

Modeling and Simulation of Polarization Mode Dispersion and Polarization Dependent Loss

by

Michael Andrew Reimer

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Physics

Waterloo, Ontario, Canada, 2007

© Michael Reimer 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Novel theoretical formulations and efficient simulation methods for polarization-mode dispersion (PMD) and polarization-dependent loss (PDL) that are directly applicable to optical network design are developed. In particular, a formalism based upon the Magnus expansion is advanced for the determination of the frequency evolution of the Mueller matrix in terms of increasing orders of PMD and PDL. Several previous models of polarization evolution are shown to be specializations of this more general formalism.

A least-squares algorithm that extracts PMD and PDL coefficients from repeated measurements of the output Stokes vector of an optical system for a random set of input polarization states is introduced and subsequently applied to the rapid experimental determination of the probability density of the differential group delay of a fiber-squeezer based PMD emulator. The applicability of Clifford algebra and Padé -approximant techniques to the efficient simulation of the wavelength-dependence of PMD and PDL is also discussed.

Acknowledgments

I would like to thank my supervisor, Dr. David Yevick, for the guidance and support that made this work possible. I also wish to thank Dr. Maurice O'Sullivan for many fruitful discussions and for sparking my interest in polarization optics. Support for this thesis was also provided by the National Sciences and Research Council of Canada (NSERC), Nortel, the Centre for Electrophotonic Materials and Devices (CEMD) and the Ontario Research and Development Challenge Fund (ORDCF).

Dedication

To Kerry: for your love, **patience** and support.

Contents

1	Introduction	1
1.1	Polarization Mode Dispersion	1
1.1.1	Birefringence in optical fibers	1
1.1.2	The PMD vector	3
1.1.3	Longitudinal evolution of the PMD vector	5
1.2	Polarization Dependent Loss	9
1.2.1	PDL statistics	10
1.2.2	Principal states of polarization	12
1.3	Conclusions	15
2	Mueller Matrix Description of PMD and PDL	17
2.1	Introduction	17
2.2	Theoretical background	19
2.2.1	Normalized stokes vectors	23
2.3	Magnus expansion	24
2.4	Mueller matrix formalism	25
2.4.1	Comparison with previous work	27
2.5	Operator symmetrization	29
2.6	Conclusions	30

3	The Magnus Expansion for PMD and PDL	32
3.1	Introduction	32
3.2	Mueller matrix formalism	33
3.3	Alternate Mueller matrix derivation	36
3.4	Power series expansion	38
3.5	Results	39
3.6	Conclusions	45
4	A Clifford Algebra Analysis of PMD and PDL	46
4.1	Introduction	46
4.2	Summary of Clifford algebra	47
4.3	PMD and PDL	50
4.4	Conclusions	54
5	A Least-Squares Analysis of the Mueller Matrix	55
5.1	Introduction	55
5.2	Least-squares formulation	57
5.3	Numerical results	59
5.4	Conclusions	63
6	Efficient Jones Matrix Expansions	65
6.1	Introduction	65
6.2	Numerical Techniques	66
6.3	Numerical Results	69
6.4	Conclusions	74
7	Alignment Methods for Biased Multicanonical Sampling	75
7.1	Introduction	75
7.2	Multicanonical sampling	77

7.3	Biased multicanonical sampling	78
7.4	Numerical results	83
7.5	Experimental measurements	89
7.6	Conclusions	91
8	Multicanonical Analyses of System Penalties	92
8.1	Introduction	92
8.2	Compensator Jones matrix	93
8.3	Numerical results	94
8.4	Conclusions	101
9	Conclusions	102
A	PMD/PDL Frequency Evolution	104
A.1	Principal states of polarization for systems with PMD and PDL	104
A.2	The Magnus and Dyson series	105
A.3	Exponential operator expansions of the Mueller matrix	107
A.4	Higher-order PMD vectors in the concatenated segment model	108
A.5	Rotating PMD vector approximation	110
B	Matrix Identities	111
B.1	Properties of the Pauli spin vector	111
	B.1.1 Dot product identities	112
	B.1.2 Jones and Stokes vectors	113
B.2	The Kronecker product	113
B.3	The Mueller matrix	115
	B.3.1 Relation to the Jones matrix	115
	B.3.2 General properties	116
	B.3.3 Coherency matrix transformations	116
B.4	Representations of unitary Jones matrices	117

C	MATLAB Function Implementation	119
C.1	Mueller matrix SVD in Minkowski space	119
C.2	Exact evaluation of Jones matrix derivatives	120
C.3	Exact evaluation of PMD vector derivatives	121
C.4	Jones matrix interpolation	122
C.5	Jones matrix Padé approximant	125
C.6	Optimum biased multicanonical normalization	127
C.7	General polarization functionality	129

List of Figures

1.1	Illustration of $p(\tau N_{seg})$, Eq. (1.17) (solid line), and $p_{\text{maxwell}}(\tau)$, Eq. (1.19) (dotted line), calculated for a concatenation of $N_{seg} = 10$ birefringent segments with $\tau_{seg} = 1$ ps, and τ_{avg} equal to the mean of the calculated $p(\tau N_{seg})$ distribution, respectively. Both curves are displayed as a function of the DGD, τ , normalized by $N_{seg}\tau_{seg}$	7
1.2	The pdf of the total PDL, $\alpha_{tot}(\omega_0)$, for a concatenation of $N_{seg} = 10, 25, 50$ and 75 birefringent and lossy segments, Eq. (1.26), each possessing $\alpha_{dB} = 0.32$ dB. The numerical pdf estimates, displayed as dashed lines, employed 3 multicanonical iterations of 5×10^5 samples, while the superimposed solid lines result from Eq. (1.28).	11
1.3	The pdf of the angle, θ_{psp} , between Stokes space PSPs, Eq. (1.35), for 1 dB (solid line), 2 dB (dashed line) and 3 dB (dashed-dotted line) of mean PDL, α_{avg} , in a concatenation of 100 birefringent and lossy segments with mean DGD $\tau_{avg} = 25$ ps.	14
1.4	The variation of (a) the mean, $\langle\theta_{\text{psp}}\rangle$, and (b) the standard deviation, σ_{psp} , of θ_{psp} , Eq. (1.35), as a function of the mean PDL, α_{avg} , in a concatenation of 100 birefringent and lossy segments displayed here with \circ markers for $\tau_{avg} = 25$ ps. Optimal fits to linear functions are shown as solid lines. . . .	15

3.1	The average relative error, $\langle \epsilon \rangle$, for 5,000 fiber realizations, as a function of the normalized frequency deviation, $\langle \tau \rangle \Delta \omega / (2\pi)$, for a mean DGD $\langle \tau \rangle$. Dashed lines - power series expansion, Eq. (3.18), solid lines - Magnus expansion, Eq. (3.11). The first to fourth order results are indicated by \times , \star , \circ and no markers, respectively.	40
3.2	Same as Fig. 3.1, except for the standard deviation of the relative error, σ_ϵ	41
3.3	The simulated probability density function (pdf) of the relative error, χ , associated with the Taylor coefficients \mathbf{N}_1 , \mathbf{N}_2 , and \mathbf{N}_3 for 2.5×10^5 fiber realizations with 30.0 ps and 10.0 dB mean PDL. The solid, dashed and dotted lines represent the relative error of \mathbf{N}_1 , \mathbf{N}_2 , and \mathbf{N}_3 , respectively.	42
5.1	Average relative error, $\langle \epsilon \rangle$, as a function of the number of Stokes vector measurements, N . The $+$, \times and \circ markers denote $\langle \epsilon \rangle$ calculated using Eqs. (5.5), (5.6) and (5.8), respectively. Dashed lines - results for 1.0% additive noise, markers only - 3.0% additive noise.	60
5.2	As in Fig. 5.1 but for the standard deviation of the relative error, σ_ϵ	61
5.3	Histograms of (top) $ \vec{\Lambda}_0 $ (circles), and $ \vec{\Omega}_0 $ (crosses), and (bottom) $ \vec{\Lambda}_1 $ (circles), and $ \vec{\Omega}_1 $ (crosses) estimated from the least-squares algorithm of Eq. (5.6).	62
6.1	The pdf of the normalized bandwidth, $\Delta\nu = \tau_{avg} \Delta\omega_\alpha / 2\pi$, where $\epsilon_\alpha(\Delta\omega_\alpha) = \epsilon_{max} = 2\%$ for each configuration, $\alpha = 1, 2, \dots, (N_\alpha = 5 \times 10^5)$, of a $N_{seg} = 100$ fiber emulator as calculated with the fourth order Chebyshev (\times), and Taylor series ($+$) expansions as well as the [2, 2]-Padé approximant (\circ) and the interpolation procedure of Eq. (6.2) (solid line).	70

6.2	The minimum value, $\Delta\nu_{\min}$, of $\Delta\nu$ that limits the maximum recorded error, Eq. (6.17), $< 2\%$ for 99.99% of the randomly generated samples graphed as a function of expansion order, N . The \times , $+$ and \circ markers represent Chebyshev, Taylor and Padé approximant expansions, respectively, while the horizontal dashed line represents $\Delta\nu_{\min}$ from Fig. 6.1, of Eq. (6.2). The Δ markers correspond to the $\Delta\nu_{\min}$ obtained if the Taylor series error is limited to $< 2\%$ instead for 95% of the samples. The superimposed solid lines are optimized linear and square-root function approximations.	71
7.1	The maximum relative error, Eq. (7.17), for the initial estimate of the normalization constants of Eq. (7.14) (solid line) and for the iterated results $A_n^{(0)} = 1$, $n = 1, \dots, N_I$ for these constants (dotted line). Results are illustrated for a 100 segment fiber emulator with $\tau_{avg} = 25$ ps and 10 biased multicanonical iterations with 2×10^4 points per iteration with $\gamma = 0.3$	84
7.2	The pdf, Eq. (7.10), obtained from the calculation of Fig. 7.1 after joining the results of 10 different biased calculations after 1 (crosses), 2000 (circles) and 20,000 (dotted-line) iterations of Eq. (7.12), The solid line is the exact result of Eq. (1.17) [54]. The normalization constants are here initialized to $A_n^{(0)} = 1$, $n = 1, \dots, N_I$	85
7.3	(a) The biased pdf estimate \hat{p}_2 before (circles) and after (dashed line) combination with the initial Monte Carlo estimate \hat{p}_1 (solid line) according to Eq. (7.10) with $\gamma = 0.3$. (b) Analogous results for the second biasing iteration in which \hat{p}_1 , \hat{p}_2 and \hat{p}_3 (crosses) are combined according to Eq. (7.10) (solid line). The vertical lines indicate the biasing region.	86
7.4	The pdf of the DGD of a 100 section fiber emulator calculated with fifty 5×10^4 -sample iterations with the iterative biased multicanonical method (dashed-dotted line) and the standard multicanonical procedure (\circ) together with the analytic result (solid line).	87

7.5	The biased multicanonical experimental setup.	89
7.6	The experimentally determined pdf of the DGD of an 8 section fiber emulator for 45,000 samples measured with the piecewise biased multicanonical method (circles) and the standard Monte Carlo procedure (crosses). The solid line indicates the corresponding numerical results for three 5×10^5 sample iterations in the standard multicanonical method.	90
8.1	Probability density function of the eye-closure penalty for $\tau_{avg} = 30$ ps without compensation (solid line), and after the compensators described by the Jones matrices of Eq. (8.2) (\circ), Eq. (8.3) (\times), and Eq. (8.4) ($+$). . . .	95
8.2	Simulation configuration.	96
8.3	Projections of the histogram of the states of our system model that contribute to the eye-closure penalty with a 10^{-5} probability density after the compensation of Eq. (8.2) onto normalized (a) PMD/SOPMD, (b) PMD/-TOPMD and (c) SOPMD/TOPMD planes.	98
8.4	Same as Fig. 8.4 except for the compensator Jones matrix of Eq. (8.3). . .	99
8.5	Same as Fig. 8.4 except for the compensator Jones matrix of Eq. (8.4). . .	100

List of Tables

3.1 Measured relative error χ (%). 44

Table of Symbols

Symbol	Definition	Page
\dagger	Hermitian conjugation	9
$*$	Complex conjugation	20
\otimes	Kronecker (direct) matrix product	20
$\text{Tr}(\cdot)$	Matrix trace	20
$\langle \cdot \rangle$	Ensemble average	14
ω	Optical frequency [rad/s]	2
ω_0	Optical carrier frequency [rad/s]	11
$\Delta\omega$	Frequency deviation relative to the optical carrier [rad/s]	18
c	Speed of light [m/s]	3
k	Wavenumber [rad/m]	5
z	Longitudinal distance [m]	8
σ	Differential core stress [Pa]	2
Y	Young's modulus [Pa]	3
ρ_{44}	Component of the elasto-optic strain tensor [unitless]	3
$\vec{\beta}$	Local fiber birefringence [rad/m]	2
α_{seg}	Segment polarization dependent loss (PDL) [Np]	12
α_{avg}	Mean polarization dependent loss after N segments [dB]	14
α_{tot}	Total polarization dependent loss after N segments [dB]	10

Symbol	Definition	Page
τ	Differential group delay (DGD) [s]	4
τ_{seg}	Segment differential group delay [s]	7
τ_{avg}	Mean differential group delay after N segments [s]	7
η	Differential attenuation slope (DAS) [Np·s]	13
\mathbf{I}_N	$N \times N$ identity matrix	27
\mathbf{T}	2×2 complex Jones matrix	3
\mathbf{M}	4×4 Mueller matrix	18
\mathbf{R}	3×3 rotation matrix	4
$\vec{\sigma}$	Pauli spin vector	3
$\tilde{\sigma}$	Extended Pauli spin vector	19
σ_0	2×2 identity matrix	3
$ t\rangle$	2×1 complex Jones vector	3
\vec{t}	3×1 vector	2
\hat{t}	3×1 unit vector	2
\tilde{t}	4×1 Stokes vector	19
$\vec{t} \times$	3×3 skew-symmetric matrix	5
$ p\rangle$	Jones space principal state of polarization (PSP)	4
\hat{p}	Stokes space principal state of polarization	4
\vec{W}	Complex principal state vector	13
$\vec{\Omega}$	Real component of complex principal state vector	4
$\vec{\Lambda}$	Imaginary component of complex principal state vector	13

Chapter 1

Introduction

An introduction to birefringence in single mode optical fiber is presented in Section 1.1.1. Section 1.1.2 introduces the polarization mode dispersion (PMD) vector and its relation to the principal states of polarization. This is followed by a review of polarization dependent loss and the complex principal state vector.

