination angle for each vector. Eyal and Tur demonstrated the direct estimation of $\vec{\Omega}_{r,i}$ using (at least) three input states across two closely spaced frequencies. As each input state has two known quantities, three inputs are the minimum required.

Given estimates of $\vec{\Omega}_{r,i}$, Eyal and Tur use a stereographic mapping to Jones space from Stokes space to determine the PSP's, DGD, and DAS. As an alternative, the spin-vector nature of $\vec{\Omega}_{r,i}$ is used as follows. Define the complex vector $\vec{\Omega} = \vec{\Omega}_r + j\vec{\Omega}_i$. The eigenvectors of the traceless operator $\vec{\Omega} \cdot \vec{\sigma}$ are the PSP's of the system, according to (8.3.7) on page 374, and the corresponding eigenvalues are related to the DGD and DAS according to (8.3.8) on page 374. Such a procedure is detailed by Bao *et al.* [7].

The APM method has characteristics of the JME method because an eigenvalue equation in Jones space is solved, and has characteristics of the MMM method because the quantities first derived satisfy a Stokes-based equation of motion.

Modulation Phase-Shift Method

A supremely elegant PMD measurement method that dovetails directly with the four-states PDL measurement method is the modulation phase-shift technique (MPS). As with PDL, the first MPS method was “all-states”, where the input polarization was scanned in attempt to align with the PSP's at each frequency. Independently, Williams [105, 109] and Nelson *et al.* [42, 80] developed the “four-states” method similar to that for PDL. That is, by measuring four known launch states, the orientation of the PSP's can be directly calculated. The only difference between the methods of Williams and Nelson *et al.* is that the latter team reconstructs the full vector $\vec{\tau}(\omega)$, while Williams calculates $\tau(\omega)$. For this reason the methods are herein categorized as scalar MPS and vector MPS. The TIA standard for the S-MPS method is found in [93]. Williams and Kofler have extended the four-states method to six states, similar to Craig's six-state PDL measurement technique (cf. §10.2), to improve the accuracy to within a 40 fs single-measurement uncertainty [110].

A suitable measurement setup is illustrated in Fig. 10.8. The output from a narrow-line tunable laser is modulated sinusoidally at 1–2 GHz. The signal is then conditioned to lie along one of four polarization states. The modulated, polarized signal is transmitted through the DUT and detected by a network analyzer and polarization insensitive detector. The network analyzer determines the phase difference between the local oscillator, given by the modulator source, and the received optical field. The phase delay ϕ equals the product of modulation frequency ω_m and delay through the DUT τ_ϕ : $\phi = \omega_m \tau_\phi$.

Recall from the time-domain analysis of PMD in §8.2.6 that for a narrow-band signal the output group delay is due to the common and differential delays in the line ((8.2.47) on page 354):

$$\tau_g = \tau_o + (\tau/2) \hat{p} \cdot \hat{s}$$

where τ_o is due to the average group index, and $\hat{s}$ and $\hat{p}$ are the launch state and input PSP's, respectively. The measured group delay τ_g can lie anywhere between or at the extrema: $\tau_g = \tau_o \pm \tau/2$. The principal aspect of

the MPS method is to equate the measured delay τ_ϕ with the narrow-band group delay τ_g that comes from a moments analysis: $\tau_\phi = \tau_g$.

The optical signal launch into the DUT is split by the birefringence along the two input PSP's. The projected intensities are $I_\pm = I_o(1 \pm \hat{p} \cdot \hat{s})$. A phasor analysis of the received field, in the absence of appreciable differential-attenuation slope (DAS), sets the relationship between the principal variables:

$$\tan \omega_m (\tau_g - \tau_o) = \hat{p} \cdot \hat{s} \tan \left(\frac{\omega_m \tau}{2} \right) \quad (10.3.12)$$

The calibrated quantities are $\hat{s}$ and ω_m , the measured quantities are τ_g for each $\hat{s}$, and the unknown values are $\hat{p}$, τ_o , and τ . Under the constraint that $p_1^2 + p_2^2 + p_3^2 = 1$, there are a total of four unknowns. At least four measurements are necessary to solve (10.3.12).

Equation (10.3.12) can be solved in the following way. Since $\hat{p}$ is a three-entry column vector, the four input states are separated into a first launch and a group of three launches. Define a coordinate system $(\hat{r}_1, \hat{r}_2, \hat{r}_3)$ such that the first launch state $S_0 = \hat{r}_1$ and the remaining three launch states are $S_i = s_{i,1}\hat{r}_1 + s_{i,2}\hat{r}_2 + s_{i,3}\hat{r}_3$, $i = 1, 2, 3$. For each launch state there is a measured group delay $\tau_{g,i}$, $i = 0, 1, 2, 3$. The first launch condition and group of subsequent launches is written as

$$\tan \omega_m (\tau_{g,0} - \tau_o) = S_0^T \mathbf{p} \tan \left(\frac{\omega_m \tau}{2} \right) \quad (10.3.13a)$$

$$\tan \omega_m (\tau_{g,i} - \tau_o) = \mathbf{S} \mathbf{p} \tan \left(\frac{\omega_m \tau}{2} \right) \quad (10.3.13b)$$

where

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{-1} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$$

where $\mathbf{S}^{-1}$ is in anticipation of the following. In order for $\mathbf{S}^{-1}$ to exist, all three launch states cannot lie on the same plane. If two of the states are linearly polarized, the third must have a circular component.

Solving for $\mathbf{p}$ in (10.3.13b) and substitution into (10.3.13a) gives an equation which can be solved for τ_o :

$$\begin{aligned} \tan \omega_m (\tau_{g,0} - \tau_o) &= \alpha_1 \tan \omega_m (\tau_{g,1} - \tau_o) + \\ &\quad \beta_1 \tan \omega_m (\tau_{g,2} - \tau_o) + \gamma_1 \tan \omega_m (\tau_{g,3} - \tau_o) \end{aligned}$$

Linearization gives an initial solution:

$$\tau_o = \frac{\alpha_1 \tau_{g,1} + \beta_1 \tau_{g,2} + \gamma_1 \tau_{g,3} - \tau_{g,0}}{\alpha_1 + \beta_1 + \gamma_1 - 1} \quad (10.3.14)$$

Once $\tau_{g,0}$ is determined, (10.3.13b) can be solved for $\mathbf{p}$ and τ under the constraint that $\mathbf{p}^T \mathbf{p} = 1$. Nelson *et al.* report that linearization of transcendental equations (10.3.13) is valid to within 6% as long as $\omega_m \tau \leq \pi/4$.

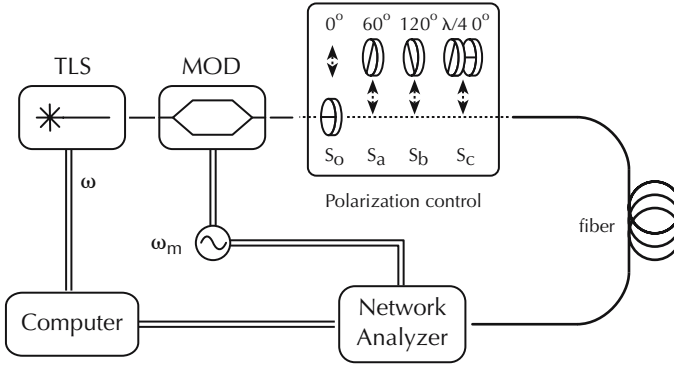


Fig. 10.8. Modulation phase-shift method to characterize $\vec{\tau}(\omega)$ [80, 105]. The line from a tunable laser source is modulated at $\omega_m \sim 1 - 2$ GHz. The field is then conditioned by one of four launch-state polarizers. At most three of the polarization states can lie in the same plane, at least one state must lie off the plane. For instance: $S_o = S_1$, $S_a = -S_1$, $S_b = S_2$, $S_c = S_3$. The signal is transmitted through the DUT and received by a network analyzer and polarization-independent detector. The analyzer measures the phase delay of the DUT path with respect to the modulated signal. Addition of a bypass around the DUT and direct intensity measurements augments the setup for PDL measurement.

Eyal *et al.* have combined PMD and PDL measurement into a single four-states MPS method [32]. Their prescription starts with the Mueller representation of the time-domain polarization transfer function (8.2.42) on page 343. Denoting the transfer function as $H(t)$, the Mueller matrix is constructed through $\mathbf{M}(t) = \frac{1}{2} \text{Tr} (H(t)\sigma_k H^\dagger(t)\sigma_i)$ for $i, k = 0, 1, 2, 3$. For an RF input frequency ω_m , the time-averaged Stokes-based transfer function is then

$$\mathbf{S}_{\text{out}} = \langle e^{j\omega_m t} \mathbf{M}(t) \rangle \mathbf{S}_{\text{in}} \quad (10.3.15)$$

By observing the amplitude and phase of the response, both PDL and PMD information can be extracted from the measurement. The advantage of this analysis is the implicit inclusion of the differential-attenuation slope (DAS). Expression (10.3.12) assumes a linear transfer function between input and output modulation amplitude. DAS, however, dilates or compresses the output modulation amplitude, distorting the transfer function. The DAS-induced amplitude change is accounted for in the Eyal analysis.

Finally, the beauty of the MPS technique is that PMD vector information can be extracted at each frequency. The frequency differencing necessary in the JME-type methods is replaced with narrow-band sinusoidal modulation. The modulation bandwidth is narrower than the step size one could achieve with JME and the phase detection of the modulated source gives a highly accurate measurement. The MPS technique is well suited for filter component testing in particular where transmission windows are substantially less than 100 GHz.

10.3.3 Polarization OTDR

Polarization-adapted optical time-domain reflectometry (P-OTDR) was pioneered in 1981 by Rogers [88] to investigate the local birefringence of communication fibers. That work was dedicated to the weak-coupling regime where $L \ll L_C$. In the mid-1990's Corsi, Galtarossa, and Palmieri extended the theory of P-OTDR into the strong-coupling regime [10, 11, 37]. Their experimental results of factory and installed fiber show that step-index, dispersion-shifted, and non-dispersion-shifted fiber all exhibit immeasurably low circular birefringence [39], validating the Wai and Menyuk stochastic model of fiber birefringence [99] for unspun fibers. They have also measured the PMD change of installed fiber over a period of several years [38].

One principal result of the Corsi *et al.* analysis is that the longitudinal evolution of the backward-travelling polarization state $\hat{s}_B(z)$ at location z obeys a precession rule about the round-trip birefringence vector $\vec{\beta}(z)$ such that

$$\frac{d\hat{s}_B}{dz} = \vec{\beta}(z) \times \hat{s}_B(z) \quad (10.3.16)$$

The practical significance of this precession rule is that it is formally equivalent to the PMD precession expression with $\omega \rightarrow z$ and $\vec{\tau} \rightarrow \vec{\beta}$. Therefore any of the PMD-vector measurement techniques reviewed above can be adapted to measure $\vec{\beta}(z)$. Figure 10.9 illustrates the experimental setup as adapted using a commercial OTDR. Recently, Gisin *et al.* have reported high-resolution measurements using a photon-counting technique adapted to P-OTDR [101].

A method complementary to P-OTDR is polarization-dependent optical frequency-domain reflectometry (P-OTFR). P-OTFR features a higher resolution than P-OTDR, but works over shorter distances. Huttner *et al.* has developed P-OTFR for birefringence measurements of fiber; their work is reported in [60–62]. The Huttner group has investigated single-mode fiber and, in particular, spun fiber having very low PMD coefficients. They have found that spun fiber exhibits a degree of circular birefringence [62].

An important study recently reported by Gisin *et al.* [71] uses the P-OFDR method to study the difference between phase and group index in various single-mode fibers. Recall that the birefringent beat length is defined as $\Lambda = \lambda_o/\Delta n$, where Δn is the refractive-index difference between the two eigenstates, while the DGD is defined as $\tau = \Delta n_g L/c$, where Δn_g is the group-index difference. Defining the “group beat length” as $\Lambda^* = \lambda/(c\tau/L)$, the ratio of beat length to group beat length Λ/Λ^* is a measure of the ratio between refractive and group birefringences. The authors report measurements at 1550 nm that show a ratio between 1.1 and 2.6, depending on fiber type. In particular, erbium-doped fiber exhibits a large ratio. So indeed any assumption that phase and group indices in optical fiber are the same is suspect.

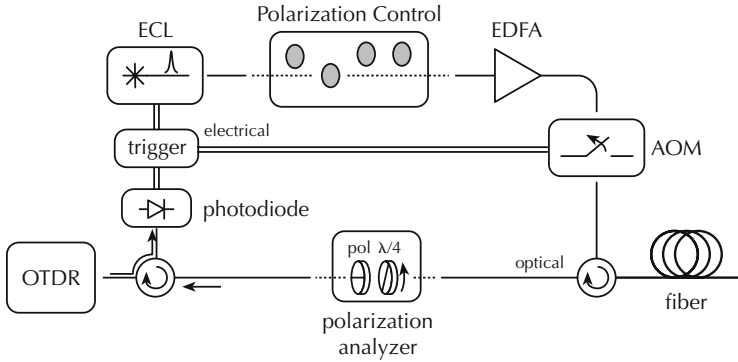


Fig. 10.9. Measurement setup for P-OTDR [39]. A commercial OTDR is adapted for polarization measurements. In particular, the instrument used by Galtarossa *et al.* did not have a pulse width as low as 5 ns. The OTDR output is detected and triggers an external-cavity short-pulse laser. The laser signal is polarized, amplified, and launched to the DUT. An acousto-optic modulator (AOM) transmits the pulse and block the ASE noise outside of the pulse time slot. The field is ultimately analyzed by a quarter-waveplate polarizer pair and returned to the OTDR for analysis.

10.4 Programmable PMD Sources

The polarization-mode dispersion present in a communications link must be accounted for when working out the link budget for a system. To satisfy the link budget at low cost the amount of PMD present needs to be low and the active components, especially transmitter/receiver (Tx/Rx) pairs, need to be tolerant to PMD. In order to validate system or Tx/Rx performance before deployment, PMD has to be generated and the system tested to demonstrate operation.

Figure 10.10 shows a testing hierarchy that optimizes the product-development cycle. In the development and validation phase of single-channel Tx/Rx pairs, a programmable PMD source is used to repeatedly address PMD states that cause trouble. A programmable source is also used to compare different products on an even basis. In the validation and deployment phase of loaded wavelength-division multiplexed systems, a PMD emulator is used to increase the confidence that the system will work over its lifetime.

The PMD emulator (PMDE) and PMD source (PMDS) complement one another. Hauer *et al.* report that a good PMD emulator should have three key properties: 1) the DGD should be Maxwellian-distributed over an ensemble of fiber realizations at any fixed optical frequency; 2) the emulator should produce accurate higher-order PMD statistics; and 3) when averaged over an ensemble of fiber realizations, the frequency autocorrelation function of the PMD emulator should tend toward zero outside a limited frequency range, in order to provide accurate PMD emulation for WDM channels [46].

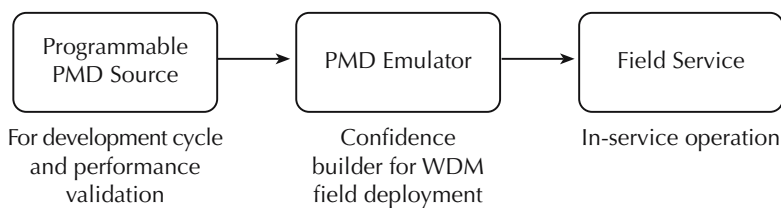


Fig. 10.10. Test hierarchy that optimizes the develop cycle for Tx/Rx pairs and for WDM system testing. A programmable PMD source targets difficult PMD states, allowing the developer to focus on the engineering issues. Product validation is also performed with the source. A PMD emulator is used to build confidence that a WDM system will work in specific environments and over lifetime. In-service fiber carries live traffic and demands PMD tolerance of the system.

A programmable PMD source, on the other hand, produces PMD in a predictable manner and typically spans a subset of all possible PMD space for a given mean DGD [28]. Additional attributes are the source's long-term stability, its repeatability, and its one-time calibration. When one uses a PMD source one accepts its limitation of PMD coordinates in exchange for predictability, repeatability, and stability. Moreover, the attribute of repeatability enables one to compare the performance of two or more different systems to the same PMD stress.

In principle a programmable PMD emulator can be built that meets the required traits for both categories above. Such an instrument, however, would be more expensive than two separate instruments. In the future, one might look for low-cost ways to combine features to create one super instrument without compromising performance.

A simple PMD emulator is a long spool of high-PMD fiber. Temperature cycling of the spool exercises a range of PMD states. One would, of course, like to have better control of the states and mean DGD. An improved PMDE couples many sections of polarization-maintaining (PM) fiber together. As few as three sections have been reported, but more typically 12, 15, or more sections are used. Mechanical rotators [63], thermo-optic heaters [45], and fiber squeezers [97] have been demonstrated. Another type is built using an integrated-optic platform and micro-ring resonators [75, 76]. Here light is divided with a polarization-beam splitter and each light component passes through a series of evanescently coupled ring resonators. The ring resonators impart DGD and co-directional couplers control the mode mixing. Still another type uses birefringent crystal and waveplates that rotate [8, 18]. Rotation of the waveplates changes the accumulated PMD through the crystals.

Since an emulator comprising 12–30 sections has far fewer correlation lengths than a typical transmission fiber, care must be used in determining the generated statistics. In the last chapter it was shown that the onset of the strong mode-coupling regime is for a length greater than 30 fiber-correlation lengths. Lima *et al.* and Biondini *et al.* have studied the difference between

“emulator” and fiber statistics and report that the tails of the emulator DGD distributions fall more quickly than fiber distributions [2, 74, 77], which leads to an under representation of high PMD states. To overcome these limitations, Yan *et al.* as well as Biondini and Kath have included importance sampling techniques to push the tails outward for correction [3, 113].

A new class of emulator is recently reported, the combined PMD and PDL emulator. Such an instrument is important to account for combined effects, especially as signal impairments can be worse than either effect in isolation. Waddy *et al.* and separately Bessa dos Santos *et al.* offer the first reports on such instruments [29, 98].

In the absence of PDL, there are three core problems when using a PMDE to develop and validate Tx/Rx-pair performance: in reference to the JPDF in Fig. 9.9 on page 406, the high-PMD states have low probability of occurrence – states as far out as $3 \langle \tau \rangle$ and $3 \langle \tau_\omega \rangle$ occur less than 0.01% of the time; emulators are not calibrated, so unless the PMD state is measured as it evolves there is no record of the states it went through; and emulators cannot reproduce the same test twice except in the statistical sense. For early development and validation applications, a programmable PMD source is necessary.

A programmable PMD source overcomes these PMDE limitations but at the expense of restriction to one- or few-channel use, and of restriction in addressable PMD space. The most basic of sources is the calibration artifact. P. Williams at the National Institutes of Standard and Technology (NIST) has developed a PMD standard for the strong mode-coupling regime [104]. The artifact is made from a stack of 35 thick quartz plates fiber pigtailed on either end. PMD measurement instrumentation can be calibrated to the artifact, setting a traceable standard. In fact, standards for PMD measurement methodologies are plentiful, but other than the Williams artifact no standards exist for PMD sources. This has impeded the industry regarding the development and commercialization of PMD compensators, both optical or electronic.

The programmable PMD source extends the stable, predictable, and repeatable nature of the artifact to a dynamic instrument. The earliest available programmable source is the JDS Uniphase “PMD emulator” [67]. This instrument, which generates only DGD, splits input light with a polarization-beam splitter and physically delays one path to the other through a Mach-Zehnder-like configuration. This instrument has been a successful product, but does not generate PMD in a meaningful way because second-order PMD is nonexistent. Gisin proposes a fix to this by looping back the light after one mode-mixing point [100]. Such an instrument generates DGD and the depolarization-component of SOPMD – these two components are the minimum necessary for product development. A drawback with both configurations is that the state-of-polarization is not stable due to the open environment of the delay line. In loop-back mode the instability will rotate the input PSP with time, which in turn changes the coupling of the signal to the generated PMD.

In order to build a stable PMD source, four physical attributes must be stabilized and controlled: the differential group delay for each stage, the birefringent phase per stage, the birefringent axes within a stage, and the polarization mode mixing between stages. In addition, to generate a clean spectrum, backreflection between components within the instrument and between the fiber-to-fiber collimation pair must be minimized [111]. Meeting all of these conditions at once has several practical ramifications that are detailed shortly.

A programmable PMD source is most useful when PMD states are specified as inputs to the instrument; given the input states the software calculates the required internal settings, e.g. mode-mixing angles, to produce the requested PMD. The alternative is to input physical parameters such as the waveplate angles and calculate the output PMD states. In fact there is a mapping between PMD states and physical states, the “forward” mapping from physical to PMD being straightforward to calculate and the “reverse” mapping from PMD to physical being, generally, multi-valued and difficult. The JDS Uniphase PMD emulator has a simple mapping between DGD and the length of the delay arm. The multistage PMD source demonstrated by Damask [28], when controlled in “wavelength-flat” mode, maps DGD and magnitude-SOPMD to physical rotation angles. This mapping is also simple. The ECHO source, a four-stage source also demonstrated by Damask [26], independently controls first- and second-order PMD and can select the balance between depolarization and PDCD components of SOPMD. As more stages are added beyond the four in ECHO, it is increasingly difficult to specify the PMD state with enough meaningful terms uniquely to reverse-map to physical parameters.

The following sections detail two successful instruments. The first instrument, simply called PMDS, does not control the birefringent phase within any section. The result is a severe limitation in the types of states that can be predictably addressed and an added complexity of the instrument. The second instrument, called ECHO for enhanced coherent higher-order [PMD source], explicitly controls the birefringent phase. This instrument is far simpler than the PMDS and generates a far broader range of predictable states.¹

10.4.1 Sources of DGD and Depolarization

The optical head of a twelve-stage programmable PMD source (PMDS) is illustrated in Fig. 10.11. The optical head is the heart of the instrument, while motion control boards, power supplies, and a chassis make it complete. This PMDS type has been built for both 10 Gb/s and 40 Gb/s applications, and was first built at Bell Laboratories, Lucent Technologies [18, 20], and subsequently by Chipman *et al.* [8]. The following sections detail how to build and operate the instrument.

¹ The author would like to redouble his acknowledgement of P. Myers, A. Boschi, R. Shelley, G. Simer, K. Rochford, and P. Marchese, without whose dedication the PMDS and ECHO sources would never have been realized.

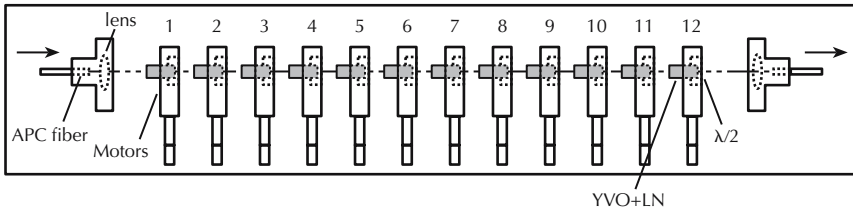


Fig. 10.11. Illustration of optical head of a twelve-stage programmable PMD source. Such a source does not have birefringent-phase control. The motorized rotary stages house fixed temperature-compensated high-birefringent crystals and rotatable true zero-order half-wave waveplates. Light is coupled in and out of the instrument via APC fibers, which are collimated with aspheric lenses having focal length $f = 5$ mm and beam diameter of 1.0 mm. Insertion loss and PDL are typically 1.8 dB and 0.1 dB, respectively.

Build and Calibration

The optical head is built with twelve independent rotary stages that house and hold temperature-compensated birefringent crystals for DGD generation and a true zero-order half-wave waveplate for mode mixing. The delay crystals are loaded into the rotary housing to minimize the optical path. All crystals and waveplates are anti-reflection (AR) coated to $R < 0.25\%$ at 1545 ± 30 nm. To reduce backreflection from the fibers and collimators, angle-polished (APC) fiber terminations and AR-coated lenses are used. The free-space optical path between collimators is ~ 30 mm and has a loss of ~ 2 dB using asphere lenses that expands the beam to 1.0 mm diameter. Once all the stages are added the insertion loss, PDL, and rotation-dependent loss (RDL) are typically 1.8 dB, 0.1 dB, and 0.2 dB, respectively.

Figure 10.12(a) illustrates the construction of each stage. Miniature, high-precision rotary stages, such as those from National Aperture [79], are used to house and hold the optics. These stages have a clear aperture 6 mm round and a top-plate that rotates. The stages are endlessly rotatable, have a repeatable resolution of 0.02° , a maximum spin rate of 4 revolutions per minute, and are driven by a miniature servo-motor. Onto each top plate, which is a separate ring that attaches to the rotary, a true zero-order half-wave waveplate is mounted. These waveplates are the polarization mode mixers. The waveplates are made from crystalline quartz with a thickness of $92 \mu\text{m}$. True zero-order waveplates, as opposed to compound zero-order plates, are used to minimize beam walk during rotation, called RDL. The waveplates are 8 mm rounds with a polished flat at the bottom aligned to the extraordinary axis of the crystal. The clear aperture of the top-plate rings is 3 mm, so there is 5 mm overlap between the waveplate and ring. This increases the resilience to mechanical shock. The waveplates are attached using a compliant UV epoxy that has minimal outgassing.

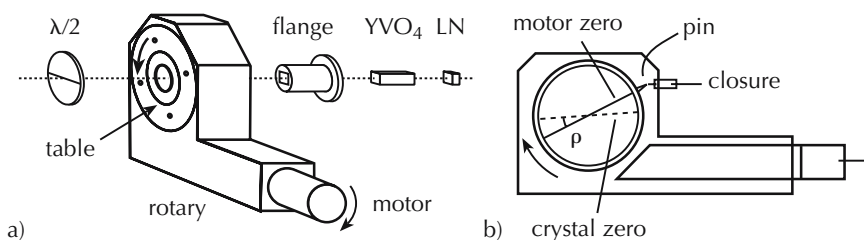


Fig. 10.12. Illustration of optics attached to the rotary stage and the absolute angular reference. a) A half-wave waveplate is mounted to the moving part of the rotary and the YVO_4 and LiNbO_3 crystals are fixed to a flange which is loaded into the body of the rotary. b) A pin and closure scheme is used to give an absolute angular reference. The calibration point of the stage is the angle between motor zero, where the pin closes the contact, and crystal zero, the orientation of the waveplate to maximize extinction on a calibration setup.

High-birefringent crystals that produce the DGD are inserted and fixed into the center bore of the rotary stage. Section §4.4 details the temperature dependence of the group index of several birefringent crystals. In particular, the combination of YVO_4 and LiNbO_3 gives a high group delay per unit length and low thermal dependence. Applicable crystal lengths are 14.801 mm of YVO_4 and 2.205 mm of LiNbO_3 per 10.0 ps of DGD (cf. Table 4.6). However, the variation of temperature coefficients from batch to batch likely exceeds the precision suggested here. For the 10 Gb/s instrument, 10.0 ps of delay is placed into each stage. The extraordinary axes of the YVO_4 and LiNbO_3 need to be aligned to compensate for temperature. The crystals are typically cut with a slightly rectangular cross-section, and the e -axis is aligned to one side.

The crystal pair is held by a custom flange that is cylindrical on the outside and rectangular on the inside. After UV epoxy is applied to the non-optical faces of the crystals, they are inserted into the flange and fixed by UV cure. To ensure that the crystals are flush, a fringe pattern at the interface between the crystals (part way into the flange) was checked. The crystals are specified to have a $\pm 0.5^\circ$ alignment of the crystalline e -axis to the physical aperture. Each crystal pair is accordingly aligned to within $\pm 1.0^\circ$. Typically, better alignment was observed.

Attachment of the crystal-loaded flange and waveplate to the rotary stage is the key part of the calibration process. The goal is to align the delay crystals across all twelve stages and to align the waveplate to each delay-crystal pair. Alignment for each stage is done one-by-one on a “calibration standard” setup [20]. The calibration standard has input and output fibers that are coupled by collimators. Two Polarcor polarizers (from Corning) are placed in the optical path in rotary stages and crossed. Using a power meter the polarizers are crossed so that the extinction ratio exceeds 60 dB. The polarizers are then permanently fastened into place.

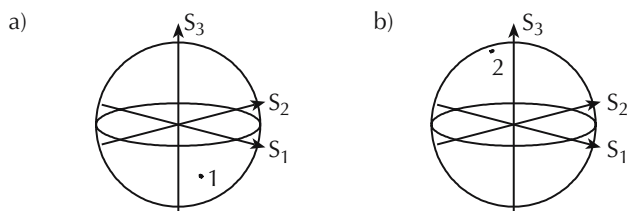


Fig. 10.13. Measured output of a PMD source over two 6 hr periods demonstrates temperature stability. a) Day time with laboratory traffic. b) Overnight.

To have a repeatable calibration point, an absolute angular reference on the rotary stage is required. For the rotary stages used here, the absolute angular reference is a mechanical closure, fixed to the housing, that is actuated by a pin, fixed to the rotary wheel, when the pin physically brushes the closure; see Fig. 10.12(b). The pin brushes the closure for about 2° of travel but first closes it with an angular precision of $\sim 0.05^\circ$. This first closure point is called “motor zero,” and is always detected by slow rotation in the same direction.

Once a rotary stage is set to motor zero, the waveplate ring is attached. The relative orientation of the waveplate e -axis to motor zero is unknown, although the polished flat gives some indication. The stage is then placed on the calibration standard and the waveplate is rotated to maximize optical extinction. Typical quartz waveplates achieve better than 50 dB extinction. The angular orientation for maximum extinction is called “crystal zero”. The difference in angle between motor and crystal zero is the calibration point for the stage. This calibration point is recorded in the instrument software. Crystal zero is found for any unknown orientation of the rotary by first rotating to motor zero, then rotating by the calibration angle to crystal zero.

Once crystal zero is found, the flange is loaded into the body of the rotary. The flange is manually rotated until maximum extinction is found and is then fixed in place. Typical extinction ratios at this point are 42 – 45 dB, but as low as 34 dB was found on occasion. Once the flange is fixed in place the rotary assembly is complete.

The optical head of the instrument is assembled using twelve rotaries all set to crystal zero. One-by-one each rotary is set in position against set pins and screwed into place. Minor adjustments are made to minimize the insertion loss, which is monitored throughout, since misalignment of the crystals will walk the beam away from the input aperture of the second fiber. Once all motors are in place, final adjustment is made to the collimators to minimize the loss and these are then locked into place. A dust cover protects the optics.

A final couple of points. There is a factor of four between the physical angle of a half-wave waveplate and its Stokes angle. One factor of two comes from the mirror-image about the e -axis of the waveplate, giving an apparent rotation of $2\times$, and the other factor comes from conversion to Stokes space from physical space. Separately, the polarization stability of this instrument

is shown in Fig. 10.13. (A good reference for the polarimetric stability of other sources is given in [114]). The temperature dependence of YVO_4 or LiNbO_3 alone is large, but the crystals as a pair greatly stabilize the birefringent phase.

Operation

Because the birefringent phase of each stage is not known and is not controlled, the class of sources called PMDS cannot predictably generate PMD that has more than one Fourier component. That is, only “wavelength-flat” states are predictably generated. A predictable, frequency-dependent DGD spectra requires phase control of the Fourier components, but this phase control is absent in the PMDS. Even though non-wavelength-flat states are not fully predictable, they can be repeated due to the instrument’s stability.

Wavelength-flat states produce DGD and pure depolarization; no PDCD or higher-order PMD is generated. For basic tests this actually has several advantages. The first is that no frequency alignment is necessary between the PMD generated by the instrument and the laser line of the transmission – the DGD and magnitude-SOPMD are constant in frequency. Second, depolarization statistically dominates PDCD so it is the more common component of SOPMD. Experimental evidence shows that depolarization also dominates the impairment of a signal in many instances. Third, the generated PMD is “engineering pessimistic” in that it is unlikely a fiber will exhibit high DGD and magnitude-SOPMD over the entire bandwidth of the signal. When a Tx/Rx pair can tolerate the PMD generated by the PMDS it will generally have an easier time of it on a live line.

Figure 10.14 shows the properties of wavelength-flat states. These states are generated by two PMD vectors $\vec{\tau}_{1,2}$. The first vector precesses about the tip of the second vector as a function of frequency (Fig. 10.14(1)). The Stokes angle $4\theta_{21}$ between the vectors is four times the physical angle θ_{21} of an intermediate half-wave waveplate. This angle is fixed in frequency. The output PMD vector is the vector sum of the components. The length is constant in frequency while the pointing direction traces a circle in Stokes space. The DGD and SOPMD can easily be determined geometrically: the DGD is the vector length following the triangle rule, and the magnitude-SOPMD is the tangential rate at the tip of $\vec{\tau}_1$ with frequency. The tangential rate is clearly $\Delta s = r\Delta\theta$, where $r = \tau_1 \sin 4\theta_{21}$ and $\Delta\theta = \tau_2\Delta\omega$. Putting this together, the DGD and magnitude-SOPMD are

$$\tau^2 = \tau_2^2 + \tau_1^2 + 2\tau_1\tau_2 \cos(4\theta_{21}) \quad (10.4.1a)$$

$$\tau_\omega = \tau_2\tau_1 \sin(4\theta_{21}) \quad (10.4.1b)$$

For fixed $\vec{\tau}_{1,2}$ the DGD and magnitude-SOPMD are parametric in θ_{21} . Patscher and Eckhardt investigated a two-stage optical compensator and demonstrated similar results [83].

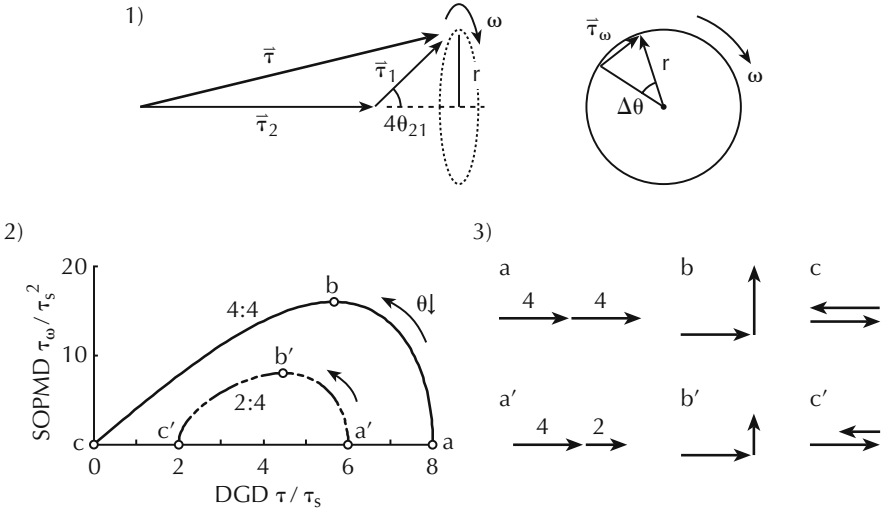


Fig. 10.14. Representations of two-section PMD states. 1) Two concatenated PMD vectors; the first vector precesses about the axis of the second with frequency. The angle between vectors is four-times the physical angle of the intermediate waveplate, and is fixed with frequency. The vector length $\vec{\tau}$ is the output PMD vector; its length is constant in frequency and its pointing direction traces a circle in Stokes space. 2) State-space in first- and second-order PMD for 4 : 4 and 2 : 4 groupings. 3) Vector representations of each corresponding state.

For each angle θ_{21} a state (τ, τ_ω) is produced. The locus of states for all angles traces a trajectory in first- and second-order-PMD state space. For example, when the component vectors are both 4 in length, the trajectory labelled 4 : 4 is traced (Fig. 10.14(2)). PMD states along a trajectory are continuous, and the maximum and minimum DGD are 8 and 0, respectively, and the maximum SOPMD is 16. Alternatively, when the vector lengths are 4 and 2, the 2 : 4 trajectory is traced. In this case the minimum DGD is not zero but $4 - 2 = 2$. The vector diagrams for various states are illustrated in Fig. 10.14(3). Finally, the state-space scales by τ_s , the delay per stage. For the PMDS described above, $\tau_s = 10.0$ ps.

The PMDS instrument makes wavelength-flat states by aligning the stages into two groups (Fig. 10.15). In this figure only eight stages are illustrated, so there are only ten unique trajectories. An important aspect is that pairs of stages can be cancelled optically by rotating the intermediate waveplate by 45° . This flips the fast and slow axes from one stage to the next. As illustrated in Fig. 10.15(b), a 4 : 2 trajectory (the same as 2 : 4) is made by allowing DGD to accrue through four consecutive stages and then mode-mixing at the junction to the fifth stage. The waveplate labelled 5 is not rotated so that DGD accrues between stages five and six. Finally, the waveplate labelled 6 is rotated by an equal and opposite amount as waveplate 4 to restore the polari-

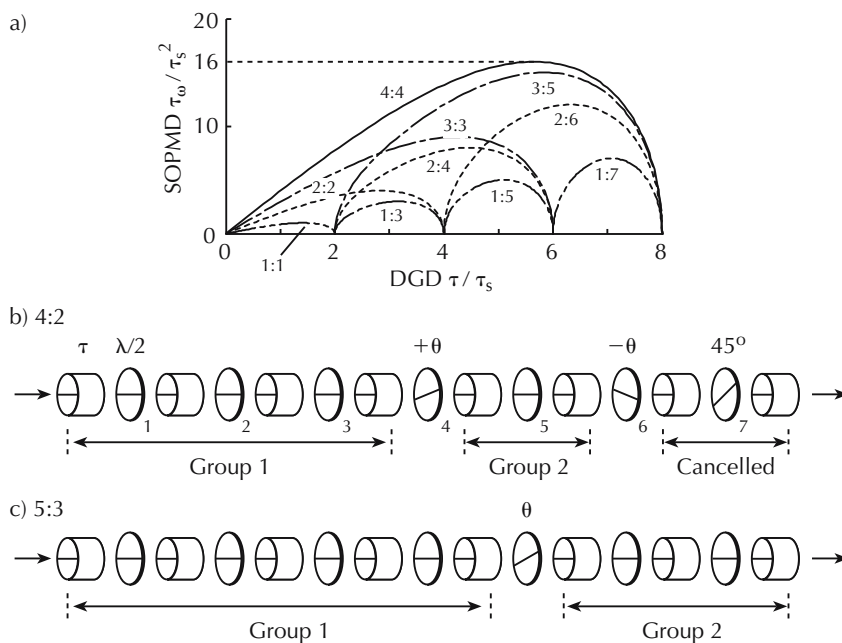


Fig. 10.15. Correspondence between PMD state-space and physical realization for an 8-stage cascade. Groups are formed by setting the intermediate waveplates to zero angle. Mode mixing happens whenever a waveplate has a non-zero angle. Pairs of stages can be optically cancelled by setting the intermediate waveplate to 45° .

metric axis. In a similar manner, a 5 : 3 group is shown. Another important feature of the PMDS is that it has a true zero PMD state. Without it, the instrument would have to be bypassed during system setup.

Experimental validation of a 10 Gb/s source is shown in Fig. 10.16(a,b) [28]. The figures show six measured spectra of two 6-stage groups. The six states of the PMDS were measured using Heffner's JME method and the Stokes data was used to calculate DGD and PSP values. The substantially wavelength-flat DGD spectrum labelled A corresponds to no mode mixing and maximum DGD. Accordingly, output PSP spectrum A points essentially in one direction. The DGD spectra B, C, D, E, and F have corresponding PSP spectra which are circles of ever increasing radius (PSP spectrum E removed for clarity). When the DGD value is zero, the corresponding PSP spectrum will trace a great circle through the $\pm S_3$ poles.

Figure 10.17 shows an overlay of 47 wavelength-flat states plotted in (τ, τ_{ω}) space. The dashed lines are theoretical trajectories derived from (10.4.1). The points are the measured first- and second-order PMD values averaged over a free-spectral range. The points in fact display the results from five repeated tests performed overnight; the tight grouping illustrates the stability of the instrument.

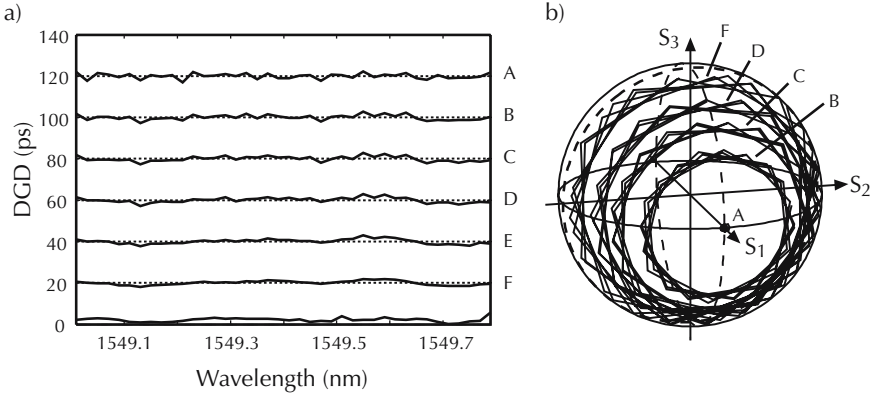


Fig. 10.16. Measured DGD and PSP spectrum from two-group operation of a 10 Gb/s PMDS. a) Seven measured DGD spectra over a free-spectral range. The spectra are generated with two 60 ps groups, where the intermediate waveplate controls the mode mixing. These spectra are “wavelength-flat,” indicating only DGD and depolarization are present. b) Six measured output PSP spectra over a free-spectral range. Letters A, B, C, D, and F correspond to respective DGD spectra. That wavelength-flat states generate pure depolarization is evidenced by the circular PSP spectra.

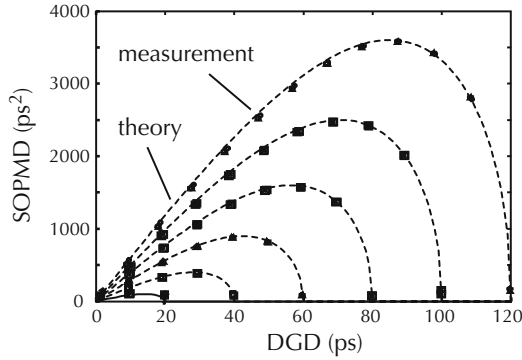


Fig. 10.17. Comparison between experiment and theory for 47 wavelength-flat PMD states. Dashed lines are theory; boxes, experiment. The experiment was repeated five times in succession overnight, so the overlap of boxes on the same state indicates the stability of the instrument.

Taken together, Figs. 10.16 and 10.17 demonstrate a central aspect of the two-group PMDS operation: the resultant PMD spectra are “pure,” with negligible wavelength dependence and pure depolarization with no PDCD. Moreover, the accuracy, stability, and repeatability evident in Fig. 10.17 allows for comparison of one system to another.

Total Delay of 1.2T

The total delay built into a PMDS instrument depends on the application. As a validation tool for Tx/Rx performance, the bit-error rate (BER) should be mapped over an entire bit time T , where $T = 100$ ps at 10 Gb/s and 25 ps at 40 Gb/s. This mapping should accurately represent both first- and second-order PMD states based on the JPDF for fiber. An increase from 1.0T to 1.2T increases the maximum SOPMD by 40%, which gives improved coverage for a JPDF scaled to a fiber with mean-PMD of 30 ps at 10 Gb/s and 7.5 ps at 40 Gb/s.

Variations

One variation is to operate the PMDS in PMD emulation mode. In this mode any and all stages are engaged so that a large amount of mode mixing is introduced. Since the rotary stages are dynamic and endlessly rotatable, rotation speeds that correspond to prime-number multiples of a unit speed drive the instrument through a virtually endless number of states. Moreover, since the instrument is calibrated, a specific path in time can be reproduced. Calculation shows that the average DGD for the 10 Gb/s instrument over all states is $\langle \tau \rangle \simeq 31.5$ ps, although the distribution tails fall faster than Maxwellian.

Another variation uses binary-weighted delay stages similar to that demonstrated by Yan *et al.* [114]. Such an instrument fills the first- and second-order PMD state space with more trajectories, giving it better coverage. One realization is an instrument with fifteen stages, the first eleven stages are as before, the next two are loaded with two $\tau_s/2$ -length crystals, and the last two loaded with two $\tau_s/4$ -length crystals. In this case, over 120 distinct trajectories are available and cover the state space well. The problems are the size and cost of the instrument, its fragility due to the short-length stages, and its difficulty to program. The two pair of binary-weighted stages divide all possible trajectories into four categories depending on their alignment or cancellation, making the instrument cumbersome to calibrate and operate.

Problems with the PMDS

The PMDS instrument was the first to demonstrate stability, predictability, and repeatability. However, problems remain, problems that ultimately call for the ECHO instrument.

Optically, the state-space coverage is poor. The 21 trajectories offer continuous PMD tuning along them, but jumping from one trajectory to another requires several motors to move at once, unless the instrument is first reset to zero, which is time consuming. It would seem unlikely to happen, but it has occurred that the instrument passes through high PMD states going between two low states, which in turn can disrupt an experiment. Even beyond this annoyance, many regions are simply not accessible, and the state density falls

off for higher PMD values. But high PMD values are precisely where the state density should be highest. Moreover, the wedge delineated by zero DGD, finite SOPMD on one side and the 6 : 6 trajectory before its peak on the other side is an entire range of relevant PMD states that are inaccessible by the instrument. These states represent high SOPMD for low DGD, which has significant probability of occurrence, as indicated on the JPDF in Fig. 9.9. Finally, the birefringent phase is not controlled at each stage, limiting the predictability of the instrument to wavelength-flat states.

Mechanically, the optical head is fragile. The crystals in the motor housings are not resilient to excessive mechanical shock or temperature variation. The rotary stages are very high quality, but motor burnout or motor-zero problems do occur. The more rotaries within any one instrument, the higher the likelihood of an instrument failure. Finally, use of twelve motors is expensive and makes for a long build and calibration time.

10.4.2 ECHO Sources

The Enhanced Coherent Higher-Order (ECHO) PMD source was developed in response to the shortcomings of the twelve-stage sources. In addition to the stabilization and control of the differential-group delay, birefringent axes within a stage, and mode-mixing between stages, ECHO calibrates and controls the birefringent phase of each stage [24]. This has several optical ramifications: higher-order PMD spectra are predictably generated, the spectra is continuously tunable in frequency without changes in shape, the first- and second-order state-space is continuous, and the state-space faithfully covers the JPDF; and several mechanical ramifications: only five rotaries are needed and all the delay crystals are mechanically aligned to a single reference.

ECHO looks very much like a birefringent filter, which has been studied by Lyot, Solc, Evans [31], and Harris [44]. But the ECHO “birefringent filter” imposes structure on the PMD spectra, not the intensity spectra. The birefringent filter was extended by Buhrer [4] to have continuous frequency tuning of the intensity spectrum by adding Evans phase shifters (cf. §4.6.3), and likewise, ECHO adopts the phase shifter to continuously tune the PMD spectrum. Tuning of the PMD spectrum at the source means that the transmission laser can be fixed in frequency, which is the preferable way to setup an experiment.

Finally, ECHO highlights the fact that the shape of PMD spectra is not determined by mode mixing alone but also by the birefringent phase. This is a key point. Structured PMD spectra generated by an unstable source, such as PM fiber, cannot be predicted *even if* the mode mixing is completely controlled. The absence of birefringent-phase stability causes the spectrum to change shape anyway. ECHO demonstrates this effect.

Opto-Mechanical Layout

The opto-mechanical layout of the ECHO optical head is fundamentally different than for the PMDS (Fig. 10.18). Regarding the optics, there are only four delay stages, rather than twelve, and three intermediate mode mixers. Like the PMDS, the delays are made from temperature-compensated YVO_4 - LiNbO_3 crystal sets. The mode mixers are true zero-order half-wave waveplates. Added to the second and third stage are Evans phase shifters. Each phase shifter has a pair of fixed quarter-wave waveplates and a rotatable half-wave waveplate mounted on a rotary stage. For stages two and three, the total birefringent phase is that from the delay crystals plus the phase imparted by the phase shifter. As shown in §8.2.7, the birefringent phases of the first and last stage do not affect the PMD spectrum, so phase shifters are not used in the two outer stages in ECHO. There is, however, a polarimetric difference whether or not end phase shifters are included, but this is immaterial.

The mechanical layout of the optical head puts all delay crystals and quarter-wave waveplates onto one solid platform (Fig. 10.18(b)). The platform has sections removed to make room for the rotary stages. Onto the platform a “crystal guide” is attached. The crystal guide gives an edge along which all crystals and waveplates are abutted. In this way the fixed optics have a single bottom and side mechanical reference. This mechanical structure makes placement of the fixed optics easy, and optical properties such as extinction ratio and phase are repeatable. To either end of the platform the collimator assemblies are attached and aligned. The rotaries are placed in the platform gaps and fixed from the bottom. The only optics attached to the rotaries are half-wave waveplates.

Like the PMDS, the delay crystals are cut with a rectangular aperture, with the e -axis aligned to one side. When aligned to the crystal guide these e -axes are horizontal. The aperture of the quarter-wave waveplates is the same and the e -axis is inclined by the requisite 45° . A true zero-order quarter-wave waveplate made from crystalline quartz is $46\text{ }\mu\text{m}$ thick at $1.55\text{ }\mu\text{m}$. In order to handle the part and fix it to the stage, the waveplate is best mounted to a host, such as BK7. To minimize internal reflection at the waveplate/host interface, the glasses should be optically contacted and anneal-bounded.

The extraordinary axes of all waveplates in the ECHO instrument have to be aligned and not crossed, an unnecessary requirement for the PMDS. Since the e -axes of the quarter-wave waveplates are at 45° , placement of the part onto the platform backwards flips the relative orientation of the plate. This, in turn, causes unwanted mirror images in the control of the phase shifter and mode mixers, as is obvious after study of Fig. 4.19 on page 185. There are at least two ways to ensure proper orientation: visual inspection of the plates through crossed polarizers on a light table; or by applying to either side two optically equivalent AR coatings having distinct colors, as proposed by Shirai for iron garnets (see page 152).