1.1 Polarization Mode Dispersion

1.1.1 Birefringence in optical fibers

The ideal isotropic optical properties of single mode fiber are disturbed by random manufacturing defects and mechanical stresses applied along the length of the fiber core. These effects establish preferential optical axes in the fiber, breaking the degeneracy and inducing birefringence between the two polarized states of the fundamental fiber mode. Incident polarized light resolved along each optical axis experiences modified waveguiding characteristics, leading to a relative difference in propagation delay, i.e. a differential group delay (DGD), between polarization modes [3, 18, 41, 45]. Though normally much smaller than

chromatic dispersion (CD) effects, this polarization mode dispersion (PMD) effect can become a major link-design concern in CD compensated systems operating at bit rates of 10 Gb/s and higher [3].

There are two origins of birefringence in optical fiber: variations of the fiber from an ideal cylindrical geometry, and the presence of residual mechanical stress or strain in the fiber core [3, 18, 63]. Geometric imperfections of the fiber are well approximated by first-order perturbation techniques in the limit of weak core ellipticity [63, 93], with preferential optical axes coincident with the major and minor axes of the elliptical cross-section. Stress-induced birefringence, on the other hand, yields a linear increase in the dielectric permittivity, with optical axes aligned with the principal directions of the mechanical perturbation [29, 99, 105, 106]. In each case, the two preferential optical axes denoted the fast and slow transmission axes, with associated propagation constants β_{fast} and β_{slow} , are distinguished by their respective transmission delays.

Variations in the core geometry, $\vec{\beta}_c$, transverse fiber stress, $\vec{\beta}_s$, due to thermal expansion gradients or applied external pressure, and fiber twist, $\vec{\beta}_t$, each contribute linearly to the total local birefringence, $\vec{\beta}$, of the fiber [29, 82, 105, 106]. Deviations of the core from an ideal cylindrical cross-section induce a birefringence with magnitude $\beta_c = |\vec{\beta}_c|$ related to the core ellipticity, e [29, 63, 63, 106],

$$\beta_c \approx \frac{0.13e^2}{b}(2\Delta)^{3/2} \quad (1.1)$$

independent of the optical frequency, ω , over the frequency range of normal fiber operation. Here, as in Ref. [29], $e = \sqrt{1 - b^2/a^2}$ with b and a the lengths of the major and minor axes, respectively, Δ the relative refractive index difference between core and cladding, while $\hat{\beta}_c$ aligns with the Stokes space representation of the minor axis. Transverse fiber stress instead generates birefringence varying linearly with both the magnitude of the differential core stress, σ , and optical frequency [29, 105],

$$\beta_s = \bar{n}^3 |\rho_{44}| \frac{\omega}{c} \frac{\sigma}{Y} (1 + N), \quad (1.2)$$

for speed of light, c , with the orientation, $\hat{\beta}_s$, determined by the direction of maximum compressive force. Material properties of silica glass enter Eq. (1.2) through Young's modulus, Y , Poisson's ratio, N , the mean refractive index of core and cladding, \bar{n} , and one component of the elasto-optic strain tensor, ρ_{44} [29]. Finally, a mechanical twist of the fiber core imparts

$$\beta_t \approx -\bar{n}^2 |\rho_{44}| T, \quad (1.3)$$

approximately independent of the optical frequency [29, 82, 106]. The orientation of $\vec{\beta}_t$ is specified by the product of the Poincaré sphere representation of left-hand circular polarization, \hat{k} , and the twist rate T , in units of rad/m, respectively, which in the case of the latter, takes on positive and negative values for right- and left-oriented twists of the fiber. Each of the $\vec{\beta}_c$, $\vec{\beta}_s$ and $\vec{\beta}_t$ contributions sum to yield the total local birefringence,

$$\vec{\beta} = \vec{\beta}_c + \vec{\beta}_s + \vec{\beta}_t, \quad (1.4)$$

of the optical fiber.

1.1.2 The PMD vector

In this section, we derive a differential relationship between the Jones matrix, $\mathbf{T}(\omega)$, and the PMD of an optical system by considering the frequency evolution of the output electric field Jones vector, $|t(\omega)\rangle = e^{-i\phi_0(\omega)} \mathbf{T}(\omega) |s\rangle$, for which the input, $|s\rangle$, is independent of ω [45, 59]. The group delay common to both polarization modes follows in this formulation from the frequency dependence of the phase, $\phi_0(\omega)$, of the fiber transfer function [30], though any overall attenuation and polarization dependent losses are neglected such that $\mathbf{T}(\omega)$ is unitary and takes the general form

$$\begin{aligned} \mathbf{T}(\omega) &= \exp \left[-\frac{i}{2} \psi(\omega) (\hat{n}(\omega) \cdot \vec{\sigma}) \right] \\ &= \cos \left(\frac{\psi}{2} \right) \sigma_0 - i \sin \left(\frac{\psi}{2} \right) \hat{n} \cdot \vec{\sigma} \end{aligned} \quad (1.5)$$

for $\psi(\omega)$, $\hat{n}(\omega)$ a real scalar and unit vector, respectively [45]. A direct calculation establishes the central result of this section,

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} = -\frac{i}{2}\vec{\Omega}(\omega) \cdot \vec{\sigma}, \quad (1.6)$$

where

$$\vec{\Omega}(\omega) \equiv \frac{d\psi}{d\omega}\hat{n} + \sin(\psi)\frac{d\hat{n}}{d\omega} + [1 - \cos(\psi)]\hat{n} \times \frac{d\hat{n}}{d\omega} \quad (1.7)$$

is termed the PMD vector of the optical system [17, 45, 79].

That $\vec{\Omega}$ characterizes the system's polarization dependent delay is evident through a further analysis of $d|t\rangle/d\omega$. Applying Eq. (1.6), we find for $\bar{\tau} \equiv d\phi_0/d\omega$ the common propagation delay

$$\begin{aligned} \frac{d|t\rangle}{d\omega} &= \left(-i\frac{d\phi_0}{d\omega}\boldsymbol{\sigma}_0 + \frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} \right) |t\rangle \\ &= -i \left(\bar{\tau}\boldsymbol{\sigma}_0 + \frac{1}{2}\vec{\Omega}(\omega) \cdot \vec{\sigma} \right) |t\rangle. \end{aligned} \quad (1.8)$$

Eq. (1.8) immediately implies the existence of two principal states of polarization (PSPs), $|p_{\pm}\rangle$, the Jones space eigenvectors of $\vec{\Omega} \cdot \vec{\sigma}$ with corresponding eigenvalues $\pm|\vec{\Omega}|$, for which

$$\frac{d|p_{\pm}\rangle}{d\omega} = -i \left[\bar{\tau} \pm \frac{|\vec{\Omega}|}{2} \right] |p_{\pm}\rangle \quad (1.9)$$

Clearly the differential propagation time, τ , between incident polarized light coupled to the $|p_{\pm}\rangle$ Jones space PSPs is identical to $|\vec{\Omega}|$ [45]. Further, the Stokes space vector $\hat{p}_{\pm} \equiv \langle p_{\pm} | \vec{\sigma} | p_{\pm} \rangle$, as demonstrated in Appendix A.1, satisfies $\hat{p}_{\pm} = \pm\vec{\Omega}/|\vec{\Omega}|$; that is, the Stokes space PSPs are anti-parallel and in the absence of PDL coincide with the direction of $\vec{\Omega}$. We conclude

$$\vec{\Omega} = \tau\hat{p}, \quad (1.10)$$

where by convention $\hat{p} \equiv \hat{p}_+$ is the slow optical transmission axis of the fiber.

Expressed in terms of the 3×3 rotation matrix, $\mathbf{R}(\omega)$, homomorphic to $\mathbf{T}(\omega)$, Eq. (1.6) has the equivalent Stokes space representation, c.f. Chapter 2,

$$\frac{d\mathbf{R}}{d\omega}\mathbf{R}^{-1} = \vec{\Omega}(\omega) \times, \quad (1.11)$$

such that $\vec{\Omega}$ effects an infinitesimal rotation of the output polarization Stokes vector, $\vec{t} \equiv \langle t|\vec{\sigma}|t\rangle$, according to $d\vec{t}/d\omega = (d\mathbf{R}/d\omega)\mathbf{R}^{-1}\vec{t} = \vec{\Omega} \times \vec{t}$. Consequently, \vec{t} precesses with increasing optical frequency about \hat{p} at a rate determined by the differential group delay, τ , a property that is often employed in the experimental determination of the PMD vector [21, 32, 47, 52, 77].

1.1.3 Longitudinal evolution of the PMD vector

The longitudinal dependence of the birefringence induces a stochastic evolution of the PMD vector along the fiber length. We analyze this variation of $\vec{\Omega}$ first under a finite segment model of fiber birefringence and subsequently in the continuum limit. This first approach leads to the concatenation rule for the addition of PMD vectors of bulk optical devices [45]. The second method instead generates a differential equation relating the longitudinal evolution of the PMD vector to the underlying fiber birefringence [37, 80]. In each case, the probability density functions of the DGD are discussed.

Over a length of fiber small with respect to the fiber's decorrelation length, L_c , [37, 66, 109] any perturbations of the core giving rise to localized birefringence can be considered to act uniformly over the short segment [29]. In this limit, the DGD, τ , between the fast and slow transmission axes

$$\tau = \frac{L}{c} \frac{d}{dk} (\beta_{\text{fast}} - \beta_{\text{slow}}) = L \frac{d|\vec{\beta}|}{d\omega} \quad (1.12)$$

for wavenumber, k , increases linearly with segment length, L [37]. However, practical optical transmission systems are much larger than the 1 – 10m decorrelation lengths typical of standard single mode fiber and random fluctuations of the birefringence along the fiber length must be considered [3]. Perhaps the most convenient model of these stochastic variations involves a concatenation of N_{seg} randomly oriented birefringent segments, each with $L \ll L_c$, where coupling between polarization modes occurs only at the $N_{\text{seg}} - 1$ discrete segment junctions. In terms of the DGD, $\tau_{(m)}$, and Stokes space slow axis, $\hat{p}_{(m)}$, of

the m^{th} short segment, the fiber's frequency domain Jones matrix, $\mathbf{T}(\omega)$, becomes in this model [22, 26, 66]

$$\begin{aligned}\mathbf{T}(\omega) &\equiv \mathbf{T}_{(N_{\text{seg}})}(\omega) \dots \mathbf{T}_{(2)}(\omega) \mathbf{T}_{(1)}(\omega) \\ &= \prod_{m=1}^{N_{\text{seg}}} \mathbf{T}_{(m)}(\omega)\end{aligned}\quad (1.13)$$

where

$$\mathbf{T}_{(m)}(\omega) = \exp \left[-\frac{i}{2} \omega \tau_{(m)} (\hat{p}_{(m)} \cdot \vec{\sigma}) \right] \quad (1.14)$$

and $\vec{\Omega}_{(m)} = \tau_{(m)} \hat{p}_{(m)}$ is the segment's frequency-independent PMD vector.

The Jones matrix, \mathbf{T} , of two concatenated fiber segments, or in general two birefringent optical devices, with corresponding Jones matrices $\mathbf{T}_{(1)}$, $\mathbf{T}_{(2)}$, and PMD vectors $\vec{\Omega}_{(1)}$, $\vec{\Omega}_{(2)}$ is according to Eq. (1.13) the product $\mathbf{T} = \mathbf{T}_{(2)} \mathbf{T}_{(1)}$. Associated with \mathbf{T} is the PMD vector, $\vec{\Omega}$, related to $\vec{\Omega}_{(1)}$ and $\vec{\Omega}_{(2)}$ through an application of Eq. (1.6),

$$\begin{aligned}\vec{\Omega} \cdot \vec{\sigma} &\equiv 2i \frac{d\mathbf{T}}{d\omega} \mathbf{T}^{-1} = 2i \frac{d\mathbf{T}_{(2)}}{d\omega} \mathbf{T}_{(2)}^{-1} + 2i \mathbf{T}_{(2)} \left[\frac{d\mathbf{T}_{(1)}}{d\omega} \mathbf{T}_{(1)}^{-1} \right] \mathbf{T}_{(2)}^{-1} \\ &= \vec{\Omega}_{(2)} \cdot \vec{\sigma} + \mathbf{T}_{(2)} \left[\vec{\Omega}_{(1)} \cdot \vec{\sigma} \right] \mathbf{T}_{(2)}^{-1} \\ &= \vec{\Omega}_{(2)} \cdot \vec{\sigma} + \left[\mathbf{R}_{(2)} \vec{\Omega}_{(1)} \right] \cdot \vec{\sigma} \\ &= \left[\vec{\Omega}_{(2)} + \mathbf{R}_{(2)} \vec{\Omega}_{(1)} \right] \cdot \vec{\sigma}\end{aligned}\quad (1.15)$$

Here, we have employed Eq. (B.53) where $\mathbf{R}_{(2)}$ is the 3×3 Stokes rotation matrix associated with $\mathbf{T}_{(2)}$, c.f. Appendix B.4. The PMD concatenation rule [45, 59] in more general notation follows immediately,

$$\vec{\Omega}_{\text{output}} = \vec{\Omega}_{\text{segment}} + \mathbf{R}_{\text{segment}} \vec{\Omega}_{\text{input}} \quad (1.16)$$

indicating that the output PMD vector of a fiber segment, $\vec{\Omega}_{\text{output}}$, is the sum of the segment's intrinsic PMD vector, $\vec{\Omega}_{\text{segment}}$, and the input PMD vector, $\vec{\Omega}_{\text{input}}$, rotated by the segment's Stokes matrix, $\mathbf{R}_{\text{segment}}$.