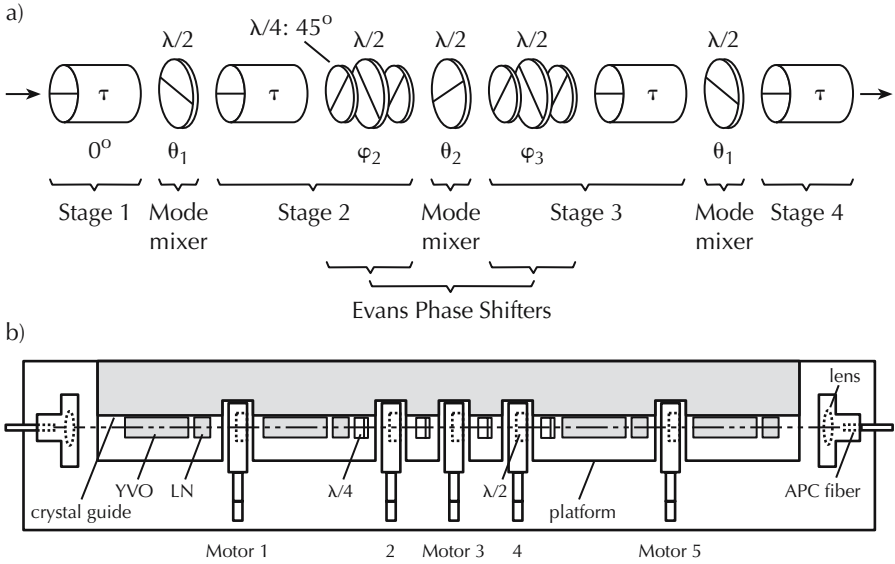


Fig. 10.18. Illustration of opto-mechanical layout of ECHO source. Four equal-length crystal delay stages are mode-mixed with three true zero-order half-wave waveplates. Two Evans phase shifters are added to the center stages to control the birefringent phase. The five rotatable waveplates fall into two groups: three waveplates for mode-mixing, two for phase control.

The ECHO instrument is calibrated as it is built [22]. The first step is to zero-out the residual birefringent phase of the center two delay sections. This is done at a “calibration frequency.” At a fixed frequency, a delay section has an integral number of birefringent beats and a fraction of a beat. This fractional part is the residual birefringent phase. For instance, a 10.0 ps delay at 194.1 THz has 1941 birefringent beats. But since the $\pm 3 \mu\text{m}$ manufacturing tolerance of the crystal length is almost the same of the $7 \mu\text{m}$ birefringent-beat length in YVO_4 , the residual birefringent phase is random. At the calibration frequency the Evans phase shifters are adjusted to drive the residual phase to zero. Once this is done, the center mode mixer is added and aligned to the optical axis of the delay, and then the outer mixers are added one by one and aligned. Once all of the calibration points are determined and all rotaries are set to crystal zero, the rotary counters are reset – all subsequent rotations refer to this zero-angle position.

Coherent PMD

In order to generate large changes in DGD and SOPMD in a small frequency band, thereby producing strong higher-order PMD states, the component PMD vectors must add constructively. As with any other optical effect, con-

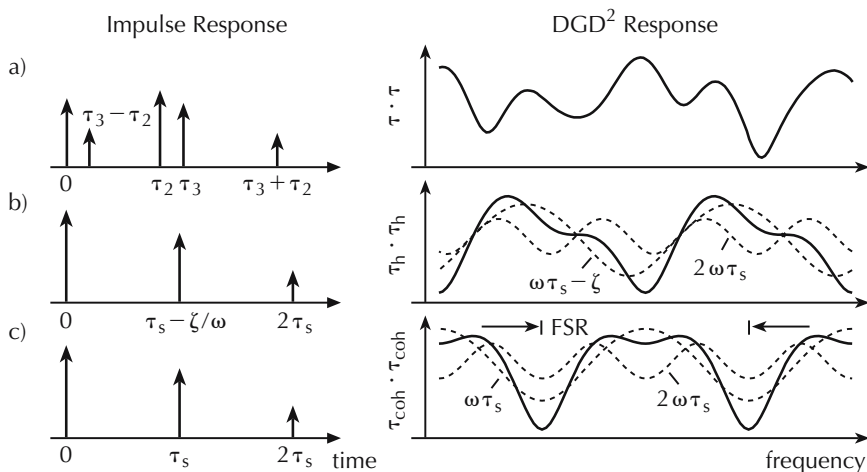


Fig. 10.19. Progression toward a coherent PMD spectrum for four delay stages. a) Four-stage incoherent spectrum. Each stage delay is different, making five Fourier components. b) Four-stage harmonic spectrum. All stage delays are the same, but the residual birefringent phase is arbitrary. c) Four-stage coherent spectrum: the fundamental and second-harmonic phases are aligned. Maximum contrast is achieved. Its evident that birefringent phase plays a key role in the shape of the spectrum.

structive interference occurs when optical phases align. An excellent example is the birefringent filter and its prerequisite coherence [44]. In terms of PMD, constructive interference happens when the phases of the Fourier components are aligned. This is called coherent polarization mode dispersion [23, 26].

There are four stages in the ECHO instrument. Recall from (8.2.75) on page 370 that the general DGD-squared spectrum for four stages has five Fourier components: a DC component, components that correspond to the delays of the center two stages, and the sum and difference of these delays. This general case is shown in Fig. 8.36(c) on page 367 and redrawn in Fig. 10.19(a). This DGD spectrum is complicated, has a lower contrast ratio, and a long free-spectral range.

The first step toward coherency is to have all Fourier components be a multiple of a unit component. This is called harmonic PMD and occurs when the stage delays τ_k are multiples of a unit stage delay τ_s : $\tau_k = n\tau_s$ where n is an integer. When $n = 1$ for each stage but the residual phases remain arbitrary, the general expression (8.2.75) reduces to

$$\vec{\tau}_h \cdot \vec{\tau}_h = b_0 + b_1 \cos(\omega\tau_s - \zeta) + b_2 \cos 2\omega\tau_s$$

where h denotes harmonic and ζ is the phase offset measured in relation to $2\omega\tau_s$. This offset is non-zero when the residual phases of the center two sections differ. Its effect is shown in Fig. 10.19(b): there are three non-degenerate Fourier components rather than five, but since the residual phases do not

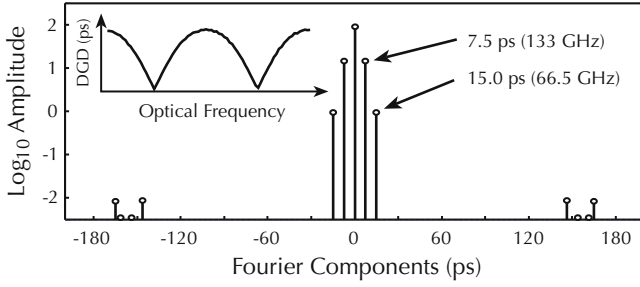


Fig. 10.20. Fourier transform of magnitude-squared DGD spectrum (inset) measured from a four-stage coherent PMD source having stage delay $\tau_s = 7.5$ ps. The principal Fourier components are DC, τ_s , and $2\tau_s$. Vertical axis is on a logarithmic scale.

match the fundamental and second-harmonic Fourier components are not phase aligned. This spectrum is harmonic but not coherent.

To go from a harmonic to coherent spectrum the residual birefringent phase of the center sections must be controlled. The normalized phase ϕ for stage k is defined

$$\varphi_k = n\phi_k$$

Coherency requires $\phi_j = \phi_k$ for all stages j and k save for the first and last stage. When $n = 1$ for all stages, the birefringent phase of each contributing stage must be the same. The Evans phase shifter makes this situation possible. The coherent four-stage magnitude-squared DGD spectrum is then

$$\vec{\tau}_{\text{coh}} \cdot \vec{\tau}_{\text{coh}} = c_0 + c_1 \cos \omega \tau_s + c_2 \cos 2\omega \tau_s \quad (10.4.2)$$

where the subscript coh denotes a coherent spectrum. One possible spectrum is illustrated in Fig. 10.19(c): the two non-DC Fourier components have the same phase, so the components of the DGD-squared spectrum align. In this case maximum constructive interference is possible and the PMD excursions will exhibit their highest contrast over the shortest FSR.

A harmonic PMD spectrum is demonstrated in Fig. 10.20. An ECHO instrument was set to maximum mode mixing and its DGD spectrum was measured. The spectrum was numerically squared and its Fourier transform taken. The amplitude of that spectrum is shown in the figure. There are strong tones at DC, τ_s , and $2\tau_s$. This spectrum is in fact coherent as well as harmonic; the phases, while not plotted, were equal to within measurement limit.

In general, for a coherent N stage concatenation the magnitude-squared DGD spectrum has the form

$$\vec{\tau} \cdot \vec{\tau} = \sum_{n=0}^N c_n \cos n\tau_s \quad (10.4.3)$$

One can see from Fig. 10.19 the fundamental importance birefringent phase has on the shape of the PMD spectra. Even for the same mode mixing, change

of the phase relationship shifts the position of the component tones, which in turn changes the spectral shape.

Theory of Operation

The following theory of operation imposes some symmetries on the control of the instrument [27]. Referring to Fig. 10.18, there are three mode mixers and two phase shifters. Once the instrument is built and calibrated, the outer two mode mixers, motors 1 and 5, are tied together so that they always register the same angle. Also, the two phase shifters are operated either in common mode or differential mode. Common mode means the phase shifters change phase by the same amount, which is tantamount to frequency tuning the spectrum. Differential mode means that the shifters change phase by equal and opposite amounts, which changes the shape of the spectrum. Control of ECHO principally deals with common-mode control.

Section §8.2.4 derived the PMD concatenation rules for a cascade. ECHO uses half-wave waveplates as mode mixers, so there is a necessary modification to the equations. Given that all delay stages (being equal) are represented by $\vec{\tau}_s$ and the waveplates by Q_k , the cumulative PMD vector $\vec{\tau}$ is

$$\vec{\tau} = \vec{\tau}_s + R_s^{(4)} Q_3 \left(\vec{\tau}_s + R_s^{(3)} Q_2 \left(\vec{\tau}_s + R_s^{(2)} Q_1 \vec{\tau}_s \right) \right) \quad (10.4.4)$$

where the vectors and operators are defined as

$$\vec{\tau}_s = \tau_s \hat{r}_s \quad (10.4.5)$$

$$Q_n = 2(\hat{q}_n \hat{q}_n \cdot) - 1 \quad (10.4.6)$$

$$R_s^{(n)} = (\hat{r}_s \hat{r}_s \cdot) + \sin \varphi_n (\hat{r}_s \times) - \cos \varphi_n (\hat{r}_s \times \hat{r}_s \times) \quad (10.4.7)$$

and where τ_s is the stage delay, φ_n is the birefringent phase of the n^{th} segment, and $\hat{q}_n$ is the direction in Stokes space to which the n^{th} half-wave waveplate is oriented. In particular, a physical rotation of a half-wave waveplate by angle $\theta/2$ corresponds to a rotation in Stokes space by 2θ . Also, Eq. (10.4.4) explicitly separates the first and third mode mixers.

The magnitude-squared DGD spectrum is

$$\begin{aligned} \frac{\vec{\tau} \cdot \vec{\tau}}{\tau_s^2} &= 4 + 2\hat{r}_s \cdot Q_3 \hat{r}_s + 2\hat{r}_s \cdot Q_2 \hat{r}_s + 2\hat{r}_s \cdot Q_1 \hat{r}_s + 2\hat{r}_s \cdot Q_3 R_s^{(3)} Q_2 \hat{r}_s \\ &\quad + 2\hat{r}_s \cdot Q_2 R_s^{(2)} Q_1 \hat{r}_s + 2\hat{r}_s \cdot Q_3 R_s^{(3)} Q_2 R_s^{(2)} Q_1 \hat{r}_s \end{aligned} \quad (10.4.8)$$

Under the coplanar assumption, where all birefringent axes lie on the equatorial plane (cf. §8.2.7), the vector products are expanded as

$$\hat{r}_s \cdot Q_j \hat{r}_s = \cos 2\theta_j \quad (10.4.9)$$

$$\hat{r}_s \cdot Q_k R_s Q_j \hat{r}_s = \cos 2\theta_k \cos 2\theta_j + \sin 2\theta_k \sin 2\theta_j \cos \varphi_j \quad (10.4.10)$$

and

$$\begin{aligned}
\hat{r}_s \cdot Q_l R_s^{(l)} Q_k R_s^{(k)} Q_j \hat{r}_s &= \cos 2\theta_l \cos 2\theta_k \cos 2\theta_j \\
&+ \sin 2\theta_l \sin 2\theta_k \cos 2\theta_j \cos \varphi_l \\
&+ \cos 2\theta_l \sin 2\theta_k \sin 2\theta_j \cos \varphi_k \\
&- \sin 2\theta_l \cos 2\theta_k \sin 2\theta_j \cos \varphi_l \cos \varphi_k \\
&+ \sin 2\theta_l \sin 2\theta_j \sin \varphi_l \sin \varphi_k \quad (10.4.11)
\end{aligned}$$

Two simplifications are now used to reduce (10.4.8) to a tractable expression. The birefringent phases of the second and third stages are split into common and differential parts, with the following definition:

$$\varphi_2 = \varphi_s + \delta\varphi, \text{ and } \varphi_3 = \varphi_s - \delta\varphi \quad (10.4.12)$$

where $\varphi_s = \omega\tau_s$. Also, the first and third mode mixers are tied such that $\theta_3 = \theta_1$. With these conditions, the magnitude-squared DGD spectrum takes the form

$$\begin{aligned}
\vec{\tau} \cdot \vec{\tau} &= 16\tau_s^2 \cos^2 \theta_1 \left(\cos^2(\theta_2 - \theta_1) \right. \\
&- 2 \sin \theta_1 \cos \theta_1 \sin \theta_2 \cos \theta_2 (1 - \cos \varphi_s \cos \delta\varphi) \\
&+ \frac{1}{2} \sin^2 \theta_1 \cos^2 \theta_2 (1 - \cos 2\varphi_s) \\
&\left. - \frac{1}{2} \sin^2 \theta_1 \sin^2 \theta_2 (1 - \cos \delta\varphi) \right) \quad (10.4.13)
\end{aligned}$$

Several observations are made about (10.4.13). First, there are only three Fourier components: DC, $\cos \varphi_s$, and $\cos 2\varphi_s$. This spectrum is harmonic. Second, the oscillatory components appear in the expression only when mode mixing angle θ_1 is not zero. When $\theta_1 = 0$ the system reduces to a two-stage concatenation, which is wavelength flat. Third, as a side-effect of tying the first and third stages together, birefringent phase error $\delta\varphi$ does not differentially phase-shift the fundamental and second-order Fourier components but instead diminishes the amplitude of the constant and fundamental Fourier component amplitudes.

Explicit calculation of the $\vec{\tau}_\omega \cdot \vec{\tau}_\omega$ spectrum is difficult because of so many higher-order harmonics. However, the vector expression for $\vec{\tau}_\omega$ can be written and is readily evaluated at $\varphi_s = 0$. The recursive $\vec{\tau}_\omega$ sequence out to four stages is

$$\begin{aligned}
\vec{\tau}(1) &= \vec{\tau}_s & \vec{\tau}_\omega(1) &= 0 \\
\vec{\tau}(2) &= \vec{\tau}_s + R_s^{(2)} Q_1 \vec{\tau}(1) & \vec{\tau}_\omega(2) &= \vec{\tau}_s \times \vec{\tau}(2) \\
\vec{\tau}(3) &= \vec{\tau}_s + R_s^{(3)} Q_2 \vec{\tau}(2) & \vec{\tau}_\omega(3) &= \vec{\tau}_s \times \vec{\tau}(3) + R_s^{(3)} Q_2 \vec{\tau}_\omega(2) \\
\vec{\tau}(4) &= \vec{\tau}_s + R_s^{(4)} Q_1 \vec{\tau}(3) & \vec{\tau}_\omega(4) &= \vec{\tau}_s \times \vec{\tau}(4) + R_s^{(4)} Q_1 \vec{\tau}_\omega(3)
\end{aligned} \quad (10.4.14)$$

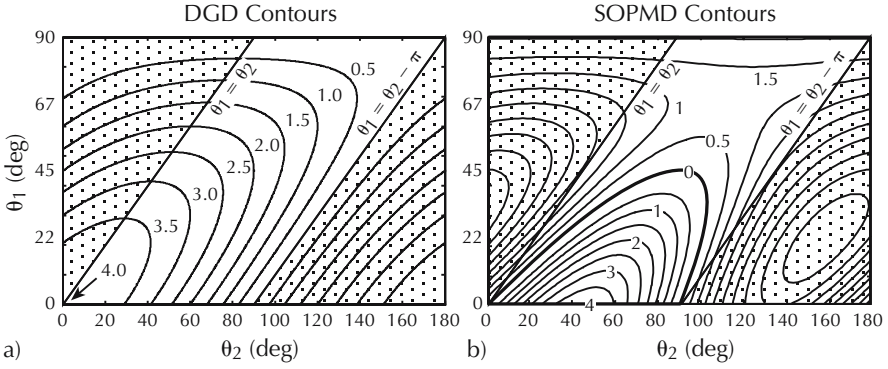


Fig. 10.21. At $\varphi_s = 0$, contours of constant τ and τ_ω scaled to $\tau_s = 1$. a) Constant τ contours, (10.4.16). Unshaded region is single-valued. b) Constant τ_ω contours, (10.4.17). Region bound by bold line is single-valued. Note the existence of contour $\tau_\omega = 0$.

At band-center, $\varphi_s = 0$, so $R_s = I$. The components of $\vec{\tau}_\omega(4)$ at this frequency are

$$\begin{aligned}\tau_{\omega 1} &= 0 \\ \tau_{\omega 2} &= 0 \\ \tau_{\omega 3} &= -2\tau_s^2 (\sin 2(\theta_2 - \theta_1)(1 + \cos 2\theta_1) - \sin 2\theta_1)\end{aligned}\tag{10.4.15}$$

The PMD coordinate (τ, τ_ω) for ECHO is defined at the calibration frequency. Since the instrument is calibrated to zero residual birefringent phase at this frequency, the precession angle is $\varphi_s = 0$. Taking the magnitude of respective $\vec{\tau}$ and $\vec{\tau}_\omega$ vectors, governed by (10.4.13) and (10.4.15), makes

$$|\vec{\tau}| = 4\tau_s |\cos \theta_1 \cos(\theta_2 - \theta_1)|\tag{10.4.16}$$

$$|\vec{\tau}_\omega| = 2\tau_s^2 |\sin 2(\theta_2 - \theta_1)(1 + \cos 2\theta_1) - \sin 2\theta_1|\tag{10.4.17}$$

These two equations map the independent variables to the dependent variables: $(\theta_1, \theta_2) \rightarrow (\tau, \tau_\omega)$. At the calibration frequency the first- and second-order PMD magnitudes are independent.

Figure 10.21 shows contours of constant τ and τ_ω . In Fig. 10.21(a) contours of constant τ are plotted as a function of (θ_1, θ_2) , where the plot is scaled to $\tau_s = 1$. The magnitude is bound between $0 \leq \tau \leq 4$. The unshaded area designates a region of monotonic, single-valued mapping of $(\theta_1, \theta_2) \rightarrow \tau$. In Fig. 10.21(b) contours of constant τ_ω are plotted as a function of (θ_1, θ_2) , similar to Fig. 10.21(a). The magnitude is bound between $0 \leq \tau_\omega \leq 4$. The special contour $\tau_\omega = 0$ exists in the parametric space, and was independently discovered and reported by [84]. The contour delineated by the dark solid line designates a region of monotonic, single-valued mapping of $(\theta_1, \theta_2) \rightarrow \tau_\omega$.

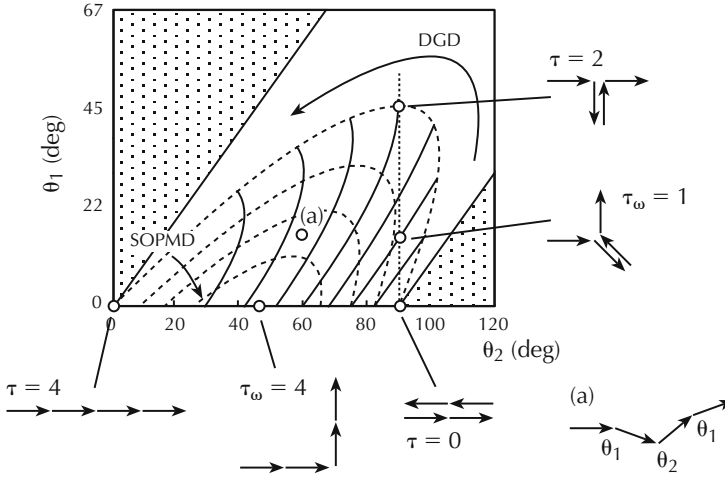


Fig. 10.22. At $\varphi_s = 0$, overlay of constant τ and τ_ω contours within single-valued region. There are two degrees of freedom, θ_1 and θ_2 , and two dependent variables, τ and τ_ω . At $\varphi_s = 0$ first- and second-order PMD are independent quantities. Several vector diagrams indicate interesting coordinates.

Figure 10.22 combines the (τ, τ_ω) contours in an area in which both coordinates are single-valued. Within this area, the mapping $(\tau, \tau_\omega) \rightarrow (\theta_1, \theta_2)$ is unique. Numerical inversion of (10.4.16-10.4.17) gives (θ_1, θ_2) for a specified (τ, τ_ω) .

There are interesting special cases on the contour map of Fig. 10.22; these are treated with the assistance of the vector diagrams of Fig. 10.23. Figure 10.23(a) shows the general case of four equal-length component PMD vectors where the mode mixing between the outer two-stage pairs is equal, i.e. $\theta_1 = \theta_3$. When $\theta_1 = \theta_2 = 0$, the vectors are aligned and create the maximum DGD of $\tau = 4\tau_s$ with concurrent $\tau_\omega = 0$ (Fig. 10.23(b)). When $\theta_1 = 0$ then the four-stage reduces to a symmetric two-stage. The two-stage maximum SOPMD is when $\theta_2 = \pi$: $\tau_\omega = (2\tau_s)^2$ with a concurrent $\tau = 4\tau_s/\sqrt{2}$ (Fig. 10.23(c)). The abscissa on Fig. 10.22 shows the locus of possible (τ, τ_ω) coordinates for the two-stage case. $\tau_\omega = 0$ is only possible at $\tau = 0$ and $\tau = 4\tau_s$. The inclusion of θ_1 as a free variable adds a necessary degree of freedom to trace the $\tau_\omega = 0$ contour over the entire range $0 \leq \tau \leq 4\tau_s$. Outside of the indicated monotonic region lies the point of maximum PDCD; such a point is illustrated in Fig. 10.23(d). When the four vectors form a square in Stokes space, the DGD is zero and the depolarization is also zero. The PDCD, however, is generated by the combined differential motions of τ_4 precession about τ_3 and these two vector's precession about τ_2 .

Four equations summarize the parameters of an ECHO source. These parameters include the extrema points described above as well as a measure of the source bandwidth:

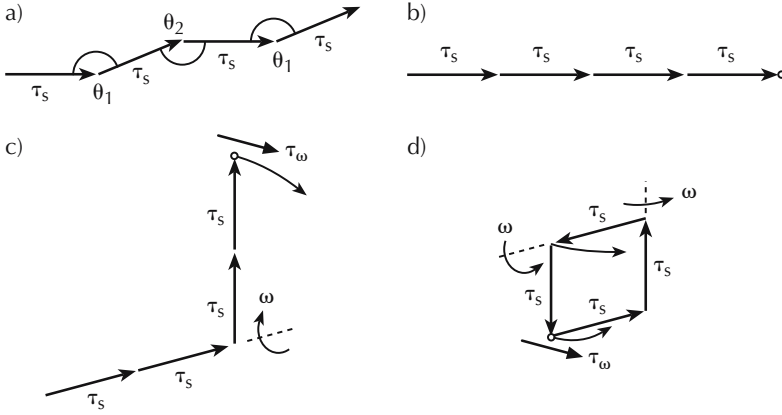


Fig. 10.23. Four-component vector diagrams. a) Arbitrary configuration but with first and third mode mixers tied. b) Maximum DGD of $4\tau_s$; all vectors align. c) Maximum two-stage SOPMD of $(2\tau_s)^2$; right angle between first and last pair of stages. d) Maximum PDCD of $2\tau_s^2$; all vectors are right angles.

$$\tau_{\max} = 4\tau_s \quad (10.4.18)$$

$$\tau_{w \max} = (2\gamma\tau_s)^2 \quad (10.4.19)$$

$$|\vec{\tau}|_{\omega \max} = 2\tau_s^2 \quad (10.4.20)$$

$$\text{FSR} = 1/\tau_s \quad (10.4.21)$$

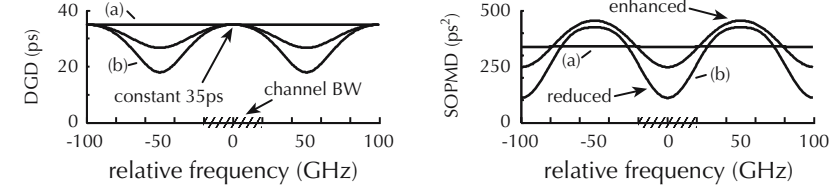
where γ is an enhancement factor due to the combined SOPMD effects of depolarization and PDCD. For a two-stage source $\gamma = 1$, but the numeric calculation of the four-stage source shows $\gamma \sim 1.09$. It is not known if the enhancement factor γ can be derived analytically.

Equations (10.4.18–10.4.21) show the inherent tradeoff for a four-stage source, where maximum PMD values tradeoff against the free-spectral range. The bandwidth of the PMD spectrum should be larger than the data channel bandwidth, which sets a maximum on τ_s . At the same time the maximum DGD and SOPMD should be representative of what a data channel would likely experience when run on a fiber with a mean PMD of $\langle\tau\rangle$. As with the PMDS-type sources, a maximum delay of $|\tau|_{\max} = 1.2T$ is a reasonable trade-off for NRZ transmission formats. However, this bandwidth is not suitable for RZ formats – this is discussed below.

Finally, by setting $\theta_1 = 0$, ECHO reverts to a symmetric two-stage source:

$$\begin{aligned} |\vec{\tau}| &= 4\tau_s |\cos \theta_2| \\ |\vec{\tau}_\omega| &= 4\tau_s^2 |\sin 2\theta_2| \end{aligned} \quad (10.4.22)$$

Performance results of various ECHO implementations are available in the technical literature [24, 26, 27].



PSP spectra:

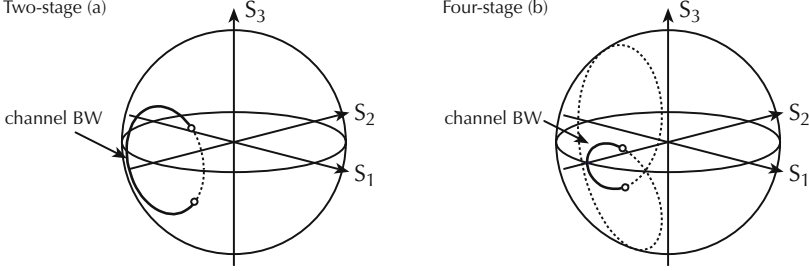


Fig. 10.24. Three calculated spectra pairs, $\tau(f - f_o)$ and $\tau_\omega(f - f_o)$, from ECHO addresses (35, 338), (35, 248), and (35, 111). The calculation uses $\tau_s = 10$ ps/stg. Observations: 1) τ is constant at 35 ps for each spectrum; 2) τ_ω is depressed at $\varphi_s = 0$ for latter two spectra; and 3) τ_ω is enhanced at $\varphi_s = \pi$ for latter two spectra.

Independent Control of $\tilde{\tau}$ And $\tilde{\tau}_\omega$

The governing equations (10.4.16-10.4.17) show that at the calibration frequency ($\varphi_s = 0$), or any integral multiple of the FSR, τ and τ_ω are independent. Figure 10.24 illustrates calculated τ and τ_ω spectra for three different states: (35, 338), (35, 248), and (35, 111). Each coordinate pair corresponds to $(\tau, \tau_\omega) \rightarrow (\text{ps}, \text{ps}^2)$. Three observations are apparent. First, at $\varphi_s = 0$ the DGD is constant at $\tau = 35$ ps, as predicted by the state address. Second, at the same relative frequency, the τ_ω values are progressively depressed from the two-stage case, that case corresponding to (35, 338). Third, the τ_ω value at $\pm\text{FSR}/2$ from center is enhanced with respect to the two-stage case, the value being the combined result of depolarization and non-zero PDCD.

The reason for the diminution of the SOPMD magnitude at center frequency is shown in the lower Poincaré plots. PMD produced by two stages has a PSP spectrum that traces circles. The angular rate of change with frequency is constant across the FSR, so the magnitude-SOPMD is constant. For the four-stage case, the PSP spectrum slows along the small arc and speeds along the large arc. On the small arc the pointing direction changes slowly, even for the same DGD. In the limit of zero SOPMD, the PSP pirouettes about a single point and the radius of the small arc is zero.

Figure 10.25 illustrates exemplar DGD, SOPMD, and PDCD spectra calculated for a 10 Gb/s instrument at two states: (35, 0) and (85, 1400). In Fig. 10.25(a), the (30, 0) state is shown because of the interesting property

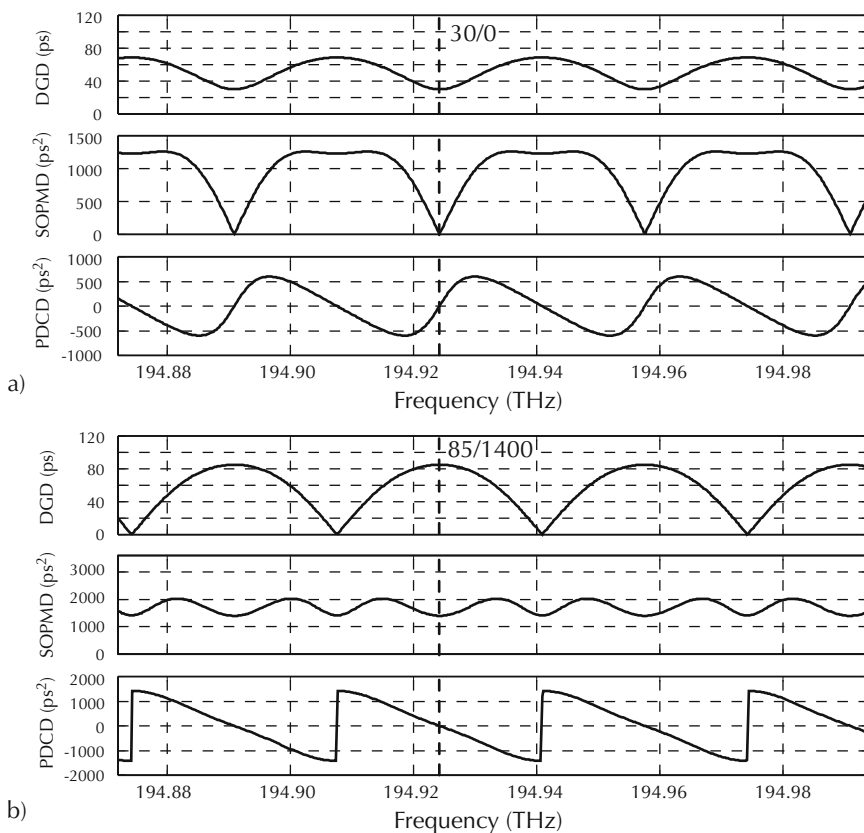


Fig. 10.25. Calculated scalar PMD spectra, $\tau_s = 10$ ps. a) State (30,0). At approximately 194.925 THz one observes $\tau = 30$ ps and $\tau_\omega = 0$ ps², the state setting. b) State (85,1400). In contrast to (a), the DGD value touches zero with large simultaneous SOPMD.

that at $\varphi_s = 0$ the τ_ω is zero while τ is finite. In a small frequency band about $\varphi_s = 0$ $\vec{\tau}$ pirouettes about a stationary position in Stokes space. Outside of this band $\vec{\tau}$ conducts its depolarizing motion. In Fig. 10.25(b), the (85,1400) state is shown because τ is zero at $\varphi_s = 0$ while τ_ω is finite. This is an important state where τ_ω dominates. In fact, it is the PDCD that dominates the SOPMD as the state is virtually devoid of depolarization; only the length of the DGD vector changes in a small frequency band about zero phase. The component PMD vector orientation is similar to that illustrated in Fig. 10.23(d).

The Role of Birefringent Phase

Birefringent phase plays a central role in determining the shape of non-flat PMD spectra and is central to construction of programmable PMD sources.

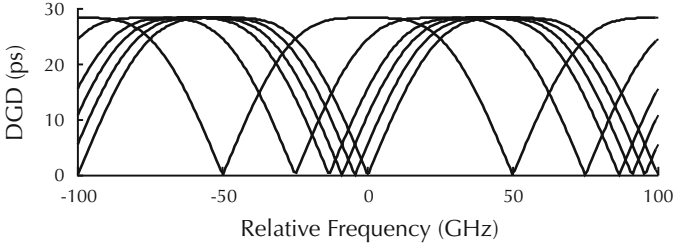


Fig. 10.26. Continuous frequency shifting. Six frequency-shifted τ spectra for $\theta_1 = \theta_2 = \pi/2$, $\tau_s = 10$ ps, and common-mode phase control. The spectra shape remains intact.

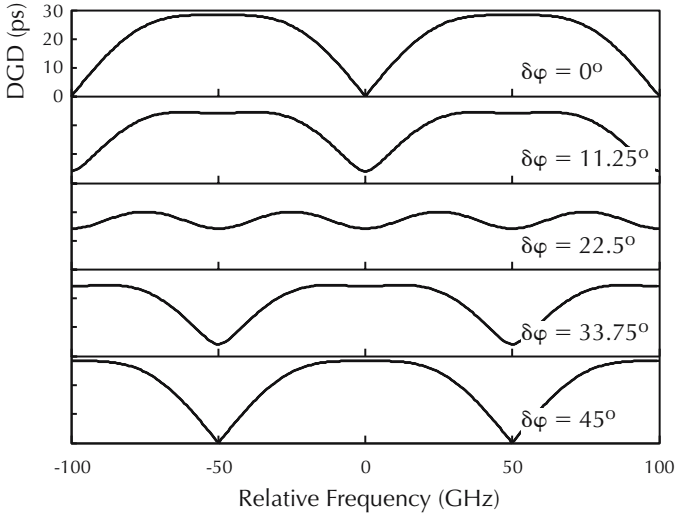


Fig. 10.27. Birefringent phase changes the DGD shape. Five τ spectra for $\theta_1 = \theta_2 = \pi/2$, $\tau_s = 10$ ps, and differential-mode phase control. The birefringent phase plays a central role in the spectral shape, which is predicted by (10.4.13).

This section demonstrates the criticality of birefringent phase using two examples: common and differential control of the birefringent phase. The Evans phase shifters in the second and third stages are used primarily to drive the concatenation into coherence. Once achieved, the phase controllers can be rotated simultaneously by the same angle. The result from this common-mode rotation is a frequency shift of the PMD spectrum [21]. Alternatively, the phase controllers can be rotated by equal and opposite amounts. The result from this differential-mode rotation is, for the highly symmetric ECHO, a change in the shape of the spectra but with zero movement of the Fourier phase of the constituent components.

Figure 10.26 shows the frequency shift of a DGD spectrum, calculated using common-mode phase control and with various phase increments. The Stokes angles of the mode mixers were all set to 90° . A detailed analysis of the Fourier components is available in [27].

In comparison to common-mode phase shift, the profound DGD shape change due to differential-mode phase control is shown in Fig. 10.27. Here the differential phase $\delta\varphi$ is varied between 0° and 180° in Stokes space; the Stokes angles of the mode mixers were all set to 90° . Unlike common-mode control, there is no frequency shift of the spectrum. Rather, the shape changes in place. The shape change is predicted by (10.4.13). In this equation the amplitude of the fundamental component vanishes when $\delta\varphi = 90^\circ$. Likewise, the amplitude of the DC component varies. In a fiber, differential phase change will change both the shape of the spectrum and its frequency centering.

Addressable PMD Space

ECHO instruments can continuously span first- and second-order PMD space within the envelope dictated by (10.4.22) and delineated by the outer-most contour in Fig. 10.17. However, this statement applies only to frequency $\varphi_s = 0$. If one includes all frequencies across all the possible spectra then a much wider addressable space is available. Figure 10.28 shows how to construct an envelope of the total addressable space. The figure is drawn with respect to a 10 Gb/s instrument but can be scaled to any other data rate. The dotted line shows the single contour for a symmetric two-stage source, derived from (10.4.22). Referring to the scalar spectra for the (30, 0) state in Fig. 10.25, all values for state (τ, τ_ω) are plotted parametrically on Fig. 10.28 along contour (a). Likewise, all values for state (85, 1400) are plotted along contour (b). As another example, the wavelength-flat state (100, 3316) is shown as just one point since there is no frequency dependence of that spectrum. The mapped

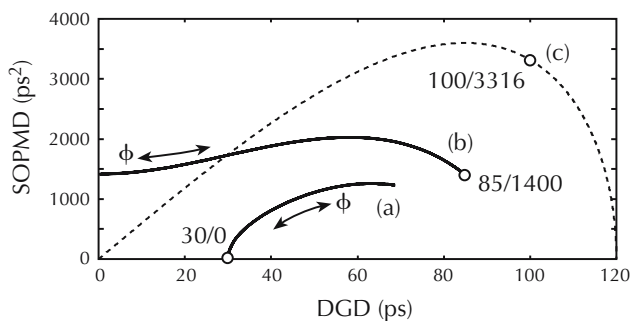


Fig. 10.28. State space for first- and second-order PMD magnitudes, scaled for $\tau_s = 30.0$ ps. Dashed line delineates two-stage contour. Contour (a) is parametric plot of scalar spectrum at address (30, 0). Likewise, contours (b) and (c) are parametric plots at addresses (85, 1400) and (100, 3316), respectively.

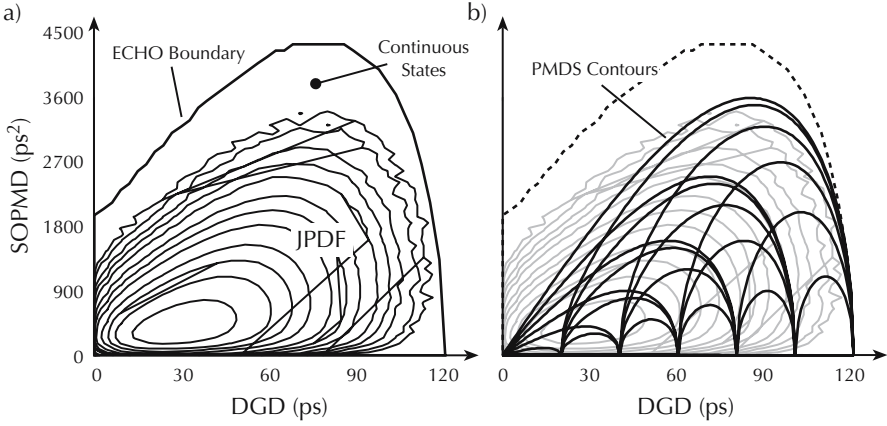


Fig. 10.29. Comparison of ECHO and PMDS addressable space in relation to fiber JPDF for $\langle \tau \rangle = 33$ ps. a) Addressable region for a 10 Gb/s ECHO lies below the boundary line and is continuous on the plane. b) Addressable region for a 10 Gb/s PMDS. The addressable states lie along lines and do not cover the entire space. Also, the low-DGD high-SOPMD wedge is not covered at all.

function is written

$$(\theta_1, \theta_2, \theta_\varphi) \longrightarrow (\tau, \tau_\omega) \quad (10.4.23)$$

where, with θ_φ being the angle of the Evans phase shifters, the left-hand side is a coordinate of physical parameters and the right-hand side is a coordinate of PMD parameters.

Following this approach for all combinations angles $(\theta_1, \theta_2, \theta_\varphi)$, where the four-stage concatenation remains coherent ($\varphi_3 = \varphi_2$) and where the first and third mode mixers are tied, all possible PMD addresses can be calculated. Figure 10.29(a) shows the results of this calculation. The region below the boundary shows the the addressable space of ECHO. The states are continuous; there are no holes in this two-dimensional surface. As a point of comparison, the addressable space for a 10 Gb/s PMDS is shown in Figure 10.29(b).

A richer mapping of physical to PMD-specific coordinates is

$$(\theta_1, \theta_2, \theta_\varphi) \longrightarrow (\tau, \tau_\omega, |\tau|_\omega) \quad (10.4.24)$$

where $|\tau|_\omega$ is the PDCD. Indeed there are three independent input variables, so one should expect three dependent variables. However, the inverse mapping of (10.4.24) is not one-to-one. One important inverse map is

$$(|\tau|_\omega; \tau, \tau_\omega) \longrightarrow (\theta_1, \theta_2, \theta_\varphi) \quad (10.4.25)$$

where τ and τ_ω remain fixed. This inverse map explores the balance between PDCD and depolarization at a fixed PMD coordinate (τ, τ_ω) . It would be very interesting to find how receiver sensitivity changes across the balance of second-order components.

Instrument Bandwidth

While it is significant that the ECHO instrument can smoothly cover a wide region of first- and second-order PMD space, this property alone is only a partial description and can be misleading. What is missing is a statement of the instrument's free-spectral range and its relation to the channel bandwidth. Figure 10.30(a) shows a spectral overlay of a 10 Gb/s ECHO DGD spectrum with a 10 Gb/s non-return to zero (NRZ) data channel bandwidth. The FSR of the source is 33.33 GHz, while the first channel null is at 10 GHz. By design the FSR is larger than the channel bandwidth.

Figure 10.30(b) shows a similar overlay with the same instrument but with a 12.7 Gb/s 33% duty-cycle return-to-zero (RZ) pulse bandwidth. The RZ channel bandwidth exceeds the FSR of the instrument. The built-in periodicity of the instrument imparts an artificial aliasing that would likely not exist in a real transmission system. Use of a 10 Gb/s ECHO source to test a 40 Gb/s data link is pointless because the channel bandwidth is many times the FSR of the instrument, even in spite of the fact that the 10 Gb/s instrument can reach suitably low first- and second-order PMD values.

While it is uneconomical to build one instrument to test both 10 Gb/s and 40 Gb/s data rates, a single source can be designed to accommodate NRZ and RZ transmission formats. Figure 10.30(c) illustrates one possibility. The center two vectors (all normalized to length 4) are split in a 3 : 1 ratio and the mode mixers between these stages are either aligned or crossed. When aligned, the four equal-length vector concatenation is recovered. When crossed, a 4 : 2 : 2 : 4 vector grouping appears. In this case the FSR is doubled. The FSR of the modified instrument can in this way “breathe” between a tight FSR and high PMD region and a looser FSR and a lower PMD region.

10.5 Receiver Performance Validation

There are two categories of information an operator of an optical communications link would like to have regarding PMD-induced impairments: what is the total outage probability (TOP) of the system, and what is the mean outage rate (R_{out}) as well as the mean outage duration (T_{out}). Total outage probability is a static estimate of the total number of severely-errored seconds (SES) a system will suffer over a period of time. Mean outage rate and duration are estimates of the dynamic behavior of the system under impairment. Since most links are operated in protected configurations, a few, or even one, severely-errored seconds may be enough to switch traffic onto the backup line. For the same number of SES a year, the question is whether the protection switch is thrown once and the full outage seconds occur, or whether the protection switch is thrown frequently for short time intervals. Frequent switching is deleterious to smooth system operation.

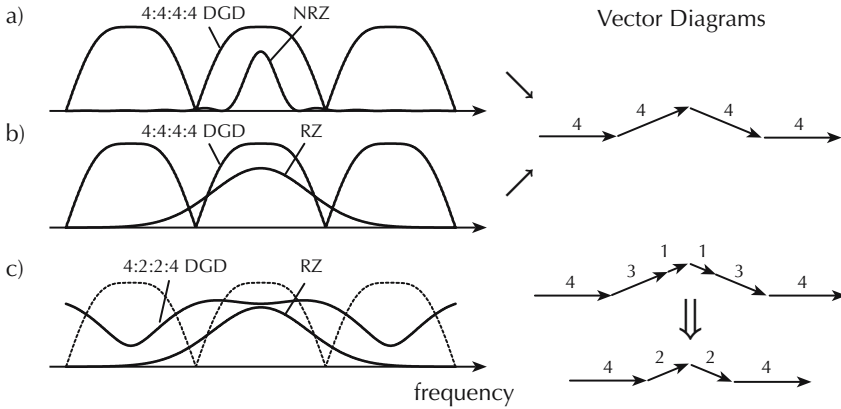


Fig. 10.30. Relation between instrument spectral periodicity and channel bandwidth. a-b) Overlay of a 10 Gb/s DGD spectrum with a 10 Gb/s NRZ and 12.7 Gb/s RZ bandwidths. In both cases the four component vector lengths are the same. c) Wide-FSR DGD spectrum and RZ bandwidth. Both with RZ bandwidth. Here the middle two stages are split in a 3 : 1 ratio. When vector of length 1 is folded back onto the vector of length 3, the net middle vector length is 2.

Static and dynamic estimates of PMD-induced system impairments can be estimated using a “receiver map” of a Tx/Rx pair and knowledge of the PMD statistics. Early estimates were derived by considering DGD alone, although there was cognizance of impairments due to SOPMD [58, 59]. Bulow subsequently showed the importance of including second-order PMD effects [5]. This philosophy is consistent with the view-point of this text: at least first- and second-order PMD must be considered to derive a meaningful outage estimate.

A Tx/Rx pair has a certain PMD tolerance that is independent of the fiber-optic line on which it operates. The receiver map isolates and quantifies this tolerance. A simple test setup to generate a receiver map is illustrated in Fig. 10.31. A single channel is driven by a bit-error-rate test setup (BERTS) and transmitted through a polarization scrambler, a programmable PMD source, and a fiber spool to introduce chromatic dispersion (CD). The channel is then noise-loaded prior to detection. In this “all-states” method, the bit-error rate (BER) must be averaged over a time interval such that the scrambler covers the entire Poincaré sphere (typically 5 mins using an Agilent 11896A polarization controller). A BER is recorded for each coordinate in PMD space and a contour plot of BER versus PMD is generated. Two such contour plots are illustrated in Fig. 10.32 [19, 25]. The contour plot is called the receiver map.

The receiver map provides qualitative and quantitative information about the Tx/Rx pair and its expected performance. Comparison of Figs. 10.32(a) and (b) shows that the latter receiver is more tolerant of PMD than the

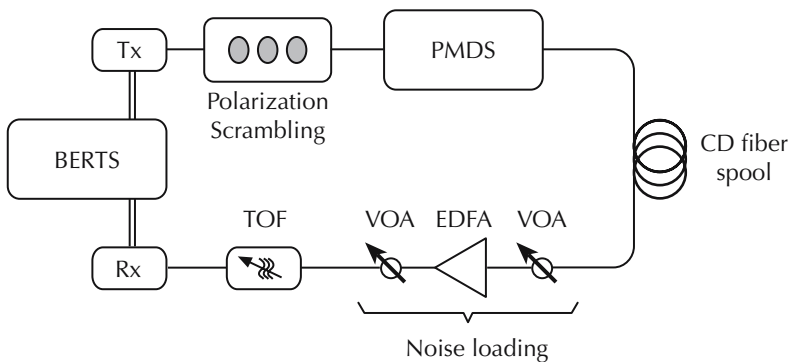


Fig. 10.31. Simple test configuration to generate a receiver map of the Tx/Rx pair. The bit-error rate is measured across a large number of PMD coordinates addressed by the PMDS. The channel is noise-loaded and chromatic dispersion can be added. For each state of the programmable PMD source, the bit-error rate (BER) is measured as an average over a uniform distribution of input polarization states. This is the so-called “all-states” method. The channel is noise-loaded using the two variable optical attenuators (VOA), an erbium amplifier (EDFA), and a tunable optical filter (TOF). Chromatic dispersion (CD) can be added parametrically to the receiver map.

former. Such behavior is usually found when a PMD compensator is added before the receiver. Other receiver comparisons have shown that some receivers are more tolerant of SOPMD than others. Finally, the receiver maps should be generated parametrically over the range of expected chromatic dispersion.

Total outage probability of a Tx/Rx pair has to include information of the mean PMD $\langle\tau\rangle$ on the fiber-optical link. The probability density for first- and second-order PMD is determined by the JPDF, which scales as $\langle\tau\rangle$. Therefore, the receiver map is compared to the JPDF as $\langle\tau\rangle$ is varied across the expected range in the physical plant. Two estimates can be generated: the expected error (rate $E[\text{BER}]$) and the total outage probability. The respective expressions are:

$$E[\text{BER}](\langle\tau\rangle) = \sum_{\tau, \tau_\omega} \text{BER}(\tau, \tau_\omega) P(\tau, \tau_\omega; \langle\tau\rangle) \quad (10.5.1)$$

$$\text{TOP}(\langle\tau\rangle) = \sum_{\tau, \tau_\omega} I(\text{BER}(\tau, \tau_\omega) > \text{TOL}) P(\tau, \tau_\omega; \langle\tau\rangle) \quad (10.5.2)$$

The expected error rate is simply the weighted average of the receiver map with the JPDF ($P(\tau, \tau_\omega; \langle\tau\rangle)$) scaled to a particular mean PMD. TOP is estimated using the JPDF and the indicator function, where $I = 0$ when the BER is below threshold TOL and $I = 1$ above the threshold.