Applying Eq. (1.16) recursively to the birefringence model of Eq. (1.13), the fiber exhibits the maximum possible delay $\tau_{\text{max}} = \tau_{(1)} + \tau_{(2)} \dots + \tau_{(N_{\text{seg}})}$ for co-parallel $\hat{p}_{(m)}$,

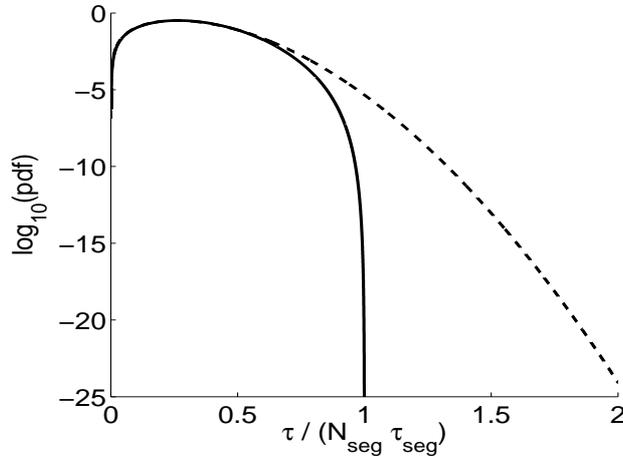


Figure 1.1: Illustration of $p(\tau|N_{seg})$, Eq. (1.17) (solid line), and $p_{\text{maxwell}}(\tau)$, Eq. (1.19) (dotted line), calculated for a concatenation of $N_{seg} = 10$ birefringent segments with $\tau_{seg} = 1$ ps, and τ_{avg} equal to the mean of the calculated $p(\tau|N_{seg})$ distribution, respectively. Both curves are displayed as a function of the DGD, τ , normalized by $N_{seg}\tau_{seg}$.

otherwise random fluctuations in the orientation of each fiber segment due to mechanical perturbations cause $\vec{\Omega}_{\text{output}}$ to execute a random walk in Stokes space. If in this model all segment PMD vectors are of equal length, $\tau_{(m)} = \tau_{seg}$, with $\hat{p}_{(m)}$ uniformly distributed on the Poincaré sphere, the probability density of the fiber's DGD, $\tau \equiv |\vec{\Omega}_{\text{output}}|$, after N_{seg} segments has the exact analytic form [54]

$$p(\tau|N_{seg}) = \frac{\tau}{2\tau_{seg}^2(N_{seg}-2)!} \sum_{m=0}^K (-1)^m \binom{N_{seg}}{m} (N_{seg}R_\tau - m)^{(N_{seg}-2)} \quad (1.17)$$

with

$$R_\tau \equiv \frac{1}{2} \left(1 - \frac{\tau}{\tau_{seg}N_{seg}} \right) \quad (1.18)$$

and $K = \text{floor}(N_{seg}R_\tau)$. It follows [54] that the mean DGD of the concatenation, $\tau_{avg} \propto \sqrt{\tau_{(1)}^2 + \tau_{(2)}^2 + \dots + \tau_{(N_{seg})}^2} \sim \tau_{seg} \sqrt{N_{seg}}$, displays a square-root dependence on the fiber length in contradistinction to the linearity of Eq. (1.12) [37, 39].

We observe that Eq. (1.17) differs materially from the Maxwellian distribution,

$$p_{\text{maxwell}}(\tau) = \frac{32\tau^2}{\pi^2\tau_{\text{avg}}^3} \exp\left(-\frac{4\tau^2}{\pi\tau_{\text{avg}}^2}\right), \quad (1.19)$$

specified here in terms of the mean DGD, τ_{avg} , usually assumed for τ . In particular, the probability of τ exceeding $N_{\text{seg}}\tau_{\text{seg}}$, i.e. $P(\tau > N_{\text{seg}}\tau_{\text{seg}})$, in Eq. (1.17) is zero (though Eq. (1.17) and Eq. (1.19) coincide as $N_{\text{seg}} \rightarrow \infty$ [54]). This point is exemplified in Fig. 1.1, which displays $p(\tau|N_{\text{seg}})$ (solid line) calculated for a concatenation of $N_{\text{seg}} = 10$ birefringent segments each possessing $\tau_{\text{seg}} = 1$ ps. The dotted line of this figure represents $p_{\text{maxwell}}(\tau)$ with τ_{avg} equal to the mean of the calculated $p(\tau|N_{\text{seg}})$ distribution. The behaviour in the vicinity of $N_{\text{seg}}\tau_{\text{seg}}$ of each curve is apparent from the comparison against the normalized $\tau/(N_{\text{seg}}\tau_{\text{seg}})$ parameter.

In the continuum limit, the longitudinal evolution of the PMD vector and by extension the DGD, τ , is determined by

$$\frac{\partial \vec{\Omega}}{\partial z} = \frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\Omega} \quad (1.20)$$

directly in terms of the underlying birefringence of the optical fiber. Viewed as a stochastic differential equation for $\vec{\Omega}$, Eq. (1.20) facilitates the statistical analysis of PMD dynamics and in particular rigorously establishes in the long fiber limit the Maxwellian probability density, Eq. (1.19), for τ [35–37, 39, 78, 80]. Eq. (1.20) is perhaps most easily derived by analyzing the evolution of the polarization Stokes vector $\vec{t}(\omega, z)$ of the propagating electric field. At each distance, z , $\vec{t}(\omega, z)$ precesses about $\vec{\beta}(\omega, z)$ at a rate equal to the magnitude of the birefringence [45, 53, 80], such that

$$\frac{\partial \vec{t}}{\partial z} = \vec{\beta}(\omega, z) \times \vec{t} \quad (1.21)$$

Differentiating Eq. (1.21) with respect to ω and $(\partial \vec{t} / \partial \omega) = \vec{\Omega}(\omega, z) \times \vec{t}$, c.f. Eq. 1.11, with respect to z yields

$$\frac{\partial^2 \vec{t}}{\partial z \partial \omega} = \frac{\partial \vec{\beta}}{\partial \omega} \times \vec{t} + \vec{\beta} \times \frac{\partial \vec{t}}{\partial \omega} = \frac{\partial \vec{\Omega}}{\partial z} \times \vec{t} + \vec{\Omega} \times \frac{\partial \vec{t}}{\partial z} \quad (1.22)$$

or after simplification

$$\frac{\partial \vec{\Omega}}{\partial z} \times \vec{t} = \left[\frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\Omega} \right] \times \vec{t} \quad (1.23)$$

That Eq. (1.23) must hold for all polarizations, \vec{t} , implies Eq. (1.20). Alternatively, Eq. (1.20) can be derived from the PMD concatenation rule, Eq. (1.16), in the limit of infinitesimal segment length, i.e. $L \rightarrow 0$ [45].

1.2 Polarization Dependent Loss

Optical communication systems often include components, notably fiber amplifiers, optical couplers and isolators, that may, unlike standard single mode fiber, possess significant polarization dependent loss (PDL). Typically, these components are interspersed amongst birefringent optical devices in a network, generating an interaction between PMD and PDL that further complicates the frequency evolution of the state of polarization [40, 50, 51]. Accordingly, we introduce in this section a model of discretized PDL and discuss the modifications necessary to the formalism of Section 1.1 in light of PMD/PDL coupling. Two important results follow from this analysis, 1) the PMD vector, Eq. (1.6), in the presence of PDL is complex valued and 2) in general, the Jones space PSPs are not orthogonal [50]. In this latter case, an expression for the angle of PSP separation is presented in terms of the real and imaginary components of the complex PMD vector.

Polarized light incident to an optical device with PDL experiences different frequency independent attenuations, α_1 and α_2 , along two orthogonal Jones space axes [40], an observation that immediately leads to the form of the component's Jones matrix

$$\mathbf{L} = \mathbf{U} \begin{bmatrix} e^{-\alpha_1} & 0 \\ 0 & e^{-\alpha_2} \end{bmatrix} \mathbf{U}^{-1} = e^{-\bar{\alpha}} \mathbf{U} \begin{bmatrix} e^{-\alpha/2} & 0 \\ 0 & e^{\alpha/2} \end{bmatrix} \mathbf{U}^\dagger \quad (1.24)$$

Above, $\bar{\alpha} = (\alpha_1 + \alpha_2)/2$ and $\alpha = (\alpha_1 - \alpha_2)$ denote the mean and differential attenuation between polarization modes, respectively, while the unitary \mathbf{U} effects a similarity

transformation prescribing the arbitrary orientation of the principal attenuation axes. The equivalent exponential representation [50]

$$\begin{aligned}
\mathbf{L} &= e^{-\bar{\alpha}} \exp \left[\frac{1}{2} \alpha (\hat{r} \cdot \vec{\sigma}) \right] \\
&= e^{-\bar{\alpha}} \left[\cosh \left(\frac{\alpha}{2} \right) \sigma_0 + \sinh \left(\frac{\alpha}{2} \right) (\hat{r} \cdot \vec{\sigma}) \right] \\
&\equiv e^{-\bar{\alpha}} \mathbf{T}^{\text{pdl}}
\end{aligned} \tag{1.25}$$

is often of more practical concern, however, as the most and least attenuated axes enter explicitly in this formulation through the Stokes space unit vectors \hat{r} and $-\hat{r}$, respectively. In view of Eq. (1.25), the concatenated segment model, Eq. (1.13), becomes

$$\begin{aligned}
\mathbf{T}(\omega) &= \prod_{m=1}^{N_{seg}} \mathbf{T}_{(m)}^{\text{pdl}} \mathbf{T}_{(m)}^{\text{pmd}}(\omega), \\
\mathbf{T}_{(m)}^{\text{pdl}} &= \exp \left[\frac{1}{2} \alpha_{(m)} (\hat{r}_{(m)} \cdot \vec{\sigma}) \right] \\
\mathbf{T}_{(m)}^{\text{pmd}}(\omega) &= \exp \left[-\frac{i}{2} \omega \tau_{(m)} (\hat{p}_{(m)} \cdot \vec{\sigma}) \right]
\end{aligned} \tag{1.26}$$

for the differential polarization behaviour of optical systems with PMD and PDL.

The total PDL, $\alpha_{tot}(\omega)$, after N_{seg} concatenated segments is the ratio of the maximum to minimum transmitted power over all input polarization states, defined in terms of the eigenvalues, $\lambda_i(\omega)$, $i = 1, 2$, of $\mathbf{T}^\dagger(\omega)\mathbf{T}(\omega)$ as [64]

$$\alpha_{tot}(\omega) = 10 \left| \log_{10} \left[\frac{\lambda_1(\omega)}{\lambda_2(\omega)} \right] \right| \tag{1.27}$$

1.2.1 PDL statistics

The optical system of Eq. (1.26) exhibits a maximum differential attenuation $\alpha_{(1)} + \alpha_{(2)} \dots + \alpha_{(N_{seg})}$ otherwise randomized coupling between polarization modes causes $\alpha_{tot}(\omega)$ to explore a characteristic probability density. As previous authors [64, 72] have shown, this probability density must in the limit $N_{seg} \rightarrow \infty$ approach the Maxwellian distribution. We

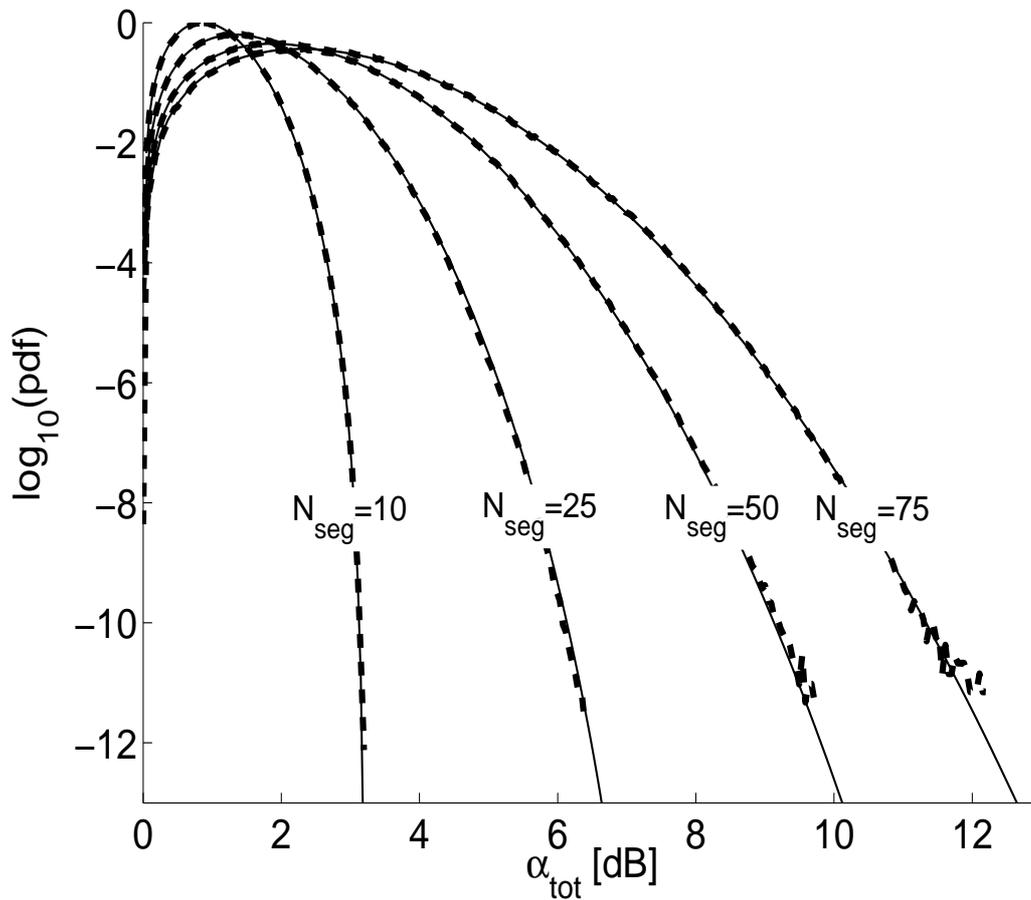


Figure 1.2: The pdf of the total PDL, $\alpha_{\text{tot}}(\omega_0)$, for a concatenation of $N_{\text{seg}} = 10, 25, 50$ and 75 birefringent and lossy segments, Eq. (1.26), each possessing $\alpha_{\text{dB}} = 0.32$ dB. The numerical pdf estimates, displayed as dashed lines, employed 3 multicanonical iterations of 5×10^5 samples, while the superimposed solid lines result from Eq. (1.28).

therefore anticipate that if in Eq. (1.26) $\alpha_{(m)} = \alpha_{seg}$, $m = 1, \dots, N_{seg}$, and $\hat{r}_{(m)}$ is uniformly distributed on the Poincaré sphere, the probability density of the total differential attenuation, $\alpha_{tot}(\omega)$, after N_{seg} segments is given by, c.f. Eq. (1.17),

$$p(\alpha_{tot}|N_{seg}) = \frac{\alpha_{tot}}{2\alpha_{dB}^2(N_{seg} - 2)!} \sum_{m=0}^K (-1)^m \binom{N_{seg}}{m} (N_{seg}R_\alpha - m)^{(N_{seg}-2)} \quad (1.28)$$

with $\alpha_{dB} \equiv \alpha_{seg} \cdot 20/\ln 10$,

$$R_\alpha \equiv \frac{1}{2} \left(1 - \frac{\alpha_{tot}}{\alpha_{dB}N_{seg}} \right) \quad (1.29)$$

and $K = \text{floor}(N_{seg}R_\alpha)$. This result is confirmed in Fig. 1.2 which illustrates the pdf of $\alpha_{tot}(\omega_0)$ for a concatenation of $N_{seg} = 10, 25, 50$ and 75 birefringent and lossy segments each possessing $\alpha_{dB} = 0.32$ dB, for $\tau_{avg} = 25$ ps and ω_0 the optical carrier. The dashed lines of this figure correspond to numerically determined estimates of the pdf, in which our multicanonical simulation [112], c.f. Chapter 7, employed 3 iterations of 5×10^5 statistical realizations of Eq. (1.26). The solid lines of Fig. 1.2 instead result from Eq. (1.28) as applied to each value of N_{seg} . Clearly, the agreement with numerical simulation supports the validity of the PDL pdf estimate, Eq. (1.28).