For a particular receiver map, $E[\text{BER}]$ and TOP can be estimated over a range of $\langle\tau\rangle$. This is illustrated in Fig. 10.33. Since the JPDF is parametric in $\langle\tau\rangle$, TOP can be calculated parametrically. Important considerations are

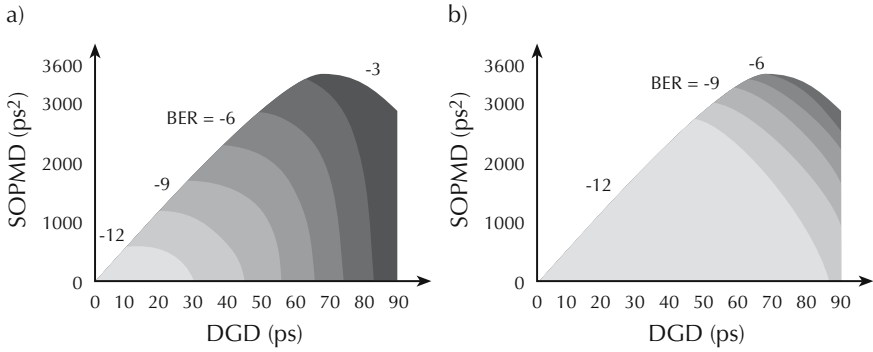


Fig. 10.32. Illustration of two receiver maps generated by a PMDS. Bands of constant bit-error rate across first- and second-order PMD space indicate the Tx/Rx tolerance to PMD. Combined with the JPDF of PMD, estimates of the total outage probability are made. a) Poor tolerance to PMD: there is a quick roll-off in BER with both first- and second-order. b) Improved tolerance.

the extent of the JPDF into low-probability regions and the estimation accuracy. The JPDF calculated by brute-force (see page 406) extends to 10^{-4} , which is not low enough to generate accurate estimates for $\langle\tau\rangle < \sim 15$ ps. The importance-sampling or direction-integration approaches resolve this problem (see page 405). The estimation accuracy depends on the density of the receiver map and coverage of the 2D PMD space. The receiver maps illustrated in Fig. 10.32 could be extended to low DGD, high SOPMD regions using an ECHO source.

Dynamic outage estimates such as R_{out} and T_{out} require a dynamic model of the PMD evolution and, most critically, a time constant with which the evolution takes place. That undersea fiber changes at a much slower rate than aerial fiber is clear. Caponi *et al.* made the first estimates based on measurements of installed terrestrial fiber [6]. Their technique uses the DGD evolution alone and the classic level-crossing-rate expression for Brownian motion. That expression requires densities for both the DGD and its temporal derivative. Caponi *et al.* use measured data to estimate the rate of change, and conjecture, after data analysis, that a particular DGD value and its temporal derivative are statistically independent.

Leo *et al.* extends the Caponi method with the conjecture that the joint-density of first- and second-order PMD and its joint temporal derivative are also independent [73, 87]. Their analysis of first- and second-order data supports the Caponi conjecture regarding DGD alone. Rather than using the one-dimensional level-crossing expression, Leo *et al.* use a receiver map and a two-dimensional indicator function. Therefore, based on measured fiber fluctuations and measured Tx/Rx performance, a simulation of fiber evolution is made in two PMD dimensions to estimate R_{out} .

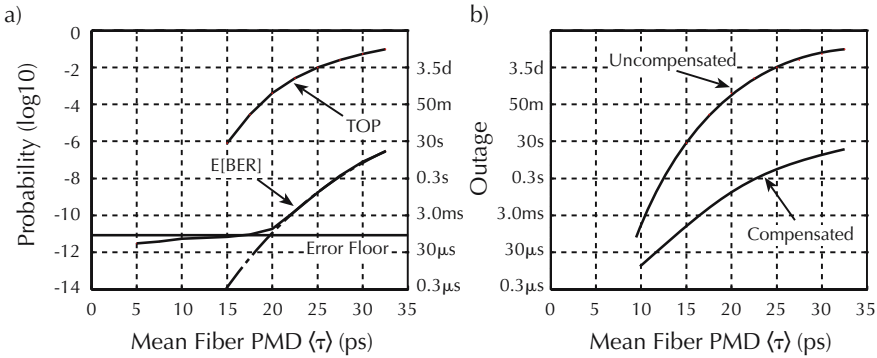


Fig. 10.33. Illustrative estimates of $E[\text{BER}]$ and TOP over a range of $\langle\tau\rangle$. The error floor relates to the minimum measured BER, and can be reduced with longer averaging times. Outage and probability are related on the abscissa. a) Comparison of TOP and $E[\text{BER}]$. b) Exemplar (un)compensated Tx/Rx TOP estimates.

The dynamic model of PMD evolution proposed by Leo lies solely on the JPDF. An alternative is to use a waveplate model of the fiber and rotate the sections in a random way. The drawback of such an evolving waveplate model is that most of the PMD states will be about the mean. Importance-sampling (IS) methods can be used for this problem as well. Earlier, IS was used to generate the JPDF for first- and second-order PMD. This is a density; what is needed is a process. Augmentation of the IS method to mimic the temporal evolution of fiber in a biased manner would be a powerful tool for robust estimates of the dynamic outage parameters.

References

1. A. J. Barlow, T. G. Arnold, T. L. Voots, and P. J. Clark, "Method and apparatus for high resolution measurement of very low levels of polarization mode dispersion (PMD) in single mode optical fibers and for calibration of PMD measuring instruments," U.S. Patent 5,654,793, Aug. 5, 1997.
2. G. Biondini, W. L. Kath, and C. R. Menyuk, "Non-maxwellian DGD distributions of PMD emulators," in *Tech. Dig., Optical Fiber Communications Conference (OFC'01)*, Anaheim, CA, Mar. 2001, paper ThA5.
3. G. Biondini and W. L. Kath, "Polarization-mode dispersion emulation with maxwellian lengths and importance sampling," *IEEE Photonics Technology Letters*, vol. 16, no. 3, pp. 789–791, Mar. 2004.
4. C. F. Buhrer, "Higher-order achromatic quarterwave combination plates and tuners," *Applied Optics*, vol. 27, no. 15, pp. 3166–3169, 1988.
5. H. Bulow, "System outage probability due to first- and second-order PMD," *IEEE Photonics Technology Letters*, vol. 10, no. 5, pp. 696–698, 1998.
6. R. Caponi, B. Riposati, A. Rossaro, and M. Schiano, "WDM design issues with highly correlated PMD spectra of buried optical cables," in *Tech. Dig., Optical Fiber Communications Conference (OFC'02)*, Anaheim, CA, Mar. 2002, paper ThI5, pp. 453–454.
7. L. Chen, O. Chen, S. Hadjifaradji, and X. Bao, "Polarization-mode dispersion measurement in a system with polarization-dependent loss or gain," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 206–208, Jan. 2004.
8. R. Chipman and R. Kinnera, "High-order polarization mode dispersion emulator," *Optical Engineering*, vol. 41, no. 5, pp. 932–937, May 2002.
9. E. Collett, "Automatic determination of the polarization state of nanosecond laser pulses," U.S. Patent 4,158,506, June 19, 1979.
10. F. Corsi, A. Galtarossa, and L. Palmieri, "Polarization mode dispersion characterization of single-mode optical fiber using backscattering technique," *Journal of Lightwave Technology*, vol. 16, no. 10, pp. 1832–1843, Oct. 1998.
11. —, "Beat length characterization based on backscattering analysis in randomly perturbed single-mode fibers," *Journal of Lightwave Technology*, vol. 17, no. 7, pp. 1172–1178, July 1999.
12. R. M. Craig, "Visualizing the limitations of four-state measurement of PDL and results of a six-state alternative," in *Symposium on Optical Fiber Measurements (SOFM 2002)*, Boulder, Colorado, Sept. 2002, pp. 121–124.
13. —, "Accurate spectral characterization of polarization-dependent loss," *Journal of Lightwave Technology*, vol. 21, no. 2, pp. 432–437, Feb. 2003.
14. R. M. Craig, S. L. Gilbert, and P. D. Hale, "High-resolution, nonmechanical approach to polarization-dependent transmission measurements," *Journal of Lightwave Technology*, vol. 16, no. 7, pp. 1285–1294, July 1998.
15. N. Cyr, "Stokes parameter analysis method, the consolidated test method for PMD measurements," in *Proceedings of the National Fiber Optical Engineering Conference*, Chicago, IL, 1999.
16. —, "Method and apparatus for measuring polarization mode dispersion of optical devices," U.S. Patent 6,204,924, Mar. 20, 2001.
17. —, "Polarization-mode dispersion measurement: Generalization of the interferometric method to any coupling regime," *Journal of Lightwave Technology*, vol. 22, no. 3, pp. 794–805, Mar. 2004.

18. J. N. Damask, "A programmable polarization-mode dispersion emulator for systematic testing of 10 Gb/s PMD compensators," in *Tech. Dig., Optical Fiber Communications Conference (OFC'00)*, Baltimore, MD, Mar. 2000, paper ThB3, pp. 28–30.
19. J. N. Damask, P. R. Myers, and T. R. Boschi, "Programmable polarization-mode-dispersion generation," in *Tech. Digest, Symposium on Optical Fiber Measurements (SOFM 2002)*, Boulder, CO, Sept. 2002.
20. J. N. Damask, "Apparatus and method for controlled generation of polarization mode dispersion," U.S. Patent 6,377,719, Apr. 23, 2002.
21. —, "Methods and apparatus for frequency shifting polarization mode dispersion spectra," U.S. Patent 2002/0 080 467 A1, June 27, 2002.
22. —, "Methods and apparatus for generating polarization mode dispersion," U.S. Patent 2002/0 191 285 A1, Dec. 19, 2002.
23. —, "Methods and apparatus for generation and control of coherent polarization mode dispersion," U.S. Patent 2002/0 118 455 A1, Aug. 29, 2002.
24. —, "Methods to construct programmable PMD sources, Part I: Technology and theory," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 997–1005, Apr. 2004.
25. J. N. Damask, G. Gray, P. Leo, G. J. Simer, K. B. Rochford, and D. Veasey, "Method to measure and estimate total outage probability for PMD-impaired systems," *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 48–50, Jan. 2003.
26. J. N. Damask, P. R. Myers, A. Boschi, and G. J. Simer, "Demonstration of a coherent PMD source," *IEEE Photonics Technology Letters*, vol. 15, no. 11, pp. 1612–1614, Nov. 2003.
27. J. N. Damask, P. R. Myers, G. J. Simer, and A. Boschi, "Methods to construct programmable PMD sources, Part II: Instrument demonstrations," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1006–1013, Apr. 2004.
28. J. N. Damask, G. J. Simer, K. B. Rochford, and P. R. Myers, "Demonstration of a programmable PMD source," *IEEE Photonics Technology Letters*, vol. 15, no. 2, pp. 296–298, Feb. 2003.
29. A. B. dos Santos and J. P. von der Weid, "PDL effects in PMD emulators made out with HiBi fibers: Building PMD/PDL emulators," *IEEE Photonics Technology Letters*, vol. 16, no. 2, pp. 452–454, Feb. 2004.
30. T. Erdogan, T. A. Strasser, and P. S. Westbrook, "In-line all-fiber polarimeter," U.S. Patent 6,211,957, Apr. 3, 2001.
31. J. W. Evans, "The birefringent filter," *Journal of the Optical Society of America*, vol. 39, no. 3, pp. 229–242, 1949.
32. A. Eyal, D. Kuperman, O. Dimenstein, and M. Tur, "Polarization dependence of the intensity modulation transfer function of an optical system with PMD and PDL," *IEEE Photonics Technology Letters*, vol. 14, no. 11, pp. 1515–1517, Nov. 2002.
33. A. Eyal and M. Tur, "Measurement of polarization mode dispersion in systems having polarization dependent loss or gain," *IEEE Photonics Technology Letters*, vol. 9, no. 9, pp. 1256–1258, Sept. 1997.
34. —, "A modified poincare sphere technique for the determination of polarization-mode dispersion in the presence of differential gain/loss," in *Tech. Dig., Optical Fiber Communications Conference (OFC'98)*, San Jose, CA, Feb. 1998, paper ThR1, p. 340.

35. D. L. Favin, B. M. Nyman, and G. M. Wolter, "System and method for measuring polarization dependent loss," U.S. Patent 5,371,597, Dec. 6, 1994.
36. K. S. Feder, P. S. Westbrook, J. Ging, P. I. Reyes, and G. E. Carver, "In-fiber spectrometer using tilted fiber gratings," *IEEE Photonics Technology Letters*, vol. 15, no. 7, pp. 933–935, July 2003.
37. A. Galtarossa and L. Palmieri, "Spatially resolved PMD measurements," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1103–1115, Apr. 2004.
38. A. Galtarossa, L. Palmieri, A. Pizzinat, M. Schiano, and T. Tambosso, "Measurement of local beat length and differential group delay in installed single-mode fibers," *Journal of Lightwave Technology*, vol. 18, no. 10, pp. 1389–1394, Oct. 2000.
39. A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, "Statistical characterization of fiber random birefringence," *Optics Letters*, vol. 25, no. 18, pp. 1322–1324, Sept. 2000.
40. N. Gisin, J. Von der Weid, and R. Passy, "Definitions and measurements of polarization mode dispersion: Interferometric versus fixed analyzer methods," *IEEE Photonics Technology Letters*, vol. 6, no. 6, pp. 730–732, 1994.
41. N. Gisin and K. Julliard, "Method and device for measuring polarization mode dispersion of an optical fiber," U.S. Patent 5,852,496, Dec. 22, 1998.
42. J. P. Gordon, R. M. Jopson, H. W. Kogelnik, and L. E. Nelson, "Polarization mode dispersion measurement," U.S. Patent 6,519,027, Feb. 11, 2003.
43. P. S. Hage, *Polarized Light: Instruments, Devices, Applications*. Bellingham, Washington: SPIE Optical Engineering Press, Jan. 1976, vol. 88, ch. Survey of Methods for the Complete Determination of a State of Polarisation, pp. 3–10.
44. S. E. Harris, E. O. Ammann, and I. C. Chang, "Optical network synthesis using birefringent crystals. i. synthesis of lossless networks of equal-length crystals," *Journal of the Optical Society of America*, vol. 54, no. 10, pp. 1267–1279, 1964.
45. M. C. Hauer, Q. Yu, and A. Willner, "Compact, all-fiber PMD emulator using an integrated series of thin-film micro-heaters," in *Tech. Dig., Optical Fiber Communications Conference (OFC'02)*, Anaheim, CA, Mar. 2002, paper ThA3.
46. M. Hauer, Q. Yu, E. Lyons, C. Lin, A. Au, H. Lee, and A. Willner, "Electrically controllable all-fiber PMD emulator using a compact array of thin-film microheaters," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1059–1065, Apr. 2004.
47. B. L. Heffner, "Automated measurement of polarization mode dispersion using Jones matrix eigenanalysis," *IEEE Photonics Technology Letters*, vol. 4, no. 9, pp. 1066–1068, 1992.
48. —, "Deterministic, analytically complete measurement of polarization-dependent transmission through optical devices," *IEEE Photonics Technology Letters*, vol. 4, no. 5, pp. 451–453, 1992.
49. —, "Accurate, automated measurement of differential group delay dispersion and principal state variation using Jones matrix eigenanalysis," *IEEE Photonics Technology Letters*, vol. 5, no. 7, pp. 814–816, 1993.
50. —, "Single-mode propagation of mutual temporal coherence: Equivalence of time and frequency measurements of polarization-mode dispersion," *Optics Letters*, vol. 19, no. 15, pp. 1104–1106, Aug. 1994.
51. —, "Optical pulse distortion measurement limitations in linear time invariant systems, and applications to polarization mode dispersion," *Optics Communications*, vol. 115, pp. 45–51, Mar. 1995.

52. —, “Influence of optical source characteristics on the measurement of polarization-mode dispersion of highly mode-coupled fibers,” *Optics Letters*, vol. 21, no. 2, pp. 113–115, Jan. 1996.
53. —, “PMD measurement techniques - a consistent comparison,” in *Tech. Dig., Optical Fiber Communications Conference (OFC’96)*, San Jose, CA, Feb. 1996, paper FA1, p. 292.
54. B. L. Heffner and P. R. Hernday, “Measurement of polarization-mode dispersion,” *Hewlett-Packard Journal*, pp. 27–33, Feb. 1995.
55. B. L. Heffner, “Method and apparatus for measuring polarization mode dispersion in optical devices,” U.S. Patent 5,227,623, July 13, 1993.
56. —, “Method and apparatus for measuring polarization sensitivity of optical devices,” U.S. Patent 5,298,972, Mar. 24, 1994.
57. —, “Polarimeter re-calibration method and apparatus,” U.S. Patent 5,296,913, Mar. 22, 1994.
58. F. Heismann, “Tutorial: Polarization mode dispersion: Fundamentals and impact on optical communication systems,” in *European Conference on Optical Communication (ECOC’98)*, vol. 2, Sept. 1998, pp. 51–79.
59. F. Heismann, D. A. Fishman, and D. L. Wilson, “Automatic compensation of first order polarization mode dispersion in a 10 gb/s transmission system,” in *European Conference on Optical Communication (ECOC’98)*, vol. 1, Sept. 1998, pp. 529–530.
60. B. Huttner, B. Gisin, and N. Gisin, “Distributed PMD measurement with a Polarization-OTDR in optical fibers,” *Journal of Lightwave Technology*, vol. 17, no. 10, pp. 1843–1848, Oct. 1999.
61. B. Huttner, J. Reecht, N. Gisin, R. Passy, and J. Weid, “Local birefringence measurements in single-mode fibers with coherent optical frequency-domain reflectometry,” *IEEE Photonics Technology Letters*, vol. 10, no. 10, pp. 1458–1460, Oct. 1998.
62. —, “Distributed beatlength measurement in single-mode fibers with optical frequency-domain reflectometry,” *Journal of Lightwave Technology*, vol. 20, no. 5, pp. 828–835, May 2002.
63. I. T. Lima, Jr., R. Khosravani, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, “Polarization mode dispersion emulator,” in *Tech. Dig., Optical Fiber Communications Conference (OFC 2000)*, Baltimore, MD, Mar. 2000, paper ThB4.
64. E. Ibragimov, G. Shtengel, and S. Suh, “Statistical correlation between first and second order PMD,” *Journal of Lightwave Technology*, vol. 20, no. 4, pp. 586–590, 2002.
65. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-12: Examinations and measurements - Polarization dependence of attenuation of a single-mode fibre optic component: Matrix calculation method*, International Electrotechnical Commission Std. IEC 61 300-3-12, 1997. [Online]. Available: <https://www.iec.ch/>
66. *Fibre optic interconnecting devices and passive components - Basic test and measurement procedures - Part 3-2: Examinations and measurements - Polarization dependence of attenuation in a single-mode fibre optic device*, International Electrotechnical Commission Std. IEC 61 300-3-2, 1999. [Online]. Available: <https://www.iec.ch/>
67. “JDS Uniphase instrumentation catalog 2003,” JDS Uniphase, Inc., Canada, 2003, PMD Emulator: PE3 or PE4. [Online]. Available: <http://www.jdsu.com/>

68. R. Jopson, L. Nelson, and H. Kogelnik, "Measurement of second-order polarization dispersion vectors in optical fibers," *IEEE Photonics Technology Letters*, vol. 12, no. 3, pp. 293–295, 2000.
69. R. M. Jopson, H. W. Kogelnik, and L. E. Nelson, "Method for measurement of first- and second-order polarization mode dispersion vectors in optical fibers," U.S. Patent 6,380,533, Apr. 30, 2002.
70. M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *Journal of Lightwave Technology*, vol. 18, no. 7, pp. 941–951, 2000.
71. M. Legre, M. Wegmuller, and N. Gisin, "Investigation of the ratio between phase and group birefringence in optical single-mode fibers," *Journal of Lightwave Technology*, vol. 21, no. 12, pp. 3374–3378, Dec. 2003.
72. P. J. Leo, G. R. Gray, G. J. Simer, and K. B. Rochford, "State of polarization changes: Classification and measurement," *Journal of Lightwave Technology*, vol. 21, no. 10, pp. 2189–2193, Oct. 2003.
73. P. J. Leo, D. L. Peterson, and K. B. Rochford, "Estimation of system outage statistics due to polarization mode dispersion," in *Symposium on Optical Fiber Measurements (SOFM 2002)*, Boulder, CO, Sept. 2002.
74. I. T. Lima, R. Khosravani, P. Ebrahimi, E. Ibragimov, C. R. Menyuk, and A. E. Willner, "Comparison of polarization mode dispersion emulators," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1872–1881, Dec. 2001.
75. C. Madsen, M. Cappuzzo, E. Laskowski, E. Chen, L. Gomez, A. Griffin, A. Wong-Foy, S. Chandrasekhar, L. Stulz, and L. Buhl, "Versatile integrated PMD emulation and compensation elements," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1041–1050, Apr. 2004.
76. C. Madsen, E. Laskowski, M. Cappuzzo, L. Buhl, S. Chandrasekhar, E. Chen, L. Gomez, A. Griffin, L. Stulz, and A. Wong-Foy, "A versatile, integrated emulator for first- and higher-order PMD," in *Tech. Dig., European Conference on Optical Communications (ECOC'03)*, Rimini, Italy, Sept. 2003, paper Th2.2.5.
77. B. S. Marks, I. T. Lima, and C. R. Menyuk, "Autocorrelation function for polarization mode dispersion emulators with rotators," *Optics Letters*, vol. 27, no. 13, pp. 1150–1152, July 2002.
78. P. Martin, G. Le Boudec, E. Taufflieb, and H. Lefevre, "Appliance for measuring polarization mode dispersion and corresponding measuring process," U.S. Patent 5,712,704, Jan. 27, 1998.
79. "National Aperture Catalogue," National Aperture, Inc., Salem, NH, 2004, MM-3M-R Rotary Stage. [Online]. Available: <http://www.naimotion.com/>
80. L. E. Nelson, R. M. Jopson, H. Kogelnik, and J. P. Gordon, "Measurement of polarization-mode dispersion vectors using the polarization-dependent signal delay method," *Optics Express*, vol. 6, no. 8, pp. 158–167, Apr. 2000. [Online]. Available: <http://www.opticsexpress.org/>
81. B. M. Nyman, D. L. Favin, and G. Wolter, "Automated system for measuring polarization dependent loss," in *Tech. Dig., Optical Fiber Communications Conference (OFC'94)*, San Jose, CA, Mar. 1994, pp. 230–231.
82. B. M. Nyman and G. Wolter, "High-resolution measurement of polarization dependent loss," *IEEE Photonics Technology Letters*, vol. 5, no. 7, pp. 817–818, July 1993.
83. J. Patscher and R. Eckhardt, "Component for second-order compensation of polarization-mode dispersion," *Electronics Letters*, vol. 33, no. 13, p. 1157, 1997.

84. P. B. Phua and H. A. Haus, "Variable differential-group-delay module without second-order PMD," *Journal of Lightwave Technology*, vol. 20, no. 9, pp. 1788–1794, 2002.
85. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization mode dispersion in long single-mode fibers," *Electronics Letters*, vol. 22, no. 19, pp. 1029–1030, 1986.
86. C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, 1994.
87. K. B. Rochford, P. J. Leo, D. L. Peterson, and P. Williams, "Recent progress in polarization mode dispersion measurement," in *Proc. 16th Intl. Conf. on Optical Fiber Sensors*, Nara, Japan, Oct. 2003.
88. A. J. Rogers, "Polarization-optical time domain reflectometry: A technique for the measurement of field distributions," *Applied Optics*, vol. 20, pp. 1060–1074, 1981.
89. G. Shtengel, 2003, Labview library for Agilent 8509 Optical Polarization Analyzer and Agilent 8614 Tunable Laser Source. [Online]. Available: <http://www.shtengel.com/gleb/Labview.htm>
90. —, private communication, 2003.
91. A. S. Siddiqui, "Optical polarimeter having four channels," U.S. Patent 5,227,623, Jan. 14, 1992.
92. *Polarization-Mode Dispersion Measurement for Single-Mode Optical Fibers by Interferometry Method*, Telecommunications Industry Association Std. TIA/EIA-455-124, 1999. [Online]. Available: <http://www.tiaonline.org/standards/>
93. *Differential Group Delay Measurement of Single-Mode Components and Devices by the Differential Phase Shift Method*, Telecommunications Industry Association Std. TIA/EIA-455-197, 2000. [Online]. Available: <http://www.tiaonline.org/standards/>
94. *Measurement of Polarization Dependent Loss (PDL) of Single-Mode Fiber Optic Components*, Telecommunications Industry Association Std. TIA/EIA-455-157, 2000. [Online]. Available: <http://www.tiaonline.org/standards/>
95. *Polarization-Mode Dispersion Measurement for Single-Mode Optical Fibers by the Fixed Analyzer Method*, Telecommunications Industry Association Std. TIA/EIA-455-113, 2001. [Online]. Available: <http://www.tiaonline.org/standards/>
96. *Polarization Mode Dispersion Measurement for Single-Mode Optical Fibers by Stokes Parameter Evaluation*, Telecommunications Industry Association Std. TIA/EIA-455-122, 2002. [Online]. Available: <http://www.tiaonline.org/standards/>
97. D. S. Waddy, L. Chen, and X. Bao, "A dynamical polarization mode dispersion emulator," *IEEE Photonics Technology Letters*, vol. 15, no. 4, pp. 534–536, Apr. 2003.
98. D. S. Waddy, L. Chen, S. Hadjifaradji, X. Bao, R. B. Walker, and S. J. Mihailov, "High-order PMD and PDL emulator," in *Tech. Dig., Optical Fiber Communications Conference (OFC'04)*, Los Angeles, CA, Feb. 2004, paper ThF6.
99. P. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 14, no. 2, pp. 148–157, Feb. 1995.

100. M. Wegmuller, S. Demma, C. Vinegoni, and N. Gisin, "Emulator of first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, no. 5, pp. 630–632, May 2002.
101. M. Wegmuller, F. Scholder, and N. Gisin, "Photon-counting OTDR for local birefringence and fault analysis in the metro environment," *Journal of Lightwave Technology*, vol. 22, no. 2, pp. 390–400, Feb. 2004.
102. P. S. Westbrook, T. A. Strasser, and T. Erdogan, "In-line polarimeter using blazed fiber gratings," *IEEE Photonics Technology Letters*, vol. 12, no. 10, pp. 1352–1354, Oct. 2000.
103. P. S. Westbrook, "System comprising in-line wavelength sensitive polarimeter," U.S. Patent 6,591,024, July 8, 2003.
104. P. A. Williams, "Mode-coupled artifact standard for polarization-mode dispersion: Design, assembly, and implementation," *Applied Optics*, vol. 38, no. 31, pp. 6498–6507, 1999.
105. —, "Modulation phase-shift measurement of PMD using only four launched polarisation states: A new algorithm," *Electronics Letters*, vol. 35, no. 18, pp. 1578–1579, 1999.
106. —, "Rotating-wave-plate stokes polarimeter for differential group delay measurements of polarization-mode dispersion," *Applied Optics*, vol. 38, no. 31, pp. 6508–6515, 1999.
107. —, "PMD measurement techniques avoiding measurement pitfalls," in *Venice Summer School on Polarization Mode Dispersion*, Venice Italy, June 2002, pp. 24–36.
108. P. A. Williams and A. J. Barlow, "Summary of current agreement among PMD measurement techniques," in *Presentation to International Electrotechnical Commission (IEC)*, Edinburgh, Scotland, Sept. 1997, paper SC86, SG1.
109. P. A. Williams, A. J. Barlow, C. Mackechnie, and J. B. Schlager, "Narrowband measurements of polarization-mode dispersion using the modulation phase shift technique," in *Tech. Digest, Symposium on Optical Fiber Measurements (SOFM 1998)*, Boulder, CO, Sept. 1998, pp. 23–26, NIST Special Publication 930.
110. P. A. Williams and J. D. Kofler, "Narrow-band measurement of differential group delay by a six-state RF phase-shift technique: 40 fs single-measurement uncertainty," *Journal of Lightwave Technology*, vol. 22, no. 2, pp. 448–456, Feb. 2004.
111. P. A. Williams and J. Kofler, "Measurement and mitigation of multiple-reflection effects on the differential group delay spectrum of optical components," in *Tech. Digest, Symposium on Optical Fiber Measurements (SOFM 2002)*, Boulder, CO, Sept. 2002, pp. 173–176.
112. P. A. Williams and C. M. Wang, "Corrections to fixed analyzer measurements of polarization mode dispersion," *Journal of Lightwave Technology*, vol. 16, no. 4, pp. 534–554, 1998.
113. L. Yan, M. Hauer, Y. Shi, X. Yao, P. Ebrahimi, Y. Wang, A. Willner, and W. Kath, "Polarization-mode-dispersion emulator using variable differential-group-delay (DGD) elements and its use for experimental importance sampling," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1051–1058, Apr. 2004.
114. L.-S. Yan, M. Hauer, C. Yeh, G. Yang, L. Lin, Z. Chen, Y. Q. Shi, X. S. Yao, A. E. Willner, and W. L. Kath, "High-speed, stable and repeatable PMD

emulator with tunable statistics,” in *Tech. Dig., Optical Fiber Communications Conference (OFC’03)*, Atlanta, GA, Mar. 2003, paper MF6.

A

Addition of Multiple Coherent Waves

There are many instances throughout the text where the simplification of a sum of coherent sine and cosine terms is necessary. Sine and cosine terms that are coherent have the same oscillatory frequency ωt but may have different amplitudes and phases. These waves can be combined into a single sine, cosine, or complex exponential expression. This appendix shows how to make the reductions.

A sum of N coherent exponentials is

$$S = \sum_{n=1}^N a_n e^{j(\omega t \pm \phi_n)} \quad (\text{A.1})$$

Expanding the sum makes

$$\begin{aligned} \sum a_n e^{\pm j\phi_n} &= \sum a_n \cos \phi_n \pm j \sum a_n \sin \phi_n \\ &= A \pm jB \end{aligned} \quad (\text{A.2})$$

where,

$$A = \sum_{n=1}^N a_n \cos \phi_n, \quad B = \sum_{n=1}^N a_n \sin \phi_n$$

Converting to polar form, the sum S simplifies to

$$\sum_{n=1}^N a_n e^{j(\omega t \pm \phi_n)} = \sqrt{A^2 + B^2} \exp(j(\omega t \pm \tan^{-1}(B/A))) \quad (\text{A.3})$$

The simplifications for sine and cosine sums requires an additional step of exponential expansion. Thus, with

$$S' = \sum_{n=1}^N a_n \sin(\omega t - \phi_n), \quad (\text{A.4})$$

Table A.1. Identities for Coherent Wave Addition

$$\sum_{n=1}^N a_n e^{j(\omega t \pm \phi_n)} = \sqrt{A^2 + B^2} \exp(j(\omega t \pm \tan^{-1} B/A))$$

$$\sum_{n=1}^N a_n \sin(\omega t \pm \phi_n) = \sqrt{A^2 + B^2} \sin(\omega t \pm \tan^{-1}(B/A))$$

$$\sum_{n=1}^N a_n \cos(\omega t \pm \phi_n) = \sqrt{A^2 + B^2} \cos(\omega t \mp \tan^{-1}(B/A))$$

$$\text{where } A = \sum_{n=1}^N a_n \cos \phi_n, \quad B = \sum_{n=1}^N a_n \sin \phi_n$$

exponential expansion of the sine terms yields

$$S' = \frac{1}{2j} \left\{ e^{j\omega t} \sum a_n e^{-j\phi_n} - e^{-j\omega t} \sum a_n e^{j\phi_n} \right\} \quad (\text{A.5})$$

Substitution of (A.2) into (A.5) yields

$$\begin{aligned} S' &= \frac{1}{2j} \left\{ A (e^{j\omega t} - e^{-j\omega t}) - jB (e^{j\omega t} + e^{-j\omega t}) \right\} \\ &= A \sin \omega t - B \cos \omega t \end{aligned} \quad (\text{A.6})$$

Recognizing that (A.6) is similar to the equation for an ellipse, the final simplification produces

$$\sum_{n=1}^N a_n \sin(\omega t \pm \phi_n) = \sqrt{A^2 + B^2} \sin(\omega t \pm \tan^{-1} B/A) \quad (\text{A.7})$$

In an analogous way, a sum of coherent cosine terms simplifies to

$$\sum_{n=1}^N a_n \cos(\omega t \pm \phi_n) = \sqrt{A^2 + B^2} \cos(\omega t \mp \tan^{-1} B/A) \quad (\text{A.8})$$

Table (A.1) summarizes the results.

B

Select Magnetic Field Profiles

Non-latching iron garnet Faraday rotation elements require the presence of an external magnetic field to saturate the magnetic domains. To generate the Faraday effect, the field lines are aligned predominantly in the direction of optical propagation. A cylindrical magnet with the center bore gives the required field profile and is also simple to analyze. This appendix gives analytic and semi-analytic expressions for the field along the centerline of the bore and in the plane perpendicular to the propagation direction, where the plane is located half-way along the bore.

A permanent magnet is described by a magnetic dipole distribution within the material and a magneto-quasi-static magnetic field. The relevant form of Maxwell's equations are then

$$\nabla \times \mathbf{H} = 0 \quad (\text{B.1a})$$

$$\nabla \cdot \mu_o \mathbf{H} = -\nabla \cdot \mu_o \mathbf{M} \quad (\text{B.1b})$$

Since the magnetic field is irrotational, the field can be defined as the gradient of a scalar potential (cf. 1.2.2b):

$$\mathbf{H} = -\nabla \Psi \quad (\text{B.2})$$

where Ψ is the scalar magnetic potential. Definition of the magnetic charge density ρ_m as

$$\rho_m = -\nabla \cdot \mu_o \mathbf{M} \quad (\text{B.3})$$

and substitution of (B.2) into (B.1b) yields Poisson's equation

$$\nabla^2 \Psi = -\frac{\rho_m}{\mu_o} \quad (\text{B.4})$$

A solution to Poisson's equation is the superposition integral

$$\Psi = \int_{V'} \frac{\rho_m(\mathbf{r}')}{4\pi\mu_o|\mathbf{r} - \mathbf{r}'|} dV' \quad (\text{B.5})$$

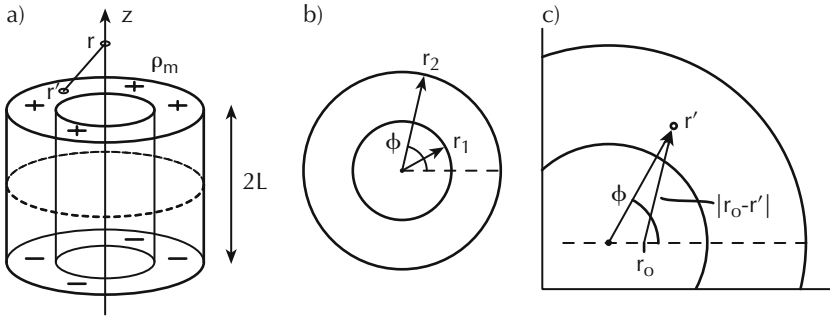


Fig. B.1. Geometry of cylindrical magnet. a) Cylindrical magnet with center bore. Magnetic charges lie on the top and bottom annular surfaces. b) Top view showing inner and outer radius. c) For in-plane calculation, position r_o is offset from the z -axis and can be related to angle ϕ (see text).

where the prime denotes a point on or within the magnetic medium and $\mathbf{r}$ is a spatial coordinate. The integral is taken over the volume of the magnetic solid. Figure B.1 illustrates the magnetic cylinder under consideration. Taking advantage of the cylindrical symmetry, the superposition integral is evaluated as

$$\Psi = \int_0^{2\pi} d\phi \int_{r_1}^{r_2} r' dr' \int_{-L}^L dz \frac{\rho_m(\mathbf{r}')}{4\pi\mu_o|\mathbf{r} - \mathbf{r}'|} \quad (\text{B.6})$$

The first evaluation of (B.6) is done along the z -axis. Integration of (B.6) along z yields two magnetic sheets, one annulus at $+L$ with positive “charges” $\mu_o M$ and the other annulus at $-L$ with negative charges $-\mu_o M$. Moreover, the distance from any point on the annular sheets to the z -axis is $|\mathbf{r} - \mathbf{r}'| = \sqrt{r'^2 + (z \mp L)^2}$, where the minus sign corresponds to the top sheet. The scalar potential, still in integral form, is

$$\Psi = \int_{r_1}^{r_2} \frac{2\pi\mu_o M r' dr'}{4\pi\mu_o \sqrt{r'^2 + (z - L)^2}} - \int_{r_1}^{r_2} \frac{2\pi\mu_o M r' dr'}{4\pi\mu_o \sqrt{r'^2 + (z + L)^2}} \quad (\text{B.7})$$

Integration and subsequently taking the gradient as prescribed by (B.2) yields the magnetic field strength along the central axis [1, 2]:

$$H_z(z) = -\frac{M}{2} \left((z - L) \left[\frac{1}{\sqrt{(z - L)^2 + r_2^2}} - \frac{1}{\sqrt{(z - L)^2 + r_1^2}} \right] - (z + L) \left[\frac{1}{\sqrt{(z + L)^2 + r_2^2}} - \frac{1}{\sqrt{(z + L)^2 + r_1^2}} \right] \right) \quad (\text{B.8})$$

Figure B.2(a) illustrates an evaluation of (B.8). A design goal would be to achieve the highest possible magnetic field in the region of $z = 0$ while minimizing the size of the magnet. The change in sign is due to the fields wrapping around either magnet end to terminate on the surface charges.

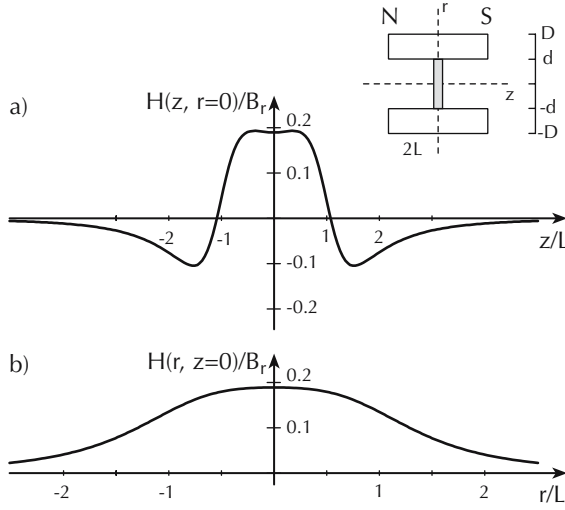


Fig. B.2. Axial and transverse magnetic field amplitude H_z of a cylindrical magnet. a) Axial field strength of H_z under the conditions $r_2 = L$ and $r_1 = L/2$. b) Transverse field strength of H_z in plane located at center of magnet. Inset shows coordinates.

The purpose of the following transverse field calculation is to explore the field uniformity within the bore but off-axis. Since an optical beam has finite width, it is insufficient to saturate the center of an iron garnet but not the outer edges. In order to keep the analysis quasi-analytic, only the field amplitude in a plane normal to z and located half-way along the bore is calculated.

The key to evaluating the superposition integral in this case is an analytic expression for $|\mathbf{r}' - \mathbf{r}|$ away from the z -axis. Referring to Fig. B.1(c), offset position r_o is related to r' and ϕ as

$$|r' - r_o| = \sqrt{\mathcal{R}^2 + (z \pm L)^2} \quad (\text{B.9})$$

where the in-plane length is

$$\mathcal{R}^2 = (r' \sin \phi)^2 + (r' \cos \phi - r_o)^2 \quad (\text{B.10})$$

With this measure, the superposition integral is

$$\Psi = \int_0^{2\pi} d\phi \int_{r_1}^{r_2} r' dr' \frac{\mu_o M}{4\pi \mu_o} \left[\frac{1}{\sqrt{\mathcal{R}^2 + (z + L)^2}} - \frac{1}{\sqrt{\mathcal{R}^2 + (z - L)^2}} \right] \quad (\text{B.11})$$

Even though (B.11) does not have an analytic form, an expression closer to $H_z(r, z = 0)$ can still be found. In particular,

$$H_z = -\frac{\partial}{\partial z} \Psi \quad (\text{B.12})$$

Carrying through the derivative with respect to z first and then setting $z = 0$ yields

$$H_z(r, z = 0) = \int_0^{2\pi} d\phi \int_{r_1}^{r_2} \frac{2\mu_o M}{4\pi\mu_o} \frac{r' dr'}{[\mathcal{R}^2 + 1]^{3/2}} \quad (\text{B.13})$$

This integral can be evaluated numerically. Applying the parameters from Fig. B.2(a) to (B.13) generates the curve given in Fig. B.2(b). Note that the z component of the magnetic field does not change sign but monotonically decays to zero far away from the magnet. Also, the uniformity of the field, for these parameters, remains within 10% of the peak within the inner radius.

A samarium-cobalt (SmCo) magnet can be an excellent choice for the permanent around an iron garnet due to its high coercivity in a small size. Length-diameter products of 1 mm^2 can readily achieve the 100–250 Oe magnetic field required for H_{sat} in iron garnets.

References

1. H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*. Englewood Cliffs, New Jersey: Prentice-Hall, 1989.
2. K. Shiraishi, F. Tajima, and S. Kawakami, "Compact faraday rotator for an optical isolator using magnets arranged with alternating polarities," *Optics Letters*, vol. 11, no. 2, pp. 82–84, 1986.

C

Efficient Calculation of PMD Spectra

Scalar and vector PMD spectra calculated in the Stokes-based PMD representation is straightforward and efficient. The concatenation rules presented in §8.2.4 starting on page 337 are derived for $\vec{\tau}$ and $\vec{\tau}_\omega$ by taking frequency derivatives analytically; numerical derivatives are therefore not necessary.

This appendix gives a vectorized code fragment written in Matlab which can be used as a core calculator for larger programs. Given the particular vectorization that follows, the code works well when there are more frequency evaluations than birefringent segments.

The differential-group delay $|\vec{\tau}|$, magnitude second-order PMD $|\vec{\tau}_\omega|$, and polarization-dependent chromatic dispersion $|\vec{\tau}|_\omega$ scalar spectra are calculated for each frequency ω by

$$|\tau|^2 = \vec{\tau} \cdot \vec{\tau} \quad (\text{C.1a})$$

$$|\tau_\omega|^2 = \vec{\tau}_\omega \cdot \vec{\tau}_\omega \quad (\text{C.1b})$$

$$|\tau|_\omega = \frac{\vec{\tau} \cdot \vec{\tau}_\omega}{\sqrt{\vec{\tau} \cdot \vec{\tau}}} \quad (\text{C.1c})$$

The output and input PSP vector spectra are

$$\hat{p}_{\text{out}} = \vec{\tau} / \tau \quad (\text{C.2a})$$

$$\hat{p}_{\text{in}} = R^\dagger \hat{p}_{\text{out}} \quad (\text{C.2b})$$

where $R = R_N R_{N-1} \dots R_1$. Each rotation operator is expanded in vector form as

$$R_k = I \cos(\omega \tau_k) + (1 - \cos(\omega \tau_k)) (\hat{r}_k \hat{r}_k \cdot) + \sin(\omega \tau_k) (\hat{r}_k \times) \quad (\text{C.3})$$

where τ_k is the DGD of a single birefringent segment and $\hat{r}_k$ is the Stokes direction of its birefringent axis. The concatenation equations (8.2.34) on page 337 are used to compute the cumulative first- and second-order PMD vector.

```

function [tau2, tauw2, pdcd, PSPout, PSPin] = CalcPMDSpec_1(w_vec, r_vec, tau_vec, phz_vec)

%
%      Inputs:
%      /w_vec/      (Trad/s) 1 x wlen vector of radial frequency range
5 %      /r_vec/      (scalar) 3 x Nseg matrix, each column is a unit Stokes vector of tau.k
%      /tau_vec/     (ps)     1 x Nseg vector of DGD for each segment
%                               (not to be confused with the PMD vector tau)
%      /phz_vec/     (rad)    1 x Nseg vector of residual birefringent phase
%                               for each segment.
10 %
%      Outputs:
%      /tau2/        (ps^2)    1 x wlen vector of DGD^2(w)
%      /tauw2/       (ps^4)    1 x wlen vector of SOPMD^2(w)
%      /pdcd/        (ps^2)    1 x wlen vector of PDCD(w)
15 %      /PSPout/     (scalar)   3 x wlen matrix of output PSP Stokes vectors
%      /PSPin/      (scalar)   3 x wlen matrix of input PSP Stokes vectors

% Defs
DEG2RAD = pi / 180; RAD2DEG = 180 / pi;
20 I2 = diag([1,1]);
I3 = diag([1,1,1]);

% Input-specific Defs
wlen = length(w_vec);
25 Nseg = length(tau_vec);

% Calculate rrdot and rcross for each segment up front
% Note: rrdot and rcross matrix has the following structure:
%
30 %      rrdot_vec = [rrdot(1) rrdot(2) ... rrdot(Nseg)]
%                      ^
%                      |
%                      im
%
% where each rrdot is a 3x3 matrix itself. Same with rcross.
35 %
% Calculate rrdot and rcross for each segment
for k = 1: Nseg,
40
    im = 3 * (k - 1) + 1; % column index into rrdot_vec
    rrdot_vec(:, [0:2]+im) = r_vec(:,k) * r_vec(:,k)'; % rrdot from dyadic

    % rcross is sum over cross-products of r_vec w/ S1, S2, and S3
    for i = 1: 3,
45         rcross_vec(:, (i-1)+im) = cross(r_vec(:,k), I3(:,i));
    end
end

50 % Precalculate the trig tables, row -> segment #; column -> freq
coswt = cos(tau_vec' * w_vec + phz_vec' * ones(size(w_vec)));
sinwt = sin(tau_vec' * w_vec + phz_vec' * ones(size(w_vec)));

% Define the tau vectors
55 for k = 1: Nseg,
    tauvec(:,k) = tau_vec(k) * r_vec(:,k);
end

```



```

60 % Now calculate the frequency response
    for iw = 1: wlen,

        % Set the frequency for the concat
        w = w_vec(iw);

65         % Initialize cumulative tau and tauw vectors.
        tau_cat = tauvec(:, 1);           % tau(1) = tau_1
        tauw_cat = zeros(size(tauvec(:, 1))); % tauw(1) = 0;

70         % Initialize cumulative R operator
        R_cat = I3;

        % We need R1 to find PSPin
        Rseg = coswt(1, iw) * I3 + ...
75         (1-coswt(1, iw)) * rrdot_vec(:, [0:2]+1) + ...
            sinwt(1, iw) * rcross_vec(:, [0:2]+1);

        % Make first concatenation
        R_cat = Rseg * R_cat;

80         % Accumulate tau, tauw, and Rseg through each segment
        for iseg = 2: Nseg,

            % column index into rrdot_vec and rcross_vec
85             im = 3 * (iseg - 1) + 1;

            % Construct R_iseg(w)
            Rseg = coswt(iseg, iw) * I3 + ...
                (1-coswt(iseg, iw)) * rrdot_vec(:, [0:2]+im) + ...
90             sinwt(iseg, iw) * rcross_vec(:, [0:2]+im);

            % Accumulate R
            R_cat = Rseg * R_cat;

95             % Accumulate tau_cat
            tau_cat = tauvec(:, iseg) + Rseg * tau_cat;

            % Accumulate tauw_cat
            tauw_cat = cross( tauvec(:, iseg), tau_cat ) + Rseg * tauw_cat;

100             % Accumulate tauw_cat

        end

        % Calculate the input PMD vectors tau and tauw
        Radj = conj( transpose( R_cat ) );
105         tau_in = Radj * tau_cat;
        tauw_in = Radj * tauw_cat;

        % Calculate scalar spectra (could do this outside the loop, too)
        tau2(iw) = tau_cat' * tau_cat;
110         tauw2(iw) = tauw_cat' * tauw_cat;
        dgd = sqrt(tau2(iw));
        pdcd(iw) = tau_cat' * tauw_cat / dgd;

        % Calculate the vector spectra
115         PSPout(:, iw) = tau_cat / dgd;
        PSPin(:, iw) = tau_in / dgd;

    end

```

The point-of-view of the preceding code is that operators $\hat{r}_k \hat{r}_k^\cdot$ and $\hat{r}_k \times$ as well as the $\omega \tau_k$ product can be evaluated outside the main loop. In this way the core loop is mainly a multiply-and-accumulate register.

After an initial setup, the operators $\hat{r}_k \hat{r}_k \cdot$ and $\hat{r}_k \times$ are evaluated for each PMD segment in the loop between **lines 39–48**.

$$\begin{aligned} \text{line 42:} \quad & (\hat{r}_k \hat{r}_k \cdot) = \hat{r}_k \hat{r}_k^T \\ \text{lines 45–47:} \quad & (\hat{r}_k \times) = \hat{r}_k \times \hat{s}_1 + \hat{r}_k \times \hat{s}_2 + \hat{r}_k \times \hat{s}_3 \end{aligned}$$

Matrices **rrdot_vec** and **rcross_vec** store the 3×3 operator associated with the k^{th} segment in a $3 \times 3k$ matrix that is indexed as a row vector on k .

The sine and cosine of the $\omega \tau_k$ product are computed before the concatenation loop. These calculations are stored in tables **coswt** and **sinwt** on **lines 51–52**. There is an important point that needs to be highlighted. Strictly speaking, the birefringent phase of a segment is $\omega \tau_k$. The radial frequency can certainly be used, such as $(2\pi)194.1$ THz. As an alternative, the birefringent phase of a segment is written $(\omega - \omega_o)\tau_k + \phi_k$, where ω_o is an arbitrary frequency and ϕ_k is a measure of the residual birefringent phase at ω_o . This form is useful when investigating the role of the birefringent phase on a PMD spectrum. The trigonometric terms on **lines 51–52** provide for a vector of residual birefringent phases that are added to $\omega \tau_k$, which if the vector is non-zero should be interpreted as $(\omega - \omega_o)\tau_k + \phi_k$.

Finally, the segment PMD vectors $\vec{\tau}_k$ are calculated in advance:

$$\text{lines 55–57:} \quad \vec{\tau}_k = \tau_k \hat{r}_k$$

The main frequency loop runs from **lines 60–118**. For each frequency the respective **coswt** and **sinwt** values are recalled, the R_k operators are constructed, vectors $\vec{\tau}$ and $\vec{\tau}_\omega$ are calculated, and the scalar and vector PMD spectra are computed and stored. The vectors $\vec{\tau}$ and $\vec{\tau}_\omega$ are generated by the nested loop that runs from **lines 82–101**; this loop runs the concatenation equations (8.2.34) on page 337. The inner loop is initialized with

$$\text{lines 67–68:} \quad \vec{\tau}(1) = \vec{\tau}_1, \quad \text{and} \quad \vec{\tau}_\omega(1) = 0$$

and **line 71**: $R = I$. Each iteration of the accumulation loop generates the PMD vectors from

$$\begin{aligned} \text{line 96:} \quad & \vec{\tau}(k) = \vec{\tau}_k + R_k \vec{\tau}(k-1) \\ \text{line 99:} \quad & \vec{\tau}_\omega(k) = \vec{\tau}_k \times \vec{\tau}(k) + R_k \vec{\tau}_\omega(k-1) \end{aligned}$$

Note that the running product of **line 93**: $R(k) = R_k R(k-1)$ is recorded. This operator is used to find the input PSP's from the output PSP's. In particular,

$$\text{lines 104–106:} \quad \vec{\tau}_s = R^\dagger \vec{\tau}_t, \quad \text{and} \quad \vec{\tau}_{s\omega} = R^\dagger \vec{\tau}_{t\omega}$$

With these preliminary calculations in place, the vector and scalar PMD spectra are computed on **lines 109–116**, following (C.1–C.2).