1.2.2 Principal states of polarization

We next analyze the frequency evolution of the output Stokes vector in the presence of PMD and PDL, $|t(\omega)\rangle = Ae^{-\alpha_0(\omega) - i\phi_0(\omega)}\mathbf{T}(\omega)|s\rangle$, for

$$A \equiv \exp \left[- \sum_{m=1}^{N_{seg}} \bar{\alpha}_{(m)} \right], \quad (1.30)$$

incorporating into the description a possibly frequency dependent attenuation, $\alpha_0(\omega)$, and phase, $\phi_0(\omega)$, common to both polarization modes. The polarization sensitivity of

$$\frac{d|t\rangle}{d\omega} = \left[- \left(\frac{d\alpha_0}{d\omega} + i \frac{d\phi_0}{d\omega} \right) \boldsymbol{\sigma}_0 + \frac{d\mathbf{T}}{d\omega} \mathbf{T}^{-1} \right] |t\rangle \quad (1.31)$$

enters through $(d\mathbf{T}/d\omega)\mathbf{T}^{-1}$, where unlike Section 1.1.2, \mathbf{T} is non-unitary with $\det(\mathbf{T}) = 1$, c.f. Eq. (1.26). Accordingly, we will demonstrate

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} = -\frac{i}{2}\vec{W} \cdot \vec{\sigma} \quad (1.32)$$

for some complex vector \vec{W} by differentiating the trivial identity $\mathbf{T}\mathbf{T}^{-1} = \boldsymbol{\sigma}_0$. Denoting $\mathbf{T}^{\text{adj}} = \mathbf{T}^{-1}$ the adjoint of \mathbf{T} , $d(\mathbf{T}\mathbf{T}^{\text{adj}})/d\omega = 0$ and

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{\text{adj}} = -\mathbf{T}\frac{d\mathbf{T}^{\text{adj}}}{d\omega} = -\left[\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{\text{adj}}\right]^{\text{adj}} \quad (1.33)$$

In terms of the complex scalar and vector, v_0 and \vec{v} , respectively, the decompositions $(d\mathbf{T}/d\omega)\mathbf{T}^{\text{adj}} = v_0\boldsymbol{\sigma}_0 + \vec{v} \cdot \vec{\sigma}$, c.f. Eq. B.5, and $[(d\mathbf{T}/d\omega)\mathbf{T}^{\text{adj}}]^{\text{adj}} = v_0\boldsymbol{\sigma}_0 - \vec{v} \cdot \vec{\sigma}$ when combined with Eq. (1.33) immediately imply $v_0 \equiv 0$. Consequently, $(d\mathbf{T}/d\omega)\mathbf{T}^{\text{adj}} = \vec{v} \cdot \vec{\sigma}$ and $\vec{W} \equiv 2i\vec{v}$. In analogy with Section 1.1.2, $\vec{W} \equiv \vec{\Omega} + i\vec{\Lambda}$, termed the complex principal state vector, for real $\vec{\Omega}$ and $\vec{\Lambda}$ characterizes the polarization frequency evolution of systems with PMD and PDL [40, 50, 51].

The two Jones space PSPs, $|p_{\pm}\rangle$, are in the presence of PDL eigenvectors of $\vec{W} \cdot \vec{\sigma}$, with corresponding complex eigenvalues $\pm\chi = \pm\sqrt{\vec{W} \cdot \vec{W}} \equiv \pm(\tau + i\eta)$. Here, the real and imaginary components, τ and η , of χ specify in turn an effective DGD and a differential attenuation slope (DAS) between the two fundamental polarization modes [40, 50, 51]. In Stokes space, the PSPs

$$\begin{aligned} \hat{p}_{\pm} &\equiv \langle p_{\pm} | \vec{\sigma} | p_{\pm} \rangle \\ &= \pm K \left(\tau\vec{\Omega} + \eta\vec{\Lambda} \pm \vec{\Omega} \times \vec{\Lambda} \right) \end{aligned} \quad (1.34)$$

for $K \equiv 2/(\tau^2 + \eta^2 + |\vec{\Omega}|^2 + |\vec{\Lambda}|^2)$, see Appendix A.1, are not anti-parallel with

$$\begin{aligned} \cos\theta_{\text{psp}} &= \hat{p}_+ \cdot \hat{p}_- \\ &= K^2 \left(|\vec{\Omega} \times \vec{\Lambda}|^2 - |\tau\vec{\Omega} + \eta\vec{\Lambda}|^2 \right) \end{aligned} \quad (1.35)$$

specifying the PSP angular separation.

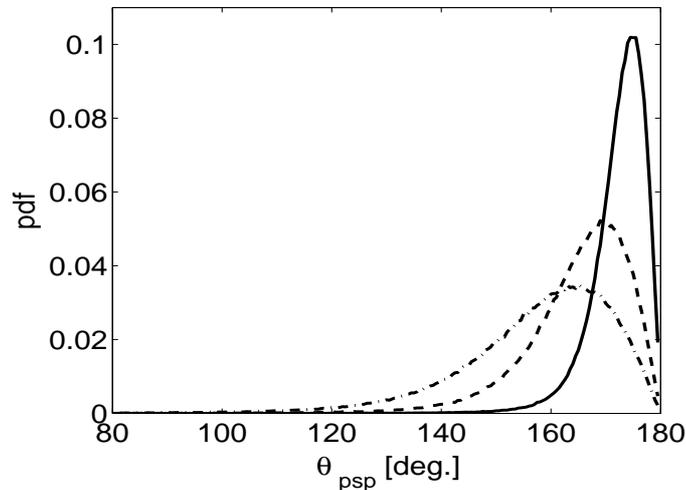


Figure 1.3: The pdf of the angle, θ_{psp} , between Stokes space PSPs, Eq. (1.35), for 1 dB (solid line), 2 dB (dashed line) and 3 dB (dashed-dotted line) of mean PDL, α_{avg} , in a concatenation of 100 birefringent and lossy segments with mean DGD $\tau_{\text{avg}} = 25$ ps.

Over 10^6 statistical realizations of $N_{\text{seg}} = 100$ concatenated segments, i.e.

$$\mathbf{T}(\omega) = \prod_{m=1}^{100} \mathbf{T}_{(m)}^{\text{pdl}} \mathbf{T}_{(m)}^{\text{pmd}}(\omega) \quad (1.36)$$

with $\hat{r}_{(m)}$ and $\hat{p}_{(m)}$ uniformly distributed on the Poincaré sphere, θ_{psp} exhibits the characteristic probability density illustrated in Fig. 1.3 for 1 dB (solid line), 2 dB (dashed line) and 3 dB (dashed-dotted line) of mean PDL, $\alpha_{\text{avg}} = \langle \alpha_{\text{tot}}(\omega_0) \rangle$, respectively. The decrease in the mean angle of PSP separation, $\langle \theta_{\text{psp}} \rangle$, for increasing values of α_{avg} clearly evident in Fig. 1.3 is examined further in Fig. 1.4a, which displays as \circ markers $\langle \theta_{\text{psp}} \rangle$ as a function of α_{avg} , and in Fig. 1.4b, which instead shows the variation of the standard deviation of PSP separation, σ_{psp} . In both cases, the solid lines indicate optimal fits to linear functions. Although deviations from linearity are apparent in Fig. 1.4b for large α_{avg} , to a good degree of precision we observe that $\langle \theta_{\text{psp}} \rangle$ decreases at the constant rate -7.5 deg/dB for $0 \leq \alpha_{\text{avg}} \leq 4$ dB, with σ_{psp} increasing by approximately 4.5 deg/dB.

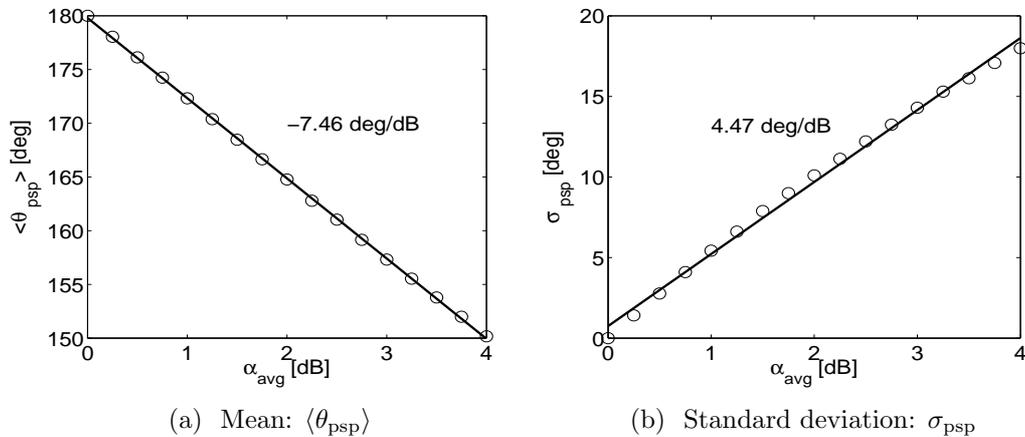


Figure 1.4: The variation of (a) the mean, $\langle \theta_{\text{psp}} \rangle$, and (b) the standard deviation, σ_{psp} , of θ_{psp} , Eq. (1.35), as a function of the mean PDL, α_{avg} , in a concatenation of 100 birefringent and lossy segments displayed here with \circ markers for $\tau_{\text{avg}} = 25$ ps. Optimal fits to linear functions are shown as solid lines.

1.3 Conclusions

We have discussed the physical origins of polarization mode dispersion in optical fibers and have introduced the mathematical models of PMD and PDL necessary for the development of later chapters. In particular, we have demonstrated that the Jones space principal states of polarization are non-orthogonal in the presence of PDL [50]. Indeed, over a range of mean PDL values relevant to optical network design the mean and standard deviation of the angle of PSP separation were shown to vary linearly with the system's average differential attenuation. In subsequent chapters we will develop a formalism based upon the Magnus expansion for analyzing the frequency dependence of the Jones or Mueller matrix in the presence of PMD and PDL. This formalism is relevant to, for example, numerical simulations of PMD and PDL induced pulse distortion, the design of joint PMD and PDL compensation devices and further leads to novel procedures for determining the PMD and PDL from measurements of the frequency dependent output Stokes vector.

The outline of this work is as follows. In Chapter 2, we formulate a Mueller matrix description of PMD and PDL in terms of the Magnus expansion. The accuracy of this model is established in Chapter 3 to fourth order in optical frequency through experiment and numerical simulation. Chapter 4 introduces Clifford algebraic techniques for the geometrical analysis of polarization transformations and further establishes a connection between PMD and PDL and the Lorentz transformation of the input state of polarization. A least-squares procedure based upon these Lorentz group properties for estimating the Mueller matrix of a fiber from repeated measurements of the output Stokes vector at adjacent optical frequencies is developed in Chapter 5 and is later applied to the high-speed measurement of PMD in optical systems in Chapter 7. We discuss in Chapter 6 efficient numerical techniques for estimating the frequency variation of the Jones matrix from its value at a minimal number of equally spaced optical frequencies and finally apply these methods to the determination of states with a worst case system penalty.

Chapter 2

Mueller Matrix Description of PMD and PDL

We now derive a differential equation that relates the Mueller matrices of an optical system at adjacent frequencies in the presence of PMD and PDL. We then demonstrate that a solution of this equation based on the Magnus expansion yields a description of the Mueller matrix in orders of the complex principal state vector that coincides with previously reported results for systems without PDL.

2.1 Introduction

In this chapter, the frequency variation of the output polarization of an optical system is parametrized in Stokes space in terms of the Taylor expansion coefficients of the complex principal state vector about a specified central frequency. Previously, such a representation has been only derived for the Jones matrices of optical systems with zero PDL, for which the Jones matrix is unitary and therefore corresponds to a three-dimensional rotation matrix. However, the relationship between the rotation angle and axis and the PMD vector is only given through a differential equation, Eq. (1.7) [45]. Considerable effort has therefore been

expended on inverting this relationship to express the 3×3 rotation matrix in terms of the PMD vector [59]. Below, we extend these analyses to systems with both PMD and PDL for which the transformation of polarization states in an optical system is described by a 4×4 Mueller matrix. The resulting formalism is applicable to any optical system described by a non-singular, frequency dependent Jones matrix.

Summarizing our main results, we demonstrate, to our knowledge for the first time, in Section 2.2 that the Mueller matrix, $\mathbf{M}(\omega)$, varies with frequency according to the differential equation

$$\frac{d\mathbf{M}}{d\omega}\mathbf{M}^{-1} = \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix}, \quad (2.1)$$

where the column vectors $\vec{\Omega}$ and $\vec{\Lambda}$ are the real and imaginary components of the complex principal state vector [50], respectively, and $\vec{\Lambda}^T$ is the transpose of $\vec{\Lambda}$. Next, in Section 2.3 we introduce the Magnus expansion, which provides an exponentiated series solution of Eq. (2.1), that, when truncated to a finite number of terms, preserves the relevant group properties of the exact solution [71, 75]. Applying this formalism yields the central result of Section 2.4, which expresses the Mueller matrix as a Taylor series in quantities directly related to the complex principal state vector. To third order in $\Delta\omega$ we find

$$\mathbf{M}(\omega) = \exp \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & \vec{b} \times \end{bmatrix} \mathbf{M}(\omega_0) \quad (2.2)$$

$$\begin{aligned} \vec{a} &= \vec{\Lambda}_0 \Delta\omega + \vec{\Lambda}_1 \frac{\Delta\omega^2}{2!} + \left(\vec{\Lambda}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Lambda}_1 - \frac{1}{2} \vec{\Lambda}_0 \times \vec{\Omega}_1 \right) \frac{\Delta\omega^3}{3!} + \dots \\ \vec{b} &= \vec{\Omega}_0 \Delta\omega + \vec{\Omega}_1 \frac{\Delta\omega^2}{2!} + \left(\vec{\Omega}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1 - \frac{1}{2} \vec{\Lambda}_1 \times \vec{\Lambda}_0 \right) \frac{\Delta\omega^3}{3!} + \dots \end{aligned}$$

In the above formula, $\vec{\Omega}_n$ and $\vec{\Lambda}_n$ are coefficients in the Taylor series expansions of $\vec{\Omega} = \vec{\Omega}(\omega)$ and $\vec{\Lambda} = \vec{\Lambda}(\omega)$, where $\omega = \omega_0 + \Delta\omega$, with ω_0 the optical carrier frequency. The corresponding Jones matrix representation is obtained through the relationship

$$\mathbf{T}(\omega) = \exp \left[-\frac{i}{2} (\vec{b} + i\vec{a}) \cdot \vec{\sigma} \right] \mathbf{T}(\omega_0), \quad (2.3)$$

where the vectors \vec{a} and \vec{b} are defined through Eq. (2.2). We then compare our results to those of previous authors [59] for the case of zero PDL. Finally, in Section 2.5 we discuss the application of the Mueller matrix formalism to the design of joint PMD and PDL compensators along the lines of Refs. [115, 118].

2.2 Theoretical background

As is well known, in a quasi-single mode optical system the input and output electric field polarizations, represented by the Jones vectors $|s\rangle$ and $|t(\omega)\rangle$, respectively, are related in the presence of both PMD and PDL by $|t(\omega)\rangle = \mathbf{T}(\omega)|s\rangle$, where $\mathbf{T}(\omega)$ denotes a complex 2×2 Jones matrix and $|s\rangle$ is frequency independent. The Jones matrix \mathbf{T} describes not only polarization evolution through the link, but also an overall, typically frequency, but not polarization, dependent attenuation, α_0 , and phase, ϕ_0 . Accordingly, we can write $\mathbf{T} = \exp(-\alpha_0 - i\phi_0)\mathbf{T}'$, in which the determinant of \mathbf{T}' is unity [50]. As the common attenuation and phase do not affect the physically interesting differential polarization behaviour, these are omitted from our subsequent discussion. We further omit primes so that \mathbf{T}' is replaced by \mathbf{T} .

The transformation between the input and output Stokes vectors, \tilde{s} and $\tilde{t}(\omega)$ is specified by the 4×4 matrix equation $\tilde{t}(\omega) = \mathbf{M}(\omega)\tilde{s}$. Further, if $\tilde{\sigma}$ signifies the Pauli spin vector extended by the 2×2 identity matrix,

$$\begin{aligned}\tilde{\sigma} &\equiv [\sigma_0, \vec{\sigma}] \\ &= [\sigma_0, \sigma_1, \sigma_2, \sigma_3]\end{aligned}\tag{2.4}$$

and

$$\begin{aligned}\sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \sigma_1 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \sigma_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \sigma_3 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\end{aligned}\tag{2.5}$$

the input and output four dimensional Stokes and Jones vectors satisfy [38] $\tilde{s} = \langle s | \tilde{\boldsymbol{\sigma}} | s \rangle$ and $\tilde{t} = \langle t | \tilde{\boldsymbol{\sigma}} | t \rangle$. The elements of the Mueller matrix \mathbf{M} are similarly defined by $m_{ij} = \text{Tr}(\boldsymbol{\sigma}_i \mathbf{T} \boldsymbol{\sigma}_j \mathbf{T}^\dagger) / 2$, $i, j = 0, 1, 2, 3$, and $\text{Tr}(\dots)$ indicates the trace, see Appendix B.3.

An alternative formulation expresses the Mueller matrix as the Kronecker matrix product [23]

$$\mathbf{M} = \mathbf{A} (\mathbf{T} \otimes \mathbf{T}^*) \mathbf{A}^\dagger, \quad (2.6)$$

where the unitary matrix \mathbf{A} equals

$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \quad (2.7)$$

and $*$, \dagger indicate complex and Hermitian conjugation, respectively. The Kronecker product maps a $M \times N$ matrix \mathbf{F} and a $P \times Q$ matrix \mathbf{G} into a $(MP) \times (NQ)$ block matrix. For example, for two 2×2 complex matrices,

$$\begin{aligned} \mathbf{F} \otimes \mathbf{G} &= \begin{bmatrix} f_{00} & f_{01} \\ f_{10} & f_{11} \end{bmatrix} \otimes \begin{bmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{bmatrix} \\ &= \begin{bmatrix} f_{00} \mathbf{G} & f_{01} \mathbf{G} \\ f_{10} \mathbf{G} & f_{11} \mathbf{G} \end{bmatrix} \\ &= \begin{bmatrix} f_{00}g_{00} & f_{00}g_{01} & f_{01}g_{00} & f_{01}g_{01} \\ f_{00}g_{10} & f_{00}g_{11} & f_{01}g_{10} & f_{01}g_{11} \\ f_{10}g_{00} & f_{10}g_{01} & f_{11}g_{00} & f_{11}g_{01} \\ f_{10}g_{10} & f_{10}g_{11} & f_{11}g_{10} & f_{11}g_{11} \end{bmatrix}. \end{aligned} \quad (2.8)$$

In general, the elements of $\mathbf{C} = \mathbf{F} \otimes \mathbf{G}$ are $c_{(Pi+k),(Qj+l)} = f_{ij}g_{kl}$, with $i = 0, 1, \dots, (M - 1)$, \dots , $l = 0, 1, \dots, (Q - 1)$ [110].

Eq. (2.6) can be conveniently derived from the component representation of the Mueller

matrix [4],

$$\begin{aligned}
m_{ij} &= \frac{1}{2} \text{Tr}(\boldsymbol{\sigma}_i \mathbf{T} \boldsymbol{\sigma}_j \mathbf{T}^\dagger) \\
&= \frac{1}{2} [\boldsymbol{\sigma}_i \mathbf{T}]_{mn} [\boldsymbol{\sigma}_j \mathbf{T}^\dagger]_{nm} \\
&= \frac{1}{2} [\boldsymbol{\sigma}_i]_{mp}(t_{pn}) [\boldsymbol{\sigma}_j]_{nq}(t_{qm})^\dagger \\
&= \frac{1}{2} [\boldsymbol{\sigma}_i^*]_{pm}(t_{pn})(t_{mq})^* [\boldsymbol{\sigma}_j]_{nq},
\end{aligned} \tag{2.9}$$

where $m, n, p, q \in \{0, 1\}$, $i, j \in \{0, 1, 2, 3\}$, repeated indices are implicitly summed and we have applied $\boldsymbol{\sigma}_i^\dagger = \boldsymbol{\sigma}_i$. Since $(t_{pn})(t_{mq})^*$ can be recast as a Kronecker matrix product according to $(\mathbf{T} \otimes \mathbf{T}^*)_{2p+m, 2n+q} = (t_{pn})(t_{mq})^*$, the last line of Eq. (2.9) is identical to

$$m_{ij} = \frac{1}{2} [\tilde{\boldsymbol{\sigma}}_i^*]_{2p+m} (\mathbf{T} \otimes \mathbf{T}^*)_{2p+m, 2n+q} [\tilde{\boldsymbol{\sigma}}_j]_{2n+q}, \tag{2.10}$$

where $\tilde{\boldsymbol{\sigma}}_i$ denotes the column vector formed from the elements of $\boldsymbol{\sigma}_i$, i.e.,

$$\tilde{\boldsymbol{\sigma}}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \tilde{\boldsymbol{\sigma}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \quad \tilde{\boldsymbol{\sigma}}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \tilde{\boldsymbol{\sigma}}_3 = \begin{bmatrix} 0 \\ -i \\ i \\ 0 \end{bmatrix}. \tag{2.11}$$

We now introduce the matrix $a_{ij} = [\tilde{\boldsymbol{\sigma}}_i^*]_j / \sqrt{2}$ such that Eq. (2.10) can be written as $m_{ij} = a_{i, 2p+m} (\mathbf{T} \otimes \mathbf{T}^*)_{2p+m, 2n+q} a_{2n+q, j}^\dagger$, or equivalently, $\mathbf{M} = \mathbf{A} (\mathbf{T} \otimes \mathbf{T}^*) \mathbf{A}^\dagger$.

Having recast the Mueller matrix as a Kronecker product we now examine the frequency derivative, denoted through the subscript ω , of the output Jones vector $|t\rangle_\omega = \mathbf{T}_\omega \mathbf{T}^{-1} |t\rangle$. The Jones space operator $\mathbf{T}_\omega \mathbf{T}^{-1}$ is a linear superposition of Pauli spin matrices [51],

$$\frac{d\mathbf{T}}{d\omega} \mathbf{T}^{-1} = -\frac{i}{2} \vec{W}(\omega) \cdot \vec{\boldsymbol{\sigma}}, \tag{2.12}$$

in which the vector $\vec{W}(\omega) \equiv \vec{\Omega}(\omega) + i\vec{\Lambda}(\omega)$ is termed the complex principal state [50]. To derive $\vec{W}(\omega)$ in terms of variables that characterize the Jones matrix $\mathbf{T}(\omega)$, we first define

$\mathbf{T}_0 = \mathbf{T}(\omega_0)$, $\lambda = \sqrt{|\vec{\beta}|^2 - |\vec{\alpha}|^2 + 2i\vec{\beta} \cdot \vec{\alpha}}$, and $\hat{w} = (\vec{\beta} + i\vec{\alpha})/\lambda$, for two real vectors $\vec{\beta}$ and $\vec{\alpha}$, such that

$$\begin{aligned} \mathbf{T}(\omega) &= \exp\left[-\frac{i}{2}(\vec{\beta} + i\vec{\alpha}) \cdot \vec{\sigma}\right] \mathbf{T}_0 \\ &= \left[\cos\left(\frac{\lambda}{2}\right) \sigma_0 - i \sin\left(\frac{\lambda}{2}\right) (\hat{w} \cdot \vec{\sigma})\right] \mathbf{T}_0. \end{aligned} \quad (2.13)$$

Here we have employed the Jones matrix representation of [10, 50], where the second expression in Eq. (2.13) follows after expanding the exponential into a power series and applying $(\hat{w} \cdot \vec{\sigma})^2 = \sigma_0$. Further, since $\hat{w}_\omega \cdot \hat{w} = 0$, $\mathbf{T}_\omega \mathbf{T}^{-1}$ can be written as

$$\mathbf{T}_\omega \mathbf{T}^{-1} = -\frac{i}{2} [\lambda_\omega \hat{w} + \sin(\lambda) \hat{w}_\omega + (1 - \cos(\lambda)) \hat{w} \times \hat{w}_\omega] \cdot \vec{\sigma}, \quad (2.14)$$

which, comparing with Eq. (2.12), yields [45] $\vec{W}(\omega) = \lambda_\omega \hat{w} + \sin(\lambda) \hat{w}_\omega + (1 - \cos(\lambda)) \hat{w} \times \hat{w}_\omega$. Our derivation and the resulting formulas for the dependence of the complex principal state vector, \vec{W} , on λ and \hat{w} can be applied so long as $\lambda \neq 0$, in which case an analogous calculation with $\mathbf{T} = \exp[\vec{\alpha} \cdot \vec{\sigma}/2] \exp[-i\vec{\beta} \cdot \vec{\sigma}/2]$ applies to all unit-determinant Jones matrices [10].

We can similarly express the Mueller matrix operator $\mathbf{M}_\omega \mathbf{M}^{-1}$ in terms of the real vectors $\vec{\Omega}$ and $\vec{\Lambda}$. The derivative of the Mueller matrix \mathbf{M} is calculated by differentiating Eq. (2.6). From the property of the Kronecker matrix product $(\mathbf{F} \otimes \mathbf{G})^{-1} = \mathbf{F}^{-1} \otimes \mathbf{G}^{-1}$, for nonsingular \mathbf{F} and \mathbf{G} , we note that $\mathbf{M}^{-1} = \mathbf{A} (\mathbf{T}^{-1} \otimes \mathbf{T}^{*-1}) \mathbf{A}^\dagger$. Further, applying $(\mathbf{F} \otimes \mathbf{G})(\mathbf{M} \otimes \mathbf{N}) = (\mathbf{F}\mathbf{M}) \otimes (\mathbf{G}\mathbf{N})$, yields

$$\begin{aligned} \mathbf{M}_\omega \mathbf{M}^{-1} &= \mathbf{A} [(\mathbf{T}_\omega \mathbf{T}^{-1}) \otimes (\mathbf{T}^* \mathbf{T}^{*-1}) + (\mathbf{T} \mathbf{T}^{-1}) \otimes (\mathbf{T}_\omega^* \mathbf{T}^{*-1})] \mathbf{A}^\dagger \\ &= -\frac{i}{2} \mathbf{A} [(\vec{\Omega} \cdot \vec{\sigma}) \otimes \sigma_0 - \sigma_0 \otimes (\vec{\Omega} \cdot \vec{\sigma})^*] \mathbf{A}^\dagger \\ &\quad + \frac{1}{2} \mathbf{A} [(\vec{\Lambda} \cdot \vec{\sigma}) \otimes \sigma_0 + \sigma_0 \otimes (\vec{\Lambda} \cdot \vec{\sigma})^*] \mathbf{A}^\dagger. \end{aligned} \quad (2.15)$$

Observing that

$$\begin{aligned}\mathbf{A}(\vec{a} \cdot \vec{\sigma}) \otimes \sigma_0 \mathbf{A}^\dagger &= \begin{bmatrix} 0 & \vec{a}^\top \\ \vec{a} & i\vec{a} \times \end{bmatrix} \\ \mathbf{A}\sigma_0 \otimes (\vec{a} \cdot \vec{\sigma})^* \mathbf{A}^\dagger &= \begin{bmatrix} 0 & \vec{a}^\top \\ \vec{a} & -i\vec{a} \times \end{bmatrix},\end{aligned}\quad (2.16)$$

for any real vector \vec{a} , we finally obtain

$$\frac{d\mathbf{M}}{d\omega} \mathbf{M}^{-1} = \begin{bmatrix} 0 & \vec{\Lambda}^\top \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} = \begin{bmatrix} 0 & \Lambda_x & \Lambda_y & \Lambda_z \\ \Lambda_x & 0 & -\Omega_z & \Omega_y \\ \Lambda_y & \Omega_z & 0 & -\Omega_x \\ \Lambda_z & -\Omega_y & \Omega_x & 0 \end{bmatrix}.\quad (2.17)$$

In the absence of PDL we recover the standard expression [45] $\mathbf{R}_\omega \mathbf{R}^{-1} = \vec{\Omega} \times$, for some 3×3 Stokes rotation matrix, \mathbf{R} .