Figures C.1–C.2 are calculated for four equal-length stages using the above code fragment. The input and output PSP vector spectra are shown as are the DGD, magnitude SOPMD, and PDCD scalar spectra. The DGD spectra in Fig. 8.33 on page 362 were calculated in the same way.

Not included in the code but easily added is the calculation of $U(\omega)$. Direct calculation of $U(\omega)$ is ideal due to the difficulty extracting U from $jU_\omega U^\dagger$. While calculating $U(\omega)$ one should concurrently calculate $U_\omega(\omega)$ so that $jU_\omega U^\dagger$ can be checked against $\vec{\tau} \cdot \vec{\sigma}$, the latter being calculated from concatenation rules on $\vec{\tau}$ by R as above. The product rule for $U(\omega)$ is trivial:

$$U(N) = U_N U_{N-1} \dots U_1 \quad (\text{C.4})$$

The frequency derivative is calculated analytically and accumulated using a recurrence relation. Matrices U and U_ω are expressed as

$$U_k = I \cos(\omega\tau_k/2) - j(\vec{\tau}_k \cdot \vec{\sigma}) \sin(\omega\tau_k/2) \quad (\text{C.5a})$$

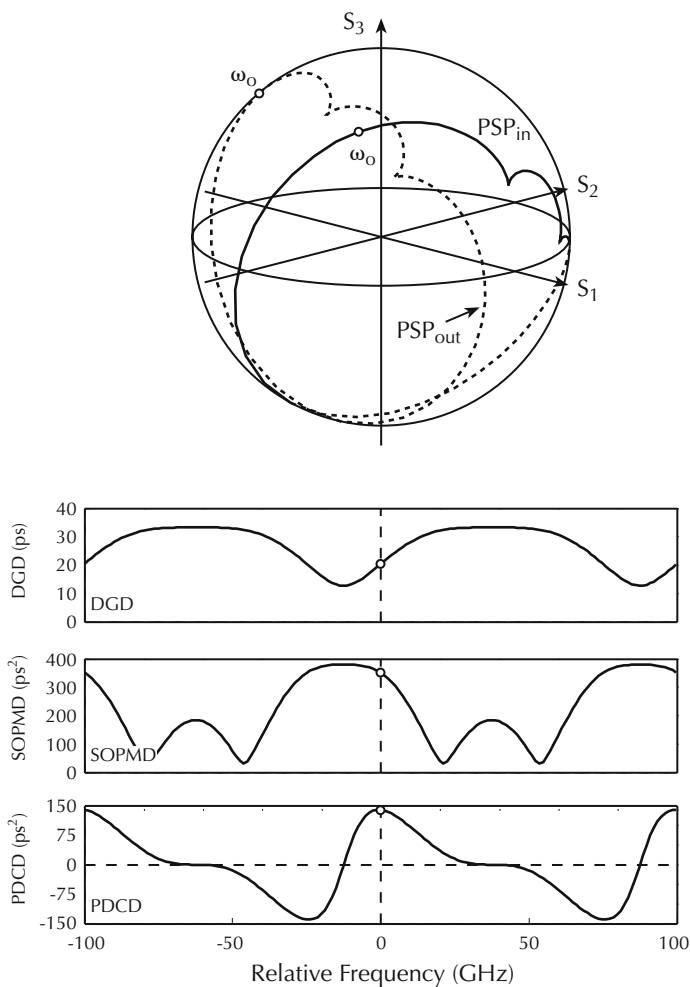
$$U_{\omega k} = -\tau_k/2 (I \sin(\omega\tau_k/2) + j(\vec{\tau}_k \cdot \vec{\sigma}) \cos(\omega\tau_k/2)) \quad (\text{C.5b})$$

As with R_k , $\vec{\tau}_k \cdot \vec{\sigma}$, $\sin(\omega\tau_k/2)$, and $\cos(\omega\tau_k/2)$ can be calculated in advance of the frequency loop. The recurrence relation for $U_\omega(k)$ is

$$U_\omega(k) = U_{\omega k} U(k-1) + U_k U_\omega(k-1) \quad (\text{C.6})$$

A quick test to verify that U and U_ω are correctly calculated is to check that $jU_\omega U^\dagger$ is Hermitian.

Calculation of $U(\omega)$ is useful, for instance, when calculating the time response of a signal that transits a PMD medium. The polarization transfer matrix in frequency and time domains are given by (8.2.41) and (8.2.42) on page 343.



$$\rightarrow \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \tau \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \tau + \phi \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \tau + \phi \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \tau \rightarrow \vec{\tau}$$

Fig. C.1. Vector and scalar spectra for four birefringent sections: $\tau = \{10, 10, 10, 10\}$ ps, $\phi = \{0, 45, 45, 0\}^\circ$, $\hat{r} = \{0, -45, -90, -135\}^\circ \times 1.5$ lying on the equator. The center frequency ω_o is indicated on both vector and scalar plots. The period of the scalar spectra is 100 GHz and the spectra have been shifted by one-eighth period.

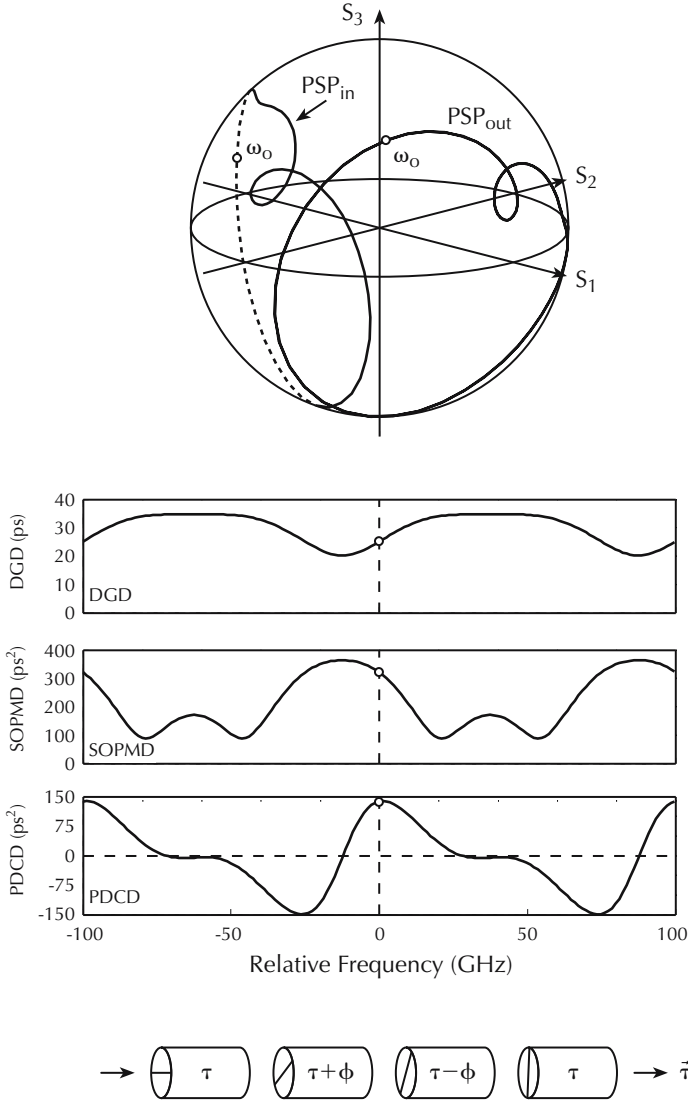


Fig. C.2. Vector and scalar spectra for four birefringent sections: $\tau = \{10, 10, 10, 10\}$ ps, $\phi = \{0, 22.5, 67.5, 0\}^\circ$, $\hat{r} = \{0, -45, -90, -135\}^\circ \times 1.25$ lying on the equator. The center frequency ω_o is indicated on both vector and scalar plots. The differential phase shift of 22.5° in the center sections about the common phase shift 45° distorts the PMD spectra.

D

Multidimensional Gaussian Deviates

Consider the gaussian random variable X . The probability density is

$$\rho_X(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \quad (\text{D.1})$$

The expectation and variance are

$$E[X] = 0, \quad \text{and} \quad \text{var}(X) = \sigma_x^2 \quad (\text{D.2})$$

Consider now a two-dimensional distribution composed of two independent identically distributed (i.i.d.) gaussian random variables (g.r.v.); denote the two deviates X_1 and X_2 , and a vector defined as $\mathbf{X} = (X_1, X_2)$. While we may be interested in the distribution of these cartesian components, an alternative is the distribution of the corresponding polar coordinates. A polar deviate is defined as $\mathbf{P} = (R, \theta)$. A one-to-one map g relates the two coordinates such that

$$\begin{aligned} g(x_1, x_2) &= (r, \theta) \\ &= (\sqrt{x_1^2 + x_2^2}, \tan^{-1} x_2/x_1) \end{aligned}$$

The inverse map $h = g^{-1}$ given in polar form is

$$\begin{aligned} h(r, \theta) &= (x_1, x_2) \\ &= (r \cos \theta, r \sin \theta) \end{aligned}$$

The joint density of the polar coordinates is related to the joint density of the cartesian coordinates through the Jacobian:

$$\rho_{\mathbf{P}}(r, \theta) = \rho_{\mathbf{X}}(h(r, \theta)) J_h \quad (\text{D.3})$$

where

$$J_h = \begin{vmatrix} \frac{\partial h_1}{\partial r} & \frac{\partial h_1}{\partial \theta} \\ \frac{\partial h_2}{\partial r} & \frac{\partial h_2}{\partial \theta} \end{vmatrix} \quad (\text{D.4})$$

In the present case, $J_h = r$. The polar joint distribution is therefore

$$\rho_{\mathbf{P}}(r, \theta) = \frac{r}{2\pi\sigma_x^2} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)$$

where the argument of the exponential is $x_1^2 + x_2^2 = r^2(\cos^2 \theta + \sin^2 \theta)$. Now, the random variables R and θ are independent, so the joint distribution is the product of the two individual distributions. The angular distribution is uniform over 2π , so the product is written as

$$\rho_\theta(\theta)\rho_R(r) = \left(\frac{1}{2\pi}\right) \left(\frac{r}{\sigma_x^2} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)\right)$$

The resultant radial distribution, known as the Rayleigh distribution, is

$$\rho_R(r) = \frac{r}{\sigma_x^2} \exp\left(-\frac{r^2}{2\sigma_x^2}\right), \quad r \geq 0 \quad (\text{D.5})$$

The moments of the Rayleigh distribution are

$$E[\rho_R^n(r)] = 2^{n/2} \sigma_x^n \Gamma\left(1 + \frac{n}{2}\right), \quad n \in \mathbb{Z}$$

where $\mathbb{Z}$ is the set of integers greater or equal to zero. Denoting the n^{th} moment as $\langle r^n \rangle$ and $\text{var}(r) = \sigma_r^2$, the basic Rayleigh distribution parameters are

$$\langle r \rangle = \sqrt{\frac{\pi}{2}} \sigma_x, \quad \langle r^2 \rangle = 2\sigma_x^2 \quad (\text{D.6a})$$

$$\sigma_r^2 = \left(2 - \frac{\pi}{2}\right) \sigma_x^2 \quad (\text{D.6b})$$

Note in particular the relation between the first and second moments:

$$\langle r^2 \rangle = \frac{4}{\pi} \langle r \rangle^2 \quad (\text{D.7})$$

Next consider the three-dimensional distribution of three i.i.d. gaussian random variables $\mathbf{X} = (X_1, X_2, X_3)$, each with variance σ_x^2 , and its polar equivalent $\mathbf{P} = (R, \theta, \phi)$. In the polar coordinate system, $\theta \in [0, \pi]$ is the declination angle from X_3 and $\phi \in [-\pi, \pi]$ is the azimuth angle. The polar to cartesian transformation h is

$$\begin{aligned} h(r, \theta, \phi) &= (x_1, x_2, x_3) \\ &= (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) \end{aligned}$$

Table D.1. Key Relations for Multivariate Gaussian Distributions

Distribution	$\rho_R(r)$	$\langle r \rangle$	$\langle r^2 \rangle$	$\text{var}(r)$	ratio
Gaussian ^(a)	$\frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)$	0	σ_x^2	σ_x^2	
Rayleigh ^(b)	$\frac{r}{\sigma_x^2} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)$	$\sqrt{\frac{\pi}{2}} \sigma_x$	$2\sigma_x^2$	$\left(2 - \frac{\pi}{2}\right) \sigma_x^2$	$\langle r^2 \rangle = \frac{4}{\pi} \langle r \rangle^2$
Maxwellian ^(b)	$\sqrt{\frac{2}{\pi}} \frac{r^2}{\sigma_x^3} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)$	$\sqrt{\frac{8}{\pi}} \sigma_x$	$3\sigma_x^2$	$\left(3 - \frac{8}{\pi}\right) \sigma_x^2$	$\langle r^2 \rangle = \frac{3\pi}{8} \langle r \rangle^2$
^(a) $r \in (-\infty, \infty)$, ^(b) $r \in [0, \infty)$					

The corresponding Jacobian is

$$J_h = r \sin \theta$$

The polar joint distribution $\rho_{\mathbf{P}}(R, \theta, \phi)$, written as the product of three independent polar random variables, is

$$\rho_{\phi}(\phi) \rho_{\theta}(\theta) \rho_R(r) = \left(\frac{1}{2\pi}\right) \left(\frac{\sin \theta}{2}\right) \left(\sqrt{\frac{\pi}{2}} \frac{r^2}{\sigma_x^3} \exp\left(-\frac{r^2}{2\sigma_x^2}\right)\right)$$

The resultant radial distribution, known at the Maxwellian distribution, is

$$\rho_R(r) = \sqrt{\frac{\pi}{2}} \frac{r^2}{\sigma_x^3} \exp\left(-\frac{r^2}{2\sigma_x^2}\right) \quad (\text{D.8})$$

The moments of the Maxwellian distribution are

$$E[\rho_R^n(r)] = \frac{2}{\sqrt{\pi}} 2^{n/2} \sigma_x^n \Gamma\left(\frac{3+n}{2}\right)$$

Therefore, the basic parameters of the Maxwellian distribution are

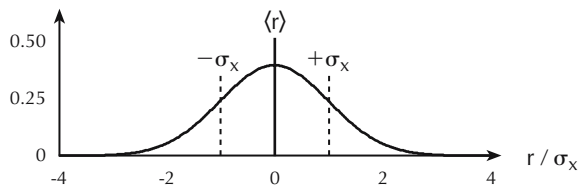
$$\langle r \rangle = \sqrt{\frac{8}{\pi}} \sigma_x, \quad \langle r^2 \rangle = 3\sigma_x^2 \quad (\text{D.9a})$$

$$\sigma_m^2 = \left(3 - \frac{8}{\pi}\right) \sigma_x^2 \quad (\text{D.9b})$$

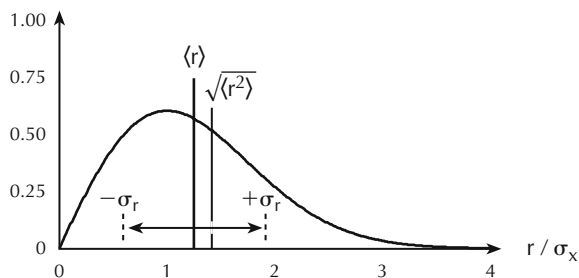
The relation between the first and second moments is

$$\langle r^2 \rangle = \frac{3\pi}{8} \langle r \rangle^2 \quad (\text{D.10})$$

Gaussian



Rayleigh



Maxwellian

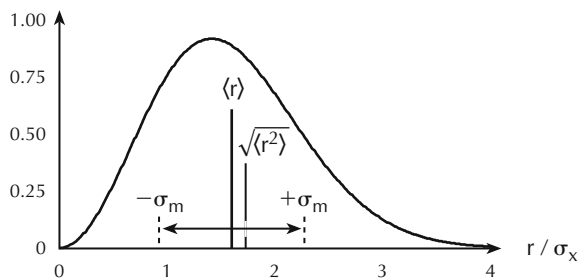


Fig. D.1. Probability densities for Gaussian, Rayleigh, and Maxwellian distributions. The gaussian distribution is symmetric about the origin while the Rayleigh and Maxwellian distributions, associated with the radius of a circle and sphere, respectively, are one-sided with $r \geq 0$. All distributions are completely determined by the component variance σ_x^2 .

Index

- Abbe number 94
- α -BBO 150, 176
 - group-index temp. co. 170
 - material properties 148
 - temperature compensation 173
- ABCD matrices
 - from q-transformation 224
 - GRIN lens 231
 - optimal lens coupling 241
 - plane-wave limit 227
- achromats
 - Koester 184
 - MgF/quartz 184
 - Pancharatnam 186
 - Shirasaki 189
- Ampère's law 2, 82
- anisotropic media 85, 136, 139, *see*
 - birefringent media
- attractor-processor method (APM) 436, 446
- autocorrelation bandwidth
 - PMD 410
- autocorrelation function *see* DGD
 - autocorrelation function, *see* PMD
 - vector
 - connection between ensemble and frequency averages 408
 - derivations 411
 - mean-square DGD 409
 - PMD vector 409
- Becquerel formula 133
- bi-circulator 294
- bi-isolator 294
- Bi:RIG *see* iron garnet
- bianisotropic media 85, 136, *see*
 - optical activity
- biaxial crystal 105
- Biot's law 141, 149
- birefringent beat length 121, 179, 389
 - of crystalline quartz 150
 - of SMF fiber 385, 450
 - of YVO₄ 465
- birefringent crystal
 - effective index 109, 268
 - high and low birefringence 147
 - positive and negative uniaxial 106
 - Poynting vector direction 109
 - properties 148
 - refraction 112
 - susceptibility tensor 106
 - temperature dependence 163, 170
 - walkoff angle 110
 - waveplate cut 116, 120
- birefringent media *see* fiber birefringence
 - constitutive relation 107
 - ordinary and extraordinary axes 107
- birefringent phase 72, 120, 130, 179, 329, 467, *see* residual birefringent phase
 - control of 185
 - Evans phase shifter 464
 - frequency dependence 182
 - relation to DGD value 122
 - relation to PMD spectrum 474

- temperature compensation 172
- temperature dependence 166, 171
- birefringent walkoff *see* walkoff angle, *see* walkoff block
- compensation 174
- crystal cut 116, 202
- effective index 117
- total internal reflection 118
- birefringent wedge *see* prism
- BK7 glass 159, 161, 184, 464
 - material properties 147
- bra and ket vectors
 - duality 40
- bracket notation 39
- Brewster's angle 100, 102
 - birefringent separation 118, 204, 275
- Brownian motion 394
 - density function 390
 - Karhunen-Loève expansion 419
 - sample paths 392, 421
- C-band 144
- calcite 274, 278
 - group-index temp. co. 170
 - material properties 148
 - temperature compensation 173
- Cayley-Klein unitary matrix 51, 332
- characteristic admittance 5
- characteristic impedance 5
- chiral media 80, 135, 150
 - constitutive relation 138
 - optic fiber 389, 422
 - wire model 136
- chirality parameter 138
- circular polarization 15, 35, 53, 430
 - and Fresnel rhomb 196
 - eigenstates 129, 139
- circulators *see* bi-circulator
 - classification 273
 - deflection type 285
 - Kaifa type 286
 - performance specs 294
 - Shirasaki-Cao type 290
 - Xie-Huang type 286, 292
 - displacement type
 - ladder type 282
 - quartz-free 283
 - strict-sense type 281
 - historical examples 277
 - polarization dependent 274
- coherency matrix 23, 35
- coherent PMD 465
- collimator *see* dual-fiber collimator, *see* fiber-to-fiber coupling
 - assemblies 214
 - air gap 216
 - epoxy joint 213
 - fused joint 217
- C-lens example 236
- C-lens type 212
- comparison chart 219
- design goals 213
- GRIN-lens example 236
- GRIN-lens type 212
- pointing direction 217, 230, 237
- complete gap 284
- component DGD *see* residual
 - birefringent phase
 - circulator 283, 292
 - delay crystals 172, 177, 456
 - isolator 258, 263, 266, 268, 270
 - Kaifa prism 204
 - Rochon prism 201
 - Wollaston prism 201
- component PDL
 - circulator 277, 282, 294
 - isolator 259, 263
 - off-axis delay crystals 174
- confocal parameter 222, 225, 262
- conservation of energy 6, 138, *see* Poynting's theorem
- constitutive relations 3, 85, 90
 - birefringent 107
 - chiral media 138
 - Drude-Born-Fedorov model 138
 - gyrotropic materials 126
 - isotropic 94
 - losslessness 86
 - optically active media 138
- coupling coefficient 242
- critical angle 101, 118, 198
- crystal classes 107
- crystalline quartz 147, 179, 455, 464
 - and MgF₂ achromat 184
 - material properties 149
- Curie temperature *see* iron garnet
- current density 2, 7

- data folding 443
- degree of polarization (DOP) 22, *see* repolarization
 - from coherency matrix 23
 - from intensity 34
 - from Stokes parameters 23
 - PDL surfaces 306
- depolarization
 - connection to partial polarization 23, 31, 324
 - connection to PDL 306
 - density conditional on DGD 404
 - effect on pulse 346, 356
 - entangled states 344
 - probability density 402
 - programmable generation 454
 - relation to mean-square SOPMD 404
 - relation to PMD autocorrelation 411
 - relation to second-order PMD 323
- depth of focus 223, 239, 262
- DGD *see* component DGD, *see* PMD and impulse response 325, 359
 - anomalous *see* PMD and PDL combined
 - impulse response verses input polarization state 354
 - relationship to birefringent phase 122
- DGD autocorrelation function 410
- DGD component of PMD
 - length of PMD vector 314, 330, 333
- DGD in fiber
 - diffusions limits 398
 - examples 324, 407
- DGD measurement 443, 445
- DGD spectrum 321, 443, 461, *see* coherent PMD
 - effect of spectrum 322
- DGD statistics *see* mean fiber DGD
 - Maxwellian density 402
 - mean-square equation of motion 399
 - mean-square growth 397
- diamagnetic media 123
 - electron equation of motion 124
 - susceptivity tensor 125
- differential attenuation slope (DAS) 371, 374, 447, 449, *see* PMD and PDL combined
- differential-group delay *see* DGD
- diffraction angle 223
- diffusion equation
 - PDL 422
 - PMD 399
 - SOP 393
- diffusion process 388
- dispersion relation 4, 92, 95, 108, 128, 156
- Drude's equation 141
- Drude-Born-Fedorov model 138
- dual-fiber collimator 198, 201
 - divergence angle 238
 - example 238
 - for circulator 284, 290
 - for polarization-beam splitter 285
- DWDM channel spacing 143
 - filter tolerancing 145, 159
- ECHO source 454, *see* coherent PMD
 - birefringent phase control 465
 - calibration 464
 - common and differential phase control 475
 - comparison with JPFD and PMDS 477
 - design criteria 471
 - frequency shift of PMD spectrum 475
 - independent control of 1st and 2nd order PMD 473
 - instrument bandwidth 478
- eigenstates 46
- electric charge density 2
- electric dipole moment 80
- electric field 2, 12
- electric-flux density 81
 - continuity condition 83
 - including media interaction 91
- electromagnetic dissipation 7
- electromagnetic stored energy 6, 87
- electron equation of motion 90, 105, 124
- elliptical polarization 15, 35, 127
- epoxy
 - heat cure 213