2.2.1 Normalized stokes vectors

Designating the four-dimensional output Stokes vector \tilde{t} by $\tilde{t} = (|\tilde{t}|, \vec{t})$, with $\vec{t} = \langle t | \vec{\sigma} | t \rangle$, the frequency derivative of \tilde{t} follows from Eq. (2.17) in that

$$\begin{aligned}\frac{d\tilde{t}}{d\omega} &= \begin{bmatrix} |\tilde{t}|_\omega \\ |\tilde{t}|_\omega \hat{t} + |\tilde{t}| \hat{t}_\omega \end{bmatrix} \\ &= \begin{bmatrix} 0 & \vec{\Lambda}^\top \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} \begin{bmatrix} |\tilde{t}| \\ \vec{t} \end{bmatrix} \\ &= \begin{bmatrix} |\tilde{t}| \vec{\Lambda} \cdot \hat{t} \\ \vec{\Lambda} |\tilde{t}| + |\tilde{t}| \vec{\Omega} \times \hat{t} \end{bmatrix},\end{aligned}\quad (2.18)$$

where $\vec{t} = |\tilde{t}| \hat{t}$. Substituting our expression for $|\tilde{t}|_\omega$ into $\hat{t} |\tilde{t}|_\omega + |\tilde{t}| \hat{t}_\omega$, we arrive at

$$\begin{aligned}\hat{t}_\omega &= \vec{\Lambda} - \hat{t}(\vec{\Lambda} \cdot \hat{t}) + \vec{\Omega} \times \hat{t} \\ &= \hat{t} \times (\vec{\Lambda} \times \hat{t}) + \vec{\Omega} \times \hat{t},\end{aligned}\quad (2.19)$$

which is often derived directly from the equation of motion of the Stokes vector [32, 62]
 $\hat{t} = \langle t | \vec{\sigma} | t \rangle / \langle t | t \rangle$.

2.3 Magnus expansion

The variation of the Mueller matrix with frequency is described by Eq. (2.17), which we write as

$$\frac{d\mathbf{M}}{d\omega} = \mathbf{H}(\omega)\mathbf{M}, \quad (2.20)$$

together with the 4×4 matrix initial condition $\mathbf{M}(\omega_0) = \mathbf{M}_0$. Denoting the frequency ordering operator by F_ω , the solution can be expressed as the ω -ordered product [25, 42]

$$\mathbf{M}(\omega) = F_\omega \exp \left[\int_{\omega_0}^{\omega} \mathbf{H}(\omega_1) d\omega_1 \right] \mathbf{M}_0. \quad (2.21)$$

While the ω -ordering is obviously superfluous for frequency-independent operators \mathbf{H} , as is the case if second and higher order PMD and PDL effects are absent, the general solution is conveniently expressed through the Magnus expansion, see Appendix A.2 [43, 71, 75, 111]

$$\mathbf{M}(\omega) = \exp \left[\sum_{n=1}^{\infty} \mathbf{B}_n(\omega) \right] \mathbf{M}_0, \quad (2.22)$$

where

$$\begin{aligned} \mathbf{B}_1(\omega) &= \int_{\omega_0}^{\omega} d\omega_1 \mathbf{H}(\omega_1) \\ \mathbf{B}_2(\omega) &= \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 [\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)] \\ \mathbf{B}_3(\omega) &= \frac{1}{6} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \int_{\omega_0}^{\omega_2} d\omega_3 ([\mathbf{H}(\omega_1), \\ &\quad [\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)]] + [[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)], \mathbf{H}(\omega_3)]), \end{aligned} \quad (2.23)$$

and $[\mathbf{F}, \mathbf{G}]$ is the commutator of the matrices \mathbf{F} and \mathbf{G} . The coefficients \mathbf{B}_n obey the recurrence relation,

$$\mathbf{B}_n(\omega) = \sum_{j=1}^{n-1} \frac{b_j}{j!} \int_{\omega_0}^{\omega} d\omega_1 \mathbf{S}_n^{(j)}(\omega_1), \quad n \geq 2 \quad (2.24)$$

in which b_j denotes the j :th Bernoulli numbers [75] and

$$\begin{aligned} \mathbf{S}_n^{(j)}(\omega) &= \sum_{m=1}^{n-j} \left[\mathbf{B}_m(\omega), \mathbf{S}_{n-m}^{(j-1)}(\omega) \right], \quad 2 \leq j \leq n-1 \\ \mathbf{S}_n^{(1)}(\omega) &= [\mathbf{B}_{n-1}(\omega), \mathbf{H}(\omega)], \\ \mathbf{S}_n^{(n-1)}(\omega) &= \left\{ \mathbf{B}_1^{(n-1)}(\omega), \mathbf{H}(\omega) \right\}. \end{aligned} \tag{2.25}$$

Here $\{\mathbf{F}^{(n)}, \mathbf{G}\}$ is the nested commutator $\{\mathbf{F}^{(n)}, \mathbf{G}\} = [\mathbf{F}, \{\mathbf{F}^{(n-1)}, \mathbf{G}\}]$ in which \mathbf{F} appears n times and $\{\mathbf{F}^{(0)}, \mathbf{G}\} = \mathbf{G}$. The Magnus expansion can also be derived by expanding the integrand of Eq. (2.21) into a Taylor series and reexponentiating the resulting terms [118].

2.4 Mueller matrix formalism

To construct the Mueller matrix $\mathbf{M}(\omega)$ in Eq. (2.2) as the exponential of a power series in $\Delta\omega$, we perform a Taylor expansion of the vectors $\vec{\Lambda}(\omega)$ and $\vec{\Omega}(\omega)$. Eq. (2.22) then generates formulas for the frequency dependence of each Taylor order of the Mueller matrix \mathbf{M} that can be compared to those of previous authors. This procedure further leads to a simple relationship between the Mueller and Jones matrix representations.

We first decompose the evolution operator \mathbf{H} into Hermitian and anti-Hermitian components. For convenience, we define the vector operators \mathbf{h} and \mathbf{c} by

$$\begin{aligned} \mathbf{H}(\omega) &= \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} \\ &\equiv \mathbf{h}(\vec{\Lambda}) + \mathbf{c}(\vec{\Omega}) \end{aligned} \tag{2.26}$$

with

$$\mathbf{h}(\vec{\Lambda}) = \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & 0 \end{bmatrix} \quad \mathbf{c}(\vec{\Omega}) = \begin{bmatrix} 0 & 0 \\ 0 & \vec{\Omega} \times \end{bmatrix}. \tag{2.27}$$

Developing $\vec{\Lambda}(\omega_0 + \Delta\omega)$ and $\vec{\Omega}(\omega_0 + \Delta\omega)$ in Eq. (2.23) in Taylor series yields for the first term in the Magnus expansion

$$\begin{aligned} \mathbf{B}_1 &= \mathbf{c}(\vec{\Omega}_0\Delta\omega + \vec{\Omega}_1\frac{\Delta\omega^2}{2!} + \vec{\Omega}_2\frac{\Delta\omega^3}{3!}) \\ &+ \mathbf{h}(\vec{\Lambda}_0\Delta\omega + \vec{\Lambda}_1\frac{\Delta\omega^2}{2!} + \vec{\Lambda}_2\frac{\Delta\omega^3}{3!}) + O(\Delta\omega^4). \end{aligned} \quad (2.28)$$

With the redefinition $\omega_i \rightarrow \omega_i - \omega_0$, the second term contains the commutators

$$\begin{aligned} \mathbf{B}_2 &= \frac{1}{2} \int_0^{\Delta\omega} d\omega_1 \int_0^{\omega_1} d\omega_2 [\mathbf{c}(\vec{\Omega}_0 + \vec{\Omega}_1\omega_1) + \mathbf{h}(\vec{\Lambda}_0 + \vec{\Lambda}_1\omega_1), \\ &\quad \mathbf{c}(\vec{\Omega}_0 + \vec{\Omega}_1\omega_2) + \mathbf{h}(\vec{\Lambda}_0 + \vec{\Lambda}_1\omega_2)], \end{aligned} \quad (2.29)$$

in which we have retained contributions up to and including third order in $\Delta\omega$ after integration. From Eq. (2.27) we obtain the commutator identities

$$\begin{aligned} [\mathbf{c}(\vec{u}), \mathbf{h}(\vec{v})] &= \mathbf{h}(\vec{u} \times \vec{v}) \\ [\mathbf{c}(\vec{u}), \mathbf{c}(\vec{v})] &= \mathbf{c}(\vec{u} \times \vec{v}) \\ [\mathbf{h}(\vec{u}), \mathbf{h}(\vec{v})] &= \mathbf{c}(\vec{v} \times \vec{u}) \end{aligned} \quad (2.30)$$

for the 4×4 matrices that are generated when the operators \mathbf{h} and \mathbf{c} are applied to two vectors \vec{u} and \vec{v} . Inserting the above formulas into Eq. (2.29) yields

$$\mathbf{B}_2 = -\frac{1}{2} \frac{\Delta\omega^3}{3!} \left\{ \mathbf{c}(\vec{\Omega}_0 \times \vec{\Omega}_1 + \vec{\Lambda}_1 \times \vec{\Lambda}_0) + \mathbf{h}(\vec{\Omega}_0 \times \vec{\Lambda}_1 + \vec{\Lambda}_0 \times \vec{\Omega}_1) \right\} + O(\Delta\omega^4). \quad (2.31)$$

Finally, after substituting \mathbf{B}_1 and \mathbf{B}_2 into Eq. (2.22) we arrive at our central result, namely,

$$\mathbf{M}(\omega) = \exp \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & \vec{b} \times \end{bmatrix} \mathbf{M}_0 \quad (2.32)$$

$$\begin{aligned} \vec{a} &= \vec{\Lambda}_0\Delta\omega + \vec{\Lambda}_1\frac{\Delta\omega^2}{2!} + \left(\vec{\Lambda}_2 - \frac{1}{2}\vec{\Omega}_0 \times \vec{\Lambda}_1 - \frac{1}{2}\vec{\Lambda}_0 \times \vec{\Omega}_1 \right) \frac{\Delta\omega^3}{3!} + O(\Delta\omega^4) \\ \vec{b} &= \vec{\Omega}_0\Delta\omega + \vec{\Omega}_1\frac{\Delta\omega^2}{2!} + \left(\vec{\Omega}_2 - \frac{1}{2}\vec{\Omega}_0 \times \vec{\Omega}_1 - \frac{1}{2}\vec{\Lambda}_1 \times \vec{\Lambda}_0 \right) \frac{\Delta\omega^3}{3!} + O(\Delta\omega^4), \end{aligned}$$

which describes the frequency evolution of the Mueller matrix to third order in $\Delta\omega$. Higher order corrections can be obtained from the recursion relation, Eq. (2.24), for the Magnus coefficients.

To convert between the Mueller and Jones formalisms, we derive the Jones matrix analog of Eq. (2.32) by noting that each Jones matrix of the form $\mathbf{T} = \exp[\mathbf{K}] \mathbf{T}_0$, for a 2×2 complex matrix \mathbf{K} , maps to a Mueller matrix $\mathbf{M} = \exp[\mathbf{A}(\mathbf{K} \otimes \boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_0 \otimes \mathbf{K}^*)\mathbf{A}^\dagger] \mathbf{M}_0$ with \mathbf{A} given by Eq. (2.7) [8]. Eq. (2.16) in conjunction with Eq. (2.32) then yields,

$$\mathbf{T}(\omega) = \exp\left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\boldsymbol{\sigma}}\right] \mathbf{T}_0, \quad (2.33)$$

in which the vectors \vec{a} and \vec{b} are defined in Eq. (2.32).

2.4.1 Comparison with previous work

To verify that we can recover the results of previous authors we consider the case of zero PDL. In terms of the 4×4 matrix \mathbf{N}_c related to the vectors \vec{a} and \vec{b} through

$$\mathbf{N}_c = \begin{bmatrix} 0 & (\vec{a} + i\vec{b})^T \\ \vec{a} + i\vec{b} & -i(\vec{a} + i\vec{b}) \times \end{bmatrix}, \quad (2.34)$$

and the complex variable $\eta = \sqrt{|\vec{a}|^2 - |\vec{b}|^2 + 2i\vec{a} \cdot \vec{b}}$, the exponentiated Mueller matrix, Eq. (2.32), becomes

$$\mathbf{M}(\omega) = \exp\begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & \vec{b} \times \end{bmatrix} \mathbf{M}_0 = (\mathbf{F}_c \mathbf{F}_c^*) \mathbf{M}_0 \quad (2.35)$$

where

$$\mathbf{F}_c \equiv \exp\left[\frac{\mathbf{N}_c}{2}\right] = \cosh\left(\frac{\eta}{2}\right) \mathbf{I}_4 + \frac{1}{\eta} \sinh\left(\frac{\eta}{2}\right) \mathbf{N}_c \quad (2.36)$$

and \mathbf{I}_N denotes the $N \times N$ identity matrix. To derive the above formulas, the exponent of Eq. (2.35) is recast as $(\mathbf{N}_c + \mathbf{N}_c^*)/2$, with $[\mathbf{N}_c, \mathbf{N}_c^*] = 0$. Subsequently, the identities of e.g. Ref. [46] are applied, in which $\exp(\mathbf{N}_c/2)$ is expanded in a power series and the result simplified according to $\mathbf{N}_c^2 = \eta^2 \mathbf{I}_4$.

If either $|\vec{a}|$ or $|\vec{b}|$ is zero, Eq. (2.35) can be simplified considerably. Setting $\vec{a} = a\hat{a}$ and $\vec{b} = b\hat{b}$, we find respectively

$$\exp \begin{bmatrix} 0 & 0 \\ 0 & \vec{b} \times \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \cos(b)\mathbf{I}_3 + \sin(b)(\hat{b} \times) + (1 - \cos(b))\hat{b}\hat{b}^T \end{bmatrix} \quad (2.37)$$

$$\exp \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & 0 \end{bmatrix} = \begin{bmatrix} \cosh(a) & \sinh(a)\hat{a}^T \\ \sinh(a)\hat{a} & (\cosh(a) - 1)\hat{a}\hat{a}^T + \mathbf{I}_3 \end{bmatrix}. \quad (2.38)$$

As expected, the 3×3 submatrix in the lower right hand corner of Eq. (2.37) is identical to $\exp(\vec{b} \times)$, c.f. Appendix B.4. Next, in view of Eq. (2.37), the Mueller matrix of Eq. (2.32) reduces to

$$\mathbf{M}(\omega) = \begin{bmatrix} 1 & 0 \\ 0 & \exp \left[\left(\vec{\Omega}_0 \Delta\omega + \vec{\Omega}_1 \frac{\Delta\omega^2}{2!} + \left(\vec{\Omega}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1 \right) \frac{\Delta\omega^3}{3!} \right) \times \right] \end{bmatrix} \mathbf{M}_0 \quad (2.39)$$

for systems with zero PDL.

We now establish that our method reproduces the PMD vector identities of Refs. [45,59] in the absence of PDL. These were obtained by repeatedly differentiating

$$\vec{\Omega} = \frac{d\psi}{d\omega} \hat{n} + \sin(\psi) \frac{d\hat{n}}{d\omega} + (1 - \cos(\psi)) \hat{n} \times \frac{d\hat{n}}{d\omega}, \quad (2.40)$$

where ψ and \hat{n} can be determined from the 3×3 Stokes matrix through $\mathbf{R}(\Delta\omega) = e^{\psi \hat{n} \times}$, and then evaluating the result at the optical carrier frequency ($\psi(\Delta\omega = 0) = 0$). This yields an infinite series of relationships between derivatives of ψ and \hat{n} and the PMD vectors of different orders. The first three of these are

$$\begin{aligned} \vec{\Omega}_0 &= \psi_1 \hat{n}_0 = \left. \frac{d(\psi \hat{n})}{d\omega} \right|_{\omega_0} \\ \vec{\Omega}_1 &= \psi_2 \hat{n}_0 + 2\psi_1 \hat{n}_1 = \left. \frac{d^2(\psi \hat{n})}{d\omega^2} \right|_{\omega_0} \\ \vec{\Omega}_2 &= \psi_3 \hat{n}_0 + 3\psi_2 \hat{n}_1 + 3\psi_1 \hat{n}_2 + \psi_1^2 \hat{n}_0 \times \hat{n}_1 \\ &= \left. \frac{d^3(\psi \hat{n})}{d\omega^3} \right|_{\omega_0} + \psi_1^2 \hat{n}_0 \times \hat{n}_1 \end{aligned} \quad (2.41)$$

where $A_n = \frac{d^n A}{d\omega^n} \Big|_{\omega_0}$. Note however that if the exponent of the Stokes matrix $\mathbf{R} = e^{\psi \hat{n} \times}$ is expanded in a power series to third order in frequency, the identities in Eq. (2.41), together with $\psi_1^2 \hat{n}_0 \times \hat{n}_1 = \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1$, and

$$\frac{d^3(\psi \hat{n})}{d\omega^3} \Big|_{\omega_0} = \vec{\Omega}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1, \quad (2.42)$$

lead to

$$\mathbf{R} = \exp \left[\left(\vec{\Omega}_0 \Delta\omega + \vec{\Omega}_1 \frac{\Delta\omega^2}{2!} + \left(\vec{\Omega}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1 \right) \frac{\Delta\omega^3}{3!} \right) \times \right], \quad (2.43)$$

in agreement with the 3×3 submatrix in the lower right hand corner of our Magnus expansion result, Eq. (2.39).

2.5 Operator symmetrization

The Jones or Mueller matrices corresponding to PMD and PDL compensators can be recast as products of exponential operators each of which corresponds to a realizable optical component [31, 60]. Because these operators are non-commuting, however, they must be properly symmetrized in order to maximize the compensation bandwidth [115, 118].

Below, we will demonstrate that Eq. (2.32) together with the Baker-Campbell-Hausdorff (BCH) identity [43],

$$e^{\mathbf{F}} e^{\mathbf{G}} = \exp \left[\mathbf{F} + \mathbf{G} + \frac{1}{2} [\mathbf{F}, \mathbf{G}] + \frac{1}{12} [\mathbf{F}, [\mathbf{F}, \mathbf{G}]] + \frac{1}{12} [\mathbf{G}, [\mathbf{G}, \mathbf{F}]] + \dots \right] \quad (2.44)$$

for two matrices \mathbf{F} and \mathbf{G} , enables the design of PMD-PDL compensators that invert the frequency dependence of an arbitrary Mueller matrix to any desired order in $\Delta\omega$. For example, to third order in $\Delta\omega$, the BCH identity can be employed to establish that the five operator approximation

$$\mathbf{M}_{comp} = e^{-\frac{1}{2} \mathbf{N}_{(3)} \Delta\omega^2} e^{-\frac{1}{2} \mathbf{N}_{(2)} \Delta\omega^2} e^{-\frac{1}{2} \mathbf{N}_{(1)} \Delta\omega} e^{-\mathbf{N}_{(0)} \Delta\omega} e^{-\frac{1}{2} \mathbf{N}_{(1)} \Delta\omega} \quad (2.45)$$

with

$$\begin{aligned} \mathbf{N}_{(0)} &= \begin{bmatrix} 0 & 0 \\ 0 & \vec{\Omega}_0 \times \end{bmatrix} & \mathbf{N}_{(1)} &= \begin{bmatrix} 0 & \vec{\Lambda}_0^T \\ \vec{\Lambda}_0 & 0 \end{bmatrix} \\ \mathbf{N}_{(2)} &= \begin{bmatrix} 0 & 0 \\ 0 & \vec{\Omega}_1 \times \end{bmatrix} & \mathbf{N}_{(3)} &= \begin{bmatrix} 0 & \vec{\Lambda}_1^T \\ \vec{\Lambda}_1 & 0 \end{bmatrix} \end{aligned} \quad (2.46)$$

corresponds to the exact inverse of the frequency-dependent Mueller matrix to $O(\Delta\omega^3)$. That is, the above formalism yields a compensator Mueller matrix [115], \mathbf{M}_{comp} , that compensates PMD and PDL to $O(\Delta\omega^3)$.

The extension of the above formalism to higher-order compensators can be simply achieved by writing Eq. (2.32) as

$$\mathbf{M}(\omega) = \exp \left[\mathbf{H}_0 \Delta\omega + \mathbf{H}_1 \frac{\Delta\omega^2}{2!} + \left(\mathbf{H}_2 - \frac{1}{2} [\mathbf{H}_0, \mathbf{H}_1] \right) \frac{\Delta\omega^3}{3!} + O(\Delta\omega^4) \right] \mathbf{M}_0. \quad (2.47)$$

The quantity $\mathbf{H}(\omega)$ is given by Eq. (2.26) while $\mathbf{H}_n = \frac{d^n \mathbf{H}}{d\omega^n} |_{\omega_0}$. To fourth order, we then find for the inverse of the frequency-dependent Mueller matrix, after repeatedly applying the BCH identity and comparing the results to the Magnus expansion, Eq. (2.47), see Appendix A.3,

$$\mathbf{M}_{comp} = e^{-\frac{1}{6} \mathbf{H}_2 \Delta\omega^3} e^{-\frac{2}{3} \mathbf{H}_0 \Delta\omega} e^{-\frac{1}{2} \mathbf{H}_1 \Delta\omega^2} e^{-\frac{1}{3} \mathbf{H}_0 \Delta\omega}. \quad (2.48)$$

Each operator in Eq. (2.48) can now be decomposed as in Eq. (2.45) into a product of exponentials each of which corresponds to a n :th order PMD or PDL element.

The extension to higher order compensation is straightforward if tedious [43,44] as the recursive form of the Magnus coefficients, Eq. (2.24), can be employed to approximate the Mueller matrix as a product of exponential operators up to any specified order.

2.6 Conclusions

We have developed a general formalism for the Mueller matrix $(d\mathbf{M}/d\omega)\mathbf{M}^{-1}$ in the presence of both PMD and PDL. Through the Magnus expansion we then obtained a recursive

method for calculating successive orders of the frequency variation of the Mueller matrix. Finally, operator symmetrization yielded PMD/PDL compensators that counteract the effects of PMD and PDL with increased accuracy. While we have in this chapter restricted our attention to joint PMD and PDL compensation, our formalism is equally relevant to, for example, numerical simulations of PMD and PDL induced pulse distortion and further leads to novel procedures for determining the PMD and PDL from measurements of the frequency dependent output Stokes vector. In Chapter 3, we therefore implement these results both numerically and experimentally, and establish the accuracy of the procedure to at least fourth order in optical frequency.

Chapter 3

The Magnus Expansion for PMD and PDL

This chapter reviews a solution method for the Magnus expansion of the differential equation for the frequency dependence of the Mueller matrix in the presence of polarization mode dispersion (PMD) and polarization dependent loss (PDL). The solution is then compared with the results of previous authors and its accuracy is established to fourth order in frequency.

3.1 Introduction

If an optical fiber is approximated by a large number of constant birefringent segments, the light polarization in each segment precesses around the axis of the segment's polarization mode dispersion vector at a rate determined by the magnitude of the differential group delay, c.f. Section 1.1.2. The resultant polarization vector therefore varies rapidly with frequency. In Stokes space in the absence of PDL, the frequency dependence of the rotation angle and axis can be obtained by solving Eq. (1.11) for the 3×3 rotation matrix relating the input and output Stokes vectors [45]. While equations have previously been derived

for this rotation matrix in terms of PMD vectors of increasing order [59], we here present a more general analysis of the 4×4 Mueller matrix transformation for systems with both PMD and PDL. Our approach is a generalization of the formalism of Chapter 2 in that we establish the equivalence between results generated from the Magnus expansion solution of the underlying differential equation and those given by earlier models of systems without PDL [45, 48, 59, 88].

This chapter is organized as follows. First, we review our solution of the differential equation for the frequency dependence of the Mueller matrix based upon the Magnus expansion [50, 88]. We then demonstrate that our procedure agrees with a generalized version of the method of Refs. [45, 59] to at least third order in frequency, and also establish the equivalence of our results to the power series expansions of Ref. [48]. We finally discuss the applicability of our technique in both numerical and experimental contexts.

3.2 Mueller matrix formalism

In an optical system characterized by two orthogonally polarized guided modes with different group velocities and losses (PMD and PDL), the Jones vectors of the output and the frequency-independent input electric field polarizations, $|t(\omega)\rangle$ and $|s\rangle$, are related through a complex 2×2 Jones matrix according to $|t(\omega)\rangle = \mathbf{T}(\omega)|s\rangle$. (As we are interested only in the difference between the attenuation and phase of the two polarizations, we normalize $\det(\mathbf{T}) = 1$ [50]). Analogously, the 4 component input and output Stokes vectors given by $\tilde{s} = \langle s|\tilde{\boldsymbol{\sigma}}|s\rangle$ and $\tilde{t} = \langle t|\tilde{\boldsymbol{\sigma}}|t\rangle$ satisfy $\tilde{t}(\omega) = \mathbf{M}(\omega)\tilde{s}$ where $\tilde{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}_0, \vec{\boldsymbol{\sigma}})$, and $\boldsymbol{\sigma}_0, \vec{\boldsymbol{\sigma}}$ denote the 2×2 identity matrix and the Pauli spin matrices, respectively. The Mueller matrix, \mathbf{M} , can be written in component form as $M_{ij} = \frac{1}{2}\text{Tr}(\boldsymbol{\sigma}_i\mathbf{T}\boldsymbol{\sigma}_j\mathbf{T}^\dagger)$ for $i, j \in \{0, 1, 2, 3\}$ or alternatively, as the Kronecker matrix product

$$\mathbf{M} = \mathbf{A} (\mathbf{T} \otimes \mathbf{T}^*) \mathbf{A}^\dagger \quad (3.1)$$

in which the unitary matrix \mathbf{A} equals

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}, \quad (3.2)$$

and $*$, \dagger , indicate complex and Hermitian conjugation, respectively [23].

In the presence of both PMD and PDL, the Jones matrix $\mathbf{T}_\omega \mathbf{T}^{-1} = -i(\vec{W} \cdot \vec{\sigma})/2$, in which the subscript ω indicates differentiation with respect to the frequency variable, is a linear superposition of the Pauli spin matrices [50, 51], where $\vec{W} = \vec{\Omega} + i\vec{\Lambda}$ for real vectors $\vec{\Omega}$ and $\vec{\Lambda}$, c.f. Sec. 2.2. Introducing the notation $\mathbf{H}(\omega)$ for the evolution operator, the analogous Mueller matrix expression $\mathbf{H}(\omega) = \mathbf{M}_\omega \mathbf{M}^{-1}$ can be derived from the Kronecker product form of Eq. (3.1) [88]. This yields

$$\mathbf{H}(\omega) = \frac{d\mathbf{M}}{d\omega} \mathbf{M}^{-1} = \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} \quad (3.3)$$

in which the 4×4 transfer matrix \mathbf{H} is expressed in block matrix form.