- UV cure 215
- Euler rotations 71
- evanescent field 102, 196
 - penetration depth 104
- Evans phase shifter 184, 464, 475
- evolution equation *see* diffusion equation
 - PDL 310, 380
 - PMD 339, 380
 - PMD and PDL 377, 380
 - SOP 339, 380
- extinction coefficient
 - from permittivity 92
- extraordinary axis *see* birefringent media
- Fabry-Perot interferometer 154, 163, 215
 - frequency response 157
 - temperature dependence 161
- Faraday angle 132, *see* specific rotation
- Faraday rotation 129, 132, 197, 251, 493, *see* nonreciprocal polarization rotation
 - comparison to optical activity 141
 - operator expression 207
- Faraday rotator 189, *see* Shirasaki achromat
 - for circulators 273, 277, 286
 - for isolators 247, 255, 259
 - garnet 135, 150
 - linear 197, 207
- Faraday's law 2, 82
- ferrimagnetic garnet 123, 133, 150, *see* iron garnet
- ferrule 213, 232, 285, 290
 - tilt angle 215, 234
- fiber autocorrelation length 385, 390, 395, 398
 - in relation to birefringence beat length 396
- fiber birefringence 385, 392
 - chirality 450
 - length scales 387
 - no chirality 389
 - origins 386
 - random birefringence model 391
 - random orientation model 389
- fiber-to-fiber coupling 239, 261
 - optimal coupling 240
- first and second order PMD
 - independent control of 473
- first-order PMD *see* component DGD, *see* DGD
- focus error 244
- four-states method
 - combined PMD and PDL measurement 449
 - PDL measurement 432
 - PMD measurement 437, 447
- free-spectral range
 - and group index 158
 - birefringent 121
 - Fabry-Perot 158, 165
 - PMD 336, 366, 460, 478
 - temperature shift 168
- Fresnel rhomb 196
- Frigo equation 372, 446
- fused silica 183, 218
 - material properties 147
- Gauss' electric law 2, 80
- Gauss' magnetic law 2, 82
- gaussian distribution 505
- gaussian optics 219
 - beam waist 222
 - confocal parameter 222
 - diffraction angle 223
- generator function 394, 399, 422
 - Stratonovich translation 394
- Gires-Tournois interferometer 154
 - frequency response 161
- Glan-Taylor prism 204, 274, 278, 280
- Glan-Thompson prism 274, 278
- Goos-Hänchen displacement 104
- Goos-Hänchen phase shift 102
- GRIN lens 211, 215, 237
 - ABCD matrix 231
 - index gradient constant 232
 - index profile 230
 - melt point 218
 - pitch 232
 - polish angle 217
- group delay
 - GT interferometer 161
 - PMD 329
- group index 93, 121, 158, 263

in fiber 450
 temperature dependence 163, 170
 group velocity 93, 163
 birefringent media 112
 gyrotropic angle 127
 gyrotropic media
 constitutive relation 126
 eigenvector orientation 127
 nonreciprocal polarization rotation 129
 permittivity tensor 122
 precession angle 130
 half-wave waveplate 179, 184, 254, 276, 279, 286, 434, 455, 464
 achromat 184
 bandwidth 182
 operator expression 207
 polarization control 193
 Helmholtz equation 3, 91, 388
 Hermite coefficients 411
 Hermitian matrix *see* Mueller matrix
 relation to PMD 332
 Hermitian operator 47, *see* skew-Hermitian operator
 relation to PMD 327
 relation to unitary 49
 spin-operator form 62
 spin-vector form 61
 impermeability 86
 permittivity 86, 107, 126
 importance sampling 405
 indicatrix 110, 116
 Poynting vector 112
 inner product 41
 interferometric (INT) method 436, 439
 invar 153
 iron garnet 150, 493, *see* Faraday rotator
 Bi:RIG 151, 248, 254, 279
 Curie temperature 124, 152
 design goals 151
 dual 252
 hysteresis 134, 152
 latching 134, 153, 284, 290
 saturation 134
 YIG 151, 254, 276

isolators *see* bi-isolator
 deflection type 254
 isolation 257
 PMD 258, 263
 ray-trace 256, 267
 technology comparison 259
 displacement type
 isolation 262
 PMD 263
 ray-trace 261, 265
 insertion loss 249
 isolation definition 250
 lens systems 253
 PMD compensated 266
 polarization-dependent 247
 polarization-independent 254, 259
 return loss 258
 temperature dependence 252
 tolerancing 249
 two stage 263
 wavelength dependence 251
 isomorphism 65
 isotropic media
 electron equation of motion 90
 propagation in 95
 reflection coefficient 98, 99
 refractive index 94
 susceptibility 91
 joint probability distribution of PMD 453, 477
 scales with mean fiber DGD 405
 Jones matrix 18, 45, *see* Hermitian matrix, *see* PMD operator, *see* unitary matrix
 from Stokes parameters 19
 on-axis PDL 302
 relation to Mueller matrix 19, 38
 spin-matrix form 61
 Jones matrix eigenanalysis (JME)
 436, 442, *see* data folding
 Heffner eigenvalue equation for PMD 442
 LabView code 442
 step size 444
 Jones to Stokes 56
 Jones vector 13, 52
 from Stokes parameters 18

- Kaifa circulator 286
- Kaifa prism 202, 284, 285
- Karhunen-Loève expansion 419
- kDB system 87
 - birefringent materials 108
 - defining coupled equations 89
 - Faraday rotation 129
 - gyrotropic materials 126
 - isotropic materials 95
 - optically active media 139
- Kolmogorov's backward equation 394
- kovar 153
- $\lambda/4, \lambda/4$ combination 192
- $\lambda/2, \lambda/4$ combination 194
- $\lambda/4, \lambda/2, \lambda/4$ combination 195
- L-band 144
- Lagrange multiplier method 433
- Langevin process 391
- lead molybdate (PbMoO_4) 149
- lens classification 211
- lens equation, simple 229
- linear polarization 15, 35
 - eigenstate 110
 - gyrotropic rotation 136, 141
- LiNbO_3 150, 185, 258, 456, 464
 - group-index temp. co. 170
 - material properties 148
 - temperature compensation 173, 177
- local birefringence vector 329, 339, 392, 397
- Lorentz force 90, 125
- Lorentz gauge 9
- Lytot depolarizer 31, 298
- magnesium fluoride (MgF_2) 149, 183
- magnetic dipole moment 82, 123
- magnetic field 2, 81, 90
- magnetic flux density 82
- magnetic material types 123
- magnetization density vector 2, 82
- magnification error 242
- magnification, lens 229, 233, 239, 262
- Maxwell's Equations 137, 493
 - complete form 2
 - in terms of D and B 85
 - in vacuum 8
 - interaction with media 84
 - time-harmonic form 11
- Maxwellian distribution
 - derivation of 506
- Maxwellian distribution of DGD 400, 417
- Maxwellian distribution of PDL 423
- mean fiber DGD
 - connection to waveplate model 417
 - def as statistical "unit" 400
 - measurement 436, *see* interferometric (INT) method, *see* wavelength-scanning (WS) method
 - measurement uncertainty 409, 414
 - relation to PMD vector measurement 444
- mean outage duration 478
- mean outage rate 478
- mean-reverting process *see* Langevin process
- mean-square DGD
 - autocorrelation function 414
 - relation to mean fiber DGD 401
 - relation to pulse broadening 363
- mean-square SOPMD
 - relation to mean fiber DGD 401
- modulation phase-shift (MPS) method 436, 447
 - for combined PMD and PDL 449
- Mueller matrix 23, 37, 449, *see*
 - four-states method
 - and trace of Hermitian 298
 - comparison between unitary and Hermitian 19, 38
 - comparison between unitary and traceless Hermitian 327
 - from Jones matrix 18, 66
 - on-axis PDL 304
 - PDL measurement 432
 - polarimeter 431
- Mueller matrix method (MMM) 436, 444
 - PDL tolerance 437
- natural light 22
- nonreciprocal polarization rotation 122, 131, 150, 247
- numerical aperture (N.A.) 211, 223, 292

- O(3) group 65, 327
- off-axis delay
 - effective index 177
- operators 44, 76
 - PMD 330
 - rotation
 - Jones form 67
 - Stokes form 68
- optical activity *see* chiral media
 - bi-isotropic 138
 - comparison to Faraday rotation 141
 - polarization rotation 140
 - reciprocal and nonreciprocal 139
- optical power 228, 230
- optically active material 85, 135, *see* tellurium dioxide (TeO_2)
- optically active rotator 189, 278, 282, 372
- ordinary axis *see* birefringent media
- orthogonal polarization states 59
 - and PDL 302
 - differential-group delay 326
- orthonormal basis 43
- outer product 42

- P.A.M. Dirac 40
- Pancharatnam achromat 186
- paraxial wave equation 220
- partial differential equation
 - connection to SDE 394
- partial polarization
 - coherent light 24
 - incoherent light 28
 - natural light 22
 - pseudo-depolarization 31
- Pasteur chirality parameter 139
- Pauli spin matrices 54
- Pauli spin operators 61
 - decomposition 61
 - exponential form 62
- Pauli spin vector *see* spin vector
- PDCD component
 - density conditional on DGD 404
 - probability density 402
 - relation to mean-square SOPMD 404
- PDL 297, *see* component PDL, *see* cumulative PDL vector, *see* repolarization
 - depolarized transmission 309
 - equation of motion 310
 - polarization transformation 304
 - polarization-state pulling 305, 310, 373
 - separation from PMD 378
 - symbol definitions 301
 - transmission coefficient 301
 - transmission surfaces 303
- PDL diffusion 422
- PDL measurement
 - four-states method 432
 - maximum discrepancy 435
 - six-states method 435
- PDL operator 300
- PDL statistics
 - probability density
 - Maxwellian approximation 424
 - precise 423
 - stochastic differential equation 422
- PDL value
 - connection between local and cumulative 302
 - in terms of cumulative PDL value 302
 - in terms of Mueller entries 433
 - in terms of transmission 299
- PDL vector
 - cumulative 301, 308
 - equation of motion 310, 377, 380
 - examples 311
 - local 300
- permeability 86, 91
 - free-space 2
- permittivity 86, 388
 - free-space 2
 - from susceptibility 91
- phase velocity 4, 93, 101, 120
 - in kDB 89, 95, 128, 129, 139
- phase-matching condition
 - birefringent media 113, 119
 - isotropic media 96, 101
- plane wave 4, 220, 227
 - polarization 12
 - time-harmonic form 11
 - vector form 12, 92
- PMD 297, *see* DGD, *see* mean fiber DGD, *see* PSP

- comparison to PMD and PDL
 - combined 377
 - frequency SOP evolution 319
 - historical development 312
 - how hard can it be? 312
 - in relation to polarization transformation 319
 - is not DGD 346
 - physical definition 314
 - separation from PDL 378
 - single section 315
 - spectral decomposition 320
 - two sections 320, 336, 459
- PMD and PDL combined 371, *see*
 - differential attenuation slope (DAS), *see* separation of PMD and PDL
 - anomalous pulse spreading 371
 - change in polarization state 373
 - equation of motion 377
 - Frigo equation 372
 - non-orthogonal PSPs 374
 - non-rigid precession 446
 - operator eigenvalue equation 374
 - spin-vector operator 376
- PMD concatenation rules
 - and Fourier analysis 365
 - first-order 337
 - including waveplates 468
 - second-order 337
- PMD diffusion
 - component probability density 402
 - Maxwellian density 402
- PMD emulator 451
- PMD evolution
 - equation of motion 380, 397
 - examples 336, 338, 341
- PMD Fourier content 364, *see*
 - coherent PMD
 - examples 367, 466, 467
 - generator function 370
 - phase shift due to mode mixing 369
- PMD impulse response 325, 358
 - connection with rms DGD 361
 - example 362
 - pulse broadening 363
- PMD measurement *see* data folding, *see* mean fiber DGD, *see* separation of PMD and PDL
 - classification 437
 - PDL tolerance 436
- PMD operator
 - eigenvalue equation 329
 - spin-vector form 330
 - traceless Hermitian 327
- PMD pulse distortion
 - distortion
 - first-order 345
 - moments analysis 352
 - second-order 347, 350, 351
 - vs. launch state 357
 - eye closure 363
 - field verse intensity response 440
 - inter-impulse interference 349
 - polarization transfer function 343
- PMD source 451
 - multi-state source
 - calibration 456
 - control 459
 - precision servo motors 455
 - temperature compensation 456
 - wavelength-flat states 458
- PMD spectrum
 - decomposition 320
 - efficient calculation of 497
 - examples 324, 407, 474
 - frequency shift 475
- PMD statistics 402
 - waveplate model 417
- PMD vector *see* PMD concatenation rules
 - as Stokes vector 321
 - autocorrelation function 409, 413
 - cartisian components 332
 - comparison between length and frequency increment 326
 - connection to PMD operator 330
 - governing eigenvalue equation 380
 - polarization precession in frequency 319
 - relation to DGD 319
 - relation to PSP 319
 - relation to unitary operator 333
 - statistical moments 402
 - stochastic differential equation 399
- PMD vector measurement *see*
 - attractor-processor method

- (APM), *see* Jones matrix eigen-analysis (JME), *see* modulation phase-shift (MPS) method, *see* Mueller matrix method (MMM), *see* Poincaré sphere analysis (PSA)
- Poincaré sphere
 - from Stokes parameters 20
- Poincaré sphere analysis (PSA) 436, 446
- polarimeter 430
 - fiber-grating type 432
 - for PDL measurement 432
 - for PMD measurement 443
- polarization beam splitter
 - prism comparison 285
- polarization control 191, *see* four-states method
 - arbitrary-to-arbitrary 192, 195
 - electro-optic 313
 - linear-to-arbitrary 194
- polarization decorrelation length 392
 - connection with fiber autocorrelation length 395
 - local and fixed frame 396
- polarization density vector 2, 80
 - birefringent media 105
 - chiral media 137
 - gyrotropic media 125
 - resonance expression 91
- polarization diffusion
 - short-range anisotropy 397
 - stochastic differential equation 393
- polarization ellipse
 - elliptical equation 13
- polarization retarders *see* Fresnel rhomb, *see* waveplate
- polarization state 39, *see* circular polarization, *see* elliptical polarization, *see* linear polarization, *see* orthogonal polarization states
 - conversion for this text 13, 53
 - measurement of 430
- polarization transfer function 343
- polarization vector *see* Jones vector, *see* Stokes vector
- polarization-dependent chromatic dispersion component *see* PD CD component
- polarization-dependent loss *see* PDL
- polarization-dependent optical frequency-domain reflectometry (P-OFDR) 450
- polarization-dependent optical time-domain reflectometry (P-OTDR) 450
- polarization-mode dispersion *see* PMD
- polarization-mode dispersion compensator 313
- polarization-state evolution *see* diffusion equation, *see* evolution equation
- polarization-state measurement *see* polarimeter
- polarization-state speed change 429
- polarizing wedge 117
- Poynting vector 6, 140
 - birefringent media 109, 114
 - gyrotropic media 128
 - isotropic media 95
 - relation to indicatrix 112
 - time averaged 8, 12
 - time-harmonic form 11
 - walk-off compensation 175
- Poynting's theorem 6, 86, *see* Poynting vector
 - time-harmonic form 11
- precession
 - about birefringent vector 69, 121, 131, 140, 316
 - about PMD vector 319, 331
 - birefringent and PMD comparison 326, 339
 - equation of motion 70
 - Evans phase shifter 184
 - non-rigid precession (PMD+PDL) 446
- precession angle *see* birefringent phase principal axis
 - Evans 184
 - Pancharatnam 186
- principal state of polarization *see* PSP
- prism *see* Kaifa prism, *see* Rochon prism, *see* Shirasaki prism, *see* Wollaston prism
 - birefringent 199
 - isotropic 198

- programmable PMD source *see* ECHO source, *see* PMD source
- projection matrix 16, 52
- projectors 42
 - spin-vector form 56
- pseudo-depolarization 31
- PSP *see* PMD
 - calculation 497
 - comparison to one-stage eigen-system 326
 - effect of spectrum 322
 - evolution 340
 - fiber spectrum 324
 - four-section spectrum 473
 - non-orthogonal *see* PMD and PDL combined
 - non-orthogonal overlap 375
 - pointing direction of PMD vector 330
 - stationary polarization transformation 318
 - two-section spectrum 320, 336, 443, 461
- PSP spectrum 321
- q-transformation 224, *see* ABCD matrices
- quarter-wave waveplate 179, 184, 189, 430, 434, 440, 464
 - operator expressions 207
 - polarization control 193
- $\hat{r}\hat{r}$ matrix 70
- $\hat{r}\times$ matrix 70
- Rayleigh distribution 392, 417
 - derivation of 505
- Rayleigh length *see* confocal parameter
- receiver map 479
- reciprocal polarization rotation 141
- reflection coefficient 96
- refractive index *see* Sellmeier equation
 - from permittivity 92
 - resonant model 93
- relative permittivity 91
- repolarization 424
 - surfaces 307
- residual birefringent phase 341, 465, 500
- Rochon prism 200, 273, 290
 - modified 201
- rotary power
 - Faraday rotation 133
 - optical activity 141
- rotation matrix
 - matrix form 67
- rotation operator *see* PMD concatenation rules
 - connection to PMD vector 332
 - connection to Stokes rotation 64, 207
- rutile (TiO_2) 147, 149, 206, 259, 279
- samarium-cobalt magnet 150, 496
- scalar potential 9, 224, 493
- scattering matrix
 - partially reflecting mirror 154
- second-order PMD
 - calculation 497
 - concatenation 337
 - decomposition 324
 - evolution equation 380
 - examples 407
 - generation 454, 459, 470
 - joint-probability density with DGD 406
 - probability density 402
 - pulse distortion 346, 356
 - pulse spreading 358
 - refs to Jones matrix form 364
 - relation between depolarization and PDCD densities 404
 - relation to PMD vector 323
 - simple impulse response 349
- Sellmeier equation 94, 141
- separation of PMD and PDL 437, 442, 446, 449
 - Jones theorem 378
- Shirasaki achromat 189
- Shirasaki circulator 279
- Shirasaki prism 204, 279
- Shirasaki-Cao circulator 290
- similarity transform 50, 327
- skew-Hermitian operator 61, 372
- SMF-28
 - birefringent beat length 385
 - effective index 215
 - mode-field diameter 223

- N.A. 211
- Snell's law 96, 115, 233
- specific rotation 135, 151
 - sensitivity 248
- spectral coverage 146, 182
- speed of light 4
- spin vector 55
 - identities 57
 - matrix form 61
 - PMD operator 330
- spun fiber 386, 389, 450
- stochastic differential equation 388
 - Ito and Stratonovich forms 393
- Stokes parameters
 - relation to ellipse 17, 430
- Stokes to Jones 56
- Stokes transformation
 - unitary 65
- Stokes vector 17, 35, 432, 442
 - from coherency matrix 23
 - from Jones vector 56
 - of pseudo-depolarizer 32
 - orthogonal states 59
- strong mode coupling 398
- SU(2) group 51
- susceptibility 107, 125, 130
 - linear relation between P and E 91
- $\vec{\tau} \times$ 332
- TE wave
 - reflection and transmission 97
- Tellegen parameter 139
- tellurium dioxide (TeO₂) 149, 150
- temperature compensation
 - crystal combinations 173
 - of birefringent phase 170, 458
 - of compound crystal 177
- temperature dependence
 - measurement of group index 163
 - quadratic model 166
- thermally expanded-core fiber 280
- tilt error 243
- TM wave
 - reflection and transmission 99
- total internal reflection 101
 - asymmetric 119
 - differential cutoff (birefringent) 118
 - retarder 196
 - Shirasaki prism 204
- total outage probability 478
- transformation matrix 18, 318, 378
 - Fabry-Perot 155
 - from scattering matrix 155
- transmission coefficient 96, 155
- uniaxial crystal 106
 - propagation in 109
- unitary matrix *see* Mueller matrix
 - calculation 501
 - Cayley-Klein form 51
 - general form 51
- unitary operator 48
 - connection to Hermitian operator 49
 - spin-operator form 68
 - spin-vector form 68
- unitary transform *see* Mueller matrix,
 - see* similarity transform
- Jones-Stokes equivalence 64, 331
- unspun fiber
 - fiber autocorrelation length 385
 - zero chirality 389, 450
- vector potential 9
 - gaussian optics 220
- Verdet constant 133
- walkoff angle 110, 204, 259, 262, 268
 - maximum 117
- walkoff block 119
 - for circulators 279, 281, 290
 - for isolators 259, 266
- wavelength 4
 - in media 93
- wavelength-division multiplexed grid 143
- wavelength-scanning (WS) method 436, 438
 - relationship to INT method 440
- wavenumber 4, 93, 110, 156, 220, 225
 - birefringent 120
 - gyrotropic 129
 - optically active 140
- waveplate 120, 179, 207, *see* half-wave
 - waveplate, *see* quarter-wave
 - waveplate
 - combinations 184
 - extinction ratio 183, 457

- frequency dependence 181
- polarization control 191
- technologies 182
- waveplate model of PMD fiber 417
- weak mode coupling 398
- Wollaston prism 199, 255, 273, 285
 - modified 201
- Xie-Huang circulator 286, 292
- YIG *see* iron garnet
- YVO₄ 147, 165, 176, 185, 201, 206, 258, 262, 456, 464
 - group-index temp. co. 166, 170
 - material properties 148
 - temperature compensation 173

Springer Series in OPTICAL SCIENCES

Volume 1

1 Solid-State Laser Engineering

By W. Koehner, 5th revised and updated ed. 1999, 472 figs., 55 tabs., XII, 746 pages

Published titles since volume 80

80 Optical Properties of Photonic Crystals

By K. Sakoda, 2nd ed., 2004, 107 figs., 29 tabs., XIV, 255 pages

81 Photonic Analog-to-Digital Conversion

By B.L. Shoop, 2001, 259 figs., 11 tabs., XIV, 330 pages

82 Spatial Solitons

By S. Trillo, W.E. Torruellas (Eds.), 2001, 194 figs., 7 tabs., XX, 454 pages

83 Nonimaging Fresnel Lenses

Design and Performance of Solar Concentrators

By R. Leutz, A. Suzuki, 2001, 139 figs., 44 tabs., XII, 272 pages

84 Nano-Optics

By S. Kawata, M. Ohtsu, M. Irie (Eds.), 2002, 258 figs., 2 tabs., XVI, 321 pages

85 Sensing with Terahertz Radiation

By D. Mittleman (Ed.), 2003, 207 figs., 14 tabs., XVI, 337 pages

86 Progress in Nano-Electro-Optics I

Basics and Theory of Near-Field Optics

By M. Ohtsu (Ed.), 2003, 118 figs., XIV, 161 pages

87 Optical Imaging and Microscopy

Techniques and Advanced Systems

By P. Török, F.-J. Kao (Eds.), 2003, 260 figs., XVII, 395 pages

88 Optical Interference Coatings

By N. Kaiser, H.K. Pulker (Eds.), 2003, 203 figs., 50 tabs., XVI, 504 pages

89 Progress in Nano-Electro-Optics II

Novel Devices and Atom Manipulation

By M. Ohtsu (Ed.), 2003, 115 figs., XIII, 188 pages

90/1 Raman Amplifiers for Telecommunications 1

Physical Principles

By M.N. Islam (Ed.), 2004, 488 figs., XXVIII, 328 pages

90/2 Raman Amplifiers for Telecommunications 2

Sub-Systems and Systems

By M.N. Islam (Ed.), 2004, 278 figs., XXVIII, 420 pages

91 Optical Super Resolution

By Z. Zalevsky, D. Mendlovic, 2004, 164 figs., XVIII, 232 pages

92 UV-Visible Reflection Spectroscopy of Liquids

By J.A. Ráty, K.-E. Peiponen, T. Asakura, 2004, 131 figs., XII, 219 pages

93 Fundamentals of Semiconductor Lasers

By T. Numai, 2004, 166 figs., XII, 264 pages

Springer Series in OPTICAL SCIENCES

- 94 **Photonic Crystals**
Physics, Fabrication and Applications
By K. Inoue, K. Ohtaka (Eds.), 2004, 209 figs., XV, 320 pages
- 95 **Ultrafast Optics IV**
Selected Contributions to the 4th International Conference
on Ultrafast Optics, Vienna, Austria
By F. Krausz, G. Korn, P. Corkum, I.A. Walmsley (Eds.), 2004, 281 figs., XIV, 506 pages
- 96 **Progress in Nano-Electro Optics III**
Industrial Applications and Dynamics of the Nano-Optical System
By M. Ohtsu (Ed.), 2004, 186 figs., 8 tabs., XIV, 224 pages
- 97 **Microoptics**
From Technology to Applications
By J. Jahns, K.-H. Brenner, 2004, 303 figs., XI, 335 pages
- 98 **X-Ray Optics**
High-Energy-Resolution Applications
By Y. Shvyd'ko, 2004, 181 figs., XIV, 404 pages
- 99 **Few-Cycle Photonics and Optical Scanning Tunneling Microscopy**
Route to Femtosecond Ångstrom Technology
By M. Yamashita, H. Shigekawa, R. Morita (Eds.) 2005, 241 figs., XX, 393 pages
- 100 **Quantum Interference and Coherence**
Theory and Experiments
By Z. Ficek and S. Swain, 2005, 178 figs., approx. 432 pages
- 101 **Polarization Optics in Telecommunications**
By J. Damask, 2005, 110 figs., XVI, 528 pages
- 102 **Lidar**
Range-Resolved Optical Remote Sensing of the Atmosphere
By C. Weitkamp (Ed.), 161 figs., approx. 416 pages
- 103 **Optical Fiber Fusion Splicing**
By A. D. Yablon, 2005, 100 figs., approx. 300 pages
- 104 **Optoelectronics of Molecules and Polymers**
By A. Moliton, 2005, 200 figs., approx. 460 pages
- 105 **Solid-State Random Lasers**
By M. Noginov, 2005, 149 figs., approx. 380 pages
- 106 **Coherent Sources of XUV Radiation**
Soft X-Ray Lasers and High-Order Harmonic Generation
By P. Jaeglé, 2005, 150 figs., approx. 264 pages
- 107 **Optical Frequency-Modulated Continuous-Wave (FMCW) Interferometry**
By J. Zheng, 2005, 137 figs., approx. 250 pages
- 108 **Laser Resonators and Beam Propagation**
Fundamentals, Advanced Concepts and Applications
By N. Hodgson and H. Weber, 2005, 497 figs., approx. 790 pages