The frequency dependence of the Mueller matrix can now be efficiently computed from Eq. (3.3) with the Magnus expansion, which provides a general solution of the form [71, 75]

$$\mathbf{M}(\omega) = \exp \left[\sum_{n=1}^{\infty} \mathbf{B}_n(\omega) \right] \mathbf{M}(\omega_0). \quad (3.4)$$

Denoting commutators with brackets, the first four expansion coefficients are then [76]

$$\begin{aligned}
\mathbf{B}_1(\omega) &= \int_{\omega_0}^{\omega} d\omega_1 \mathbf{H}(\omega_1) \\
\mathbf{B}_2(\omega) &= \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 [\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)], \\
\mathbf{B}_3(\omega) &= \frac{1}{6} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \int_{\omega_0}^{\omega_2} d\omega_3 \left([\mathbf{H}(\omega_1), [\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)]] \right. \\
&\quad \left. + [[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)], \mathbf{H}(\omega_3)] \right), \\
\mathbf{B}_4(\omega) &= \frac{1}{12} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \int_{\omega_0}^{\omega_2} d\omega_3 \int_{\omega_0}^{\omega_3} d\omega_4 \left([\mathbf{H}(\omega_1), [[\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)], \mathbf{H}(\omega_4)]] \right. \\
&\quad \left. + [[\mathbf{H}(\omega_1), [\mathbf{H}(\omega_2), \mathbf{H}(\omega_3)]], \mathbf{H}(\omega_4)] + [[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)], [\mathbf{H}(\omega_3), \mathbf{H}(\omega_4)]] \right)
\end{aligned} \tag{3.5}$$

where $[\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)]$ represents the matrix commutator of $\mathbf{H}(\omega_1)$ and $\mathbf{H}(\omega_2)$. Further, each \mathbf{B}_n is related to the coefficients $\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{(n-1)}$ through recursion relations [75]. Expanding $\mathbf{H}(\omega)$ into a Taylor series about the optical carrier, ω_0 , and evaluating the coefficients $\mathbf{B}_1, \dots, \mathbf{B}_4$ immediately leads to our central result, which to fifth order in the deviation, $\Delta\omega$, of the optical frequency from ω_0 is

$$\mathbf{M}(\omega) = e^{\mathbf{N}(\Delta\omega)} \mathbf{M}(\omega_0) \tag{3.6}$$

with

$$\begin{aligned}
\mathbf{N}(\Delta\omega) &= \mathbf{H}_0 \Delta\omega + \mathbf{H}_1 \frac{\Delta\omega^2}{2!} \\
&\quad + \left(\mathbf{H}_2 - \frac{1}{2} [\mathbf{H}_0, \mathbf{H}_1] \right) \frac{\Delta\omega^3}{3!} \\
&\quad + \left(\mathbf{H}_3 - [\mathbf{H}_0, \mathbf{H}_2] \right) \frac{\Delta\omega^4}{4!} \\
&\quad + \left(\mathbf{H}_4 - [\mathbf{H}_1, \mathbf{H}_2] - \frac{3}{2} [\mathbf{H}_0, \mathbf{H}_3] + \frac{1}{6} [\mathbf{H}_0, [\mathbf{H}_0, \mathbf{H}_2]] \right. \\
&\quad \left. + \frac{1}{2} [\mathbf{H}_1, [\mathbf{H}_1, \mathbf{H}_0]] + \frac{1}{6} [\mathbf{H}_0, [\mathbf{H}_0, [\mathbf{H}_0, \mathbf{H}_1]]] \right) \frac{\Delta\omega^5}{5!} \\
&\quad + O(\Delta\omega^6).
\end{aligned} \tag{3.7}$$

In the above equations, the subscript n denotes $\frac{d^n}{d\omega^n}|_{\omega_0}$. Corrections of order greater than 6 can similarly be calculated from the recursive form of the Magnus coefficients.

Alternatively, Eq. (3.3) can be employed to derive an expression for the Mueller matrix directly in terms of the vectors $\vec{\Omega}$ and $\vec{\Lambda}$. To fourth order, this yields

$$\mathbf{M}(\omega) = \exp \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & \vec{b} \times \end{bmatrix} \mathbf{M}(\omega_0). \quad (3.8)$$

Here \vec{a} and \vec{b} designate

$$\begin{aligned} \vec{a} &= \vec{\Lambda}_0 \Delta\omega + \vec{\Lambda}_1 \frac{\Delta\omega^2}{2!} \\ &+ \left(\vec{\Lambda}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Lambda}_1 - \frac{1}{2} \vec{\Lambda}_0 \times \vec{\Omega}_1 \right) \frac{\Delta\omega^3}{3!} \\ &+ \left(\vec{\Lambda}_3 - \vec{\Omega}_0 \times \vec{\Lambda}_2 - \vec{\Lambda}_0 \times \vec{\Omega}_2 \right) \frac{\Delta\omega^4}{4!} + O(\Delta\omega^5) \end{aligned} \quad (3.9)$$

and

$$\begin{aligned} \vec{b} &= \vec{\Omega}_0 \Delta\omega + \vec{\Omega}_1 \frac{\Delta\omega^2}{2!} \\ &+ \left(\vec{\Omega}_2 - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_1 - \frac{1}{2} \vec{\Lambda}_1 \times \vec{\Lambda}_0 \right) \frac{\Delta\omega^3}{3!} \\ &+ \left(\vec{\Omega}_3 - \vec{\Omega}_0 \times \vec{\Omega}_2 - \vec{\Lambda}_2 \times \vec{\Lambda}_0 \right) \frac{\Delta\omega^4}{4!} + O(\Delta\omega^5). \end{aligned} \quad (3.10)$$

The equivalent Jones matrix representation then follows from Eq. (3.1) [8], namely,

$$\mathbf{T}(\omega) = \exp \left[-\frac{i}{2} (\vec{b} + i\vec{a}) \cdot \vec{\sigma} \right] \mathbf{T}(\omega_0). \quad (3.11)$$

3.3 Alternate Mueller matrix derivation

To demonstrate that the results of Eq. (3.6) and Refs. [45, 59] coincide, we generalize the procedure first advanced in Ref. [88] by including PDL. In this method, $\vec{W}(\omega)$ is first expressed in terms of variables that characterize the Jones matrix $\mathbf{T}(\omega)$. Employing

the Jones matrix representation of Ref. [50], we write $\lambda = \sqrt{|\vec{\beta}|^2 - |\vec{\alpha}|^2 + 2i\vec{\beta} \cdot \vec{\alpha}}$, and $\hat{w} = (\vec{\beta} + i\vec{\alpha})/\lambda$, for two real vectors $\vec{\beta}$ and $\vec{\alpha}$, such that

$$\begin{aligned} \mathbf{T}(\omega) &= \exp \left[-\frac{i}{2}(\vec{\beta} + i\vec{\alpha}) \cdot \vec{\sigma} \right] \mathbf{T}(\omega_0) \\ &= \left[\cos \left(\frac{\lambda}{2} \right) \sigma_0 - i \sin \left(\frac{\lambda}{2} \right) (\hat{w} \cdot \vec{\sigma}) \right] \mathbf{T}(\omega_0). \end{aligned} \quad (3.12)$$

Here we have replaced the exponential by its power series and applied $(\hat{w} \cdot \vec{\sigma})^2 = \sigma_0$. Further, $\hat{w}_\omega \cdot \hat{w} = 0$, so that $\mathbf{T}_\omega \mathbf{T}^{-1}$ can be expanded as

$$\frac{d\mathbf{T}}{d\omega} \mathbf{T}^{-1} = -\frac{i}{2} \left[\frac{d\lambda}{d\omega} \hat{w} + \sin(\lambda) \frac{d\hat{w}}{d\omega} + (1 - \cos(\lambda)) \hat{w} \times \frac{d\hat{w}}{d\omega} \right] \cdot \vec{\sigma}, \quad (3.13)$$

and consequently $\vec{W}(\omega) = \lambda_\omega \hat{w} + \sin(\lambda) \hat{w}_\omega + (1 - \cos(\lambda)) \hat{w} \times \hat{w}_\omega$ [45, 84, 88]. We can now, as in Ref. [59], differentiate $\vec{W}(\omega)$ M times and then set $\omega = \omega_0$, at which $\lambda(\omega_0) = 0$. In this manner we find the derivatives of λ and \hat{w} in terms of different orders of the complex principal state vector. For $M = 0, 1$ and 2 ,

$$\begin{aligned} \vec{W}_0 &= \lambda_1 \hat{w}_0 = \frac{d(\lambda \hat{w})}{d\omega} \Big|_{\omega_0} \\ \vec{W}_1 &= \lambda_2 \hat{w}_0 + 2\lambda_1 \hat{w}_1 = \frac{d^2(\lambda \hat{w})}{d\omega^2} \Big|_{\omega_0} \\ \vec{W}_2 &= \lambda_3 \hat{w}_0 + 3\lambda_2 \hat{w}_1 + 3\lambda_1 \hat{w}_2 + \lambda_1^2 \hat{w}_0 \times \hat{w}_1 \\ &= \frac{d^3(\lambda \hat{w})}{d\omega^3} \Big|_{\omega_0} + \frac{1}{2} \vec{W}_0 \times \vec{W}_1, \end{aligned} \quad (3.14)$$

where as before $A_n \equiv (d^n A/d\omega^n) |_{\omega_0}$. Taylor expanding the exponent of the Jones matrix, Eq. (3.12), to $O(\Delta\omega^3)$ and inserting the identities of Eq. (3.14), we have

$$\begin{aligned} \vec{\beta} + i\vec{\alpha} &= \vec{W}_0 \Delta\omega + \vec{W}_1 \frac{\Delta\omega^2}{2!} \\ &+ (\vec{W}_2 - \frac{1}{2} \vec{W}_0 \times \vec{W}_1) \frac{\Delta\omega^3}{3!} + O(\Delta\omega^4) \end{aligned} \quad (3.15)$$

Since $\vec{W} = \vec{\Omega} + i\vec{\Lambda}$, we obtain the Jones space result $\mathbf{T}(\omega) = \exp \left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\sigma} \right] \mathbf{T}(\omega_0)$, in which \vec{a} and \vec{b} are obtained by truncating Eq. (3.9) and Eq. (3.10) to third order in $\Delta\omega$. The Mueller matrix version of Eq. (3.8) is then recovered by applying Eq. (3.1).

While the procedure above yields the same results as Section 3.2 to all orders of $\Delta\omega$, repeatedly differentiating $\vec{W}(\omega)$ to form the power series analog of Eq. (3.15) becomes increasingly difficult for fourth and higher-order terms. The Magnus expansion, on the other hand, directly expresses the Mueller matrix in terms of the relevant quantities $\vec{\Omega}_n$ and $\vec{\Lambda}_n$ to arbitrary order.

3.4 Power series expansion

In Ref. [48], a series expansion for the frequency-dependent Jones matrix for systems affected by PMD was introduced and evaluated to fourth-order in $\Delta\omega$. In this section we demonstrate the equivalence of this procedure to an analogous expansion of the matrix exponential presented in Eq. (3.11).

Specializing to the case of zero PDL, and setting $\vec{b} = b\hat{b}$ with $\vec{\Lambda} = 0$, the Jones matrix of Eq. (3.11) reduces to

$$\begin{aligned} \mathbf{U} &\equiv \mathbf{T}(\omega)\mathbf{T}(\omega_0)^{-1} \\ &= \exp\left[-\frac{i}{2}(\vec{b} \cdot \vec{\sigma})\right] \\ &= \cos\left(\frac{b}{2}\right)\sigma_0 - i\sin\left(\frac{b}{2}\right)(\hat{b} \cdot \vec{\sigma}). \end{aligned} \tag{3.16}$$

We now construct the vector, \vec{u} , and scalar, u_0 , components of the Stokes vector formed from the elements of the Jones matrix, \mathbf{U} , according to

$$\begin{aligned} u_0 &= \Re(\mathbf{U}_{00}) = \cos\left(\frac{b}{2}\right) \\ \vec{u} &= -\begin{bmatrix} \Im(\mathbf{U}_{00}) \\ \Im(\mathbf{U}_{01}) \\ \Re(\mathbf{U}_{01}) \end{bmatrix} = \sin\left(\frac{b}{2}\right)\hat{b}. \end{aligned} \tag{3.17}$$

The second equality in each expression can be derived from a comparison with Eq. (3.16). In terms of the unit vector $\hat{p} \equiv \vec{\Omega}_0/|\vec{\Omega}_0|$, we find after expanding the quantities u_0 and \vec{u} ,

in a power series, c.f. Eq. (3.10), and identifying terms of equal orders in $\Delta\omega$

$$\begin{aligned}
u_0 &= 1 - |\vec{\Omega}_0|^2 \frac{\Delta\omega^2}{8} - \vec{\Omega}_0 \cdot \vec{\Omega}_1 \frac{\Delta\omega^3}{8} \\
&\quad + \left(\frac{1}{16} |\vec{\Omega}_0|^4 - \frac{3}{4} |\vec{\Omega}_1|^2 - \vec{\Omega}_0 \cdot \vec{\Omega}_2 \right) \frac{\Delta\omega^4}{24} + \dots \\
\hat{p} \cdot \vec{u} &= |\vec{\Omega}_0| \frac{\Delta\omega}{2} + \hat{p} \cdot \vec{\Omega}_1 \frac{\Delta\omega^2}{4} \\
&\quad + \left(\hat{p} \cdot \vec{\Omega}_2 - \frac{1}{4} |\vec{\Omega}_0|^3 \right) \frac{\Delta\omega^3}{12} \\
&\quad + \left(\hat{p} \cdot \vec{\Omega}_3 - \frac{3}{2} |\vec{\Omega}_0| \vec{\Omega}_0 \cdot \vec{\Omega}_1 \right) \frac{\Delta\omega^4}{48} + \dots \\
\vec{u}_\perp &\equiv \vec{u} - (\hat{p} \cdot \vec{u}) \hat{p} \\
&= \vec{\Omega}_{1\perp} \frac{\Delta\omega^2}{4} + \left(\vec{\Omega}_{2\perp} - \frac{1}{2} \vec{\Omega}_0 \times \vec{\Omega}_{1\perp} \right) \frac{\Delta\omega^3}{12} \\
&\quad + \left(\vec{\Omega}_{3\perp} - \vec{\Omega}_0 \times \vec{\Omega}_2 - \frac{1}{2} |\vec{\Omega}_0|^2 \vec{\Omega}_{1\perp} \right) \frac{\Delta\omega^4}{48} + \dots
\end{aligned} \tag{3.18}$$

In the above equations, $\vec{\Omega}_{n\perp} = \vec{\Omega}_n - (\hat{p} \cdot \vec{\Omega}_n) \hat{p}$. If we set $\vec{\tau} = \vec{\Omega}/2$, Eqs. (3.18) reproduce the results of Ref. [48], indicating that the Magnus expansion result, Eq. (3.11), correctly characterizes the higher-order phase delays of the cross-coupled principal state of polarization (PSP) components [48]. However, the power series expansion of Eqs. (3.18) does not preserve the unitary group property of the Jones matrix, Eq. (3.16), reducing the accuracy of the procedure as compared to Eq. (3.11).

3.5 Results

We now demonstrate the accuracy of the Magnus expansion by computing the average, $\langle \epsilon \rangle$, and standard deviation, σ_ϵ , of the relative error

$$\epsilon = \frac{\|\mathbf{U}_{\text{est}} - \mathbf{U}_{\text{act}}\|}{\|\mathbf{U}_{\text{act}}\|} \times 100\%, \tag{3.19}$$

in which $\mathbf{U} = \mathbf{T}(\omega) \mathbf{T}(\omega_0)^{-1}$ and $\|\dots\|$ is identified with the Frobenius matrix norm, as a function of $\Delta\omega$. In the above formula, the estimated Jones matrix, \mathbf{U}_{est} , results from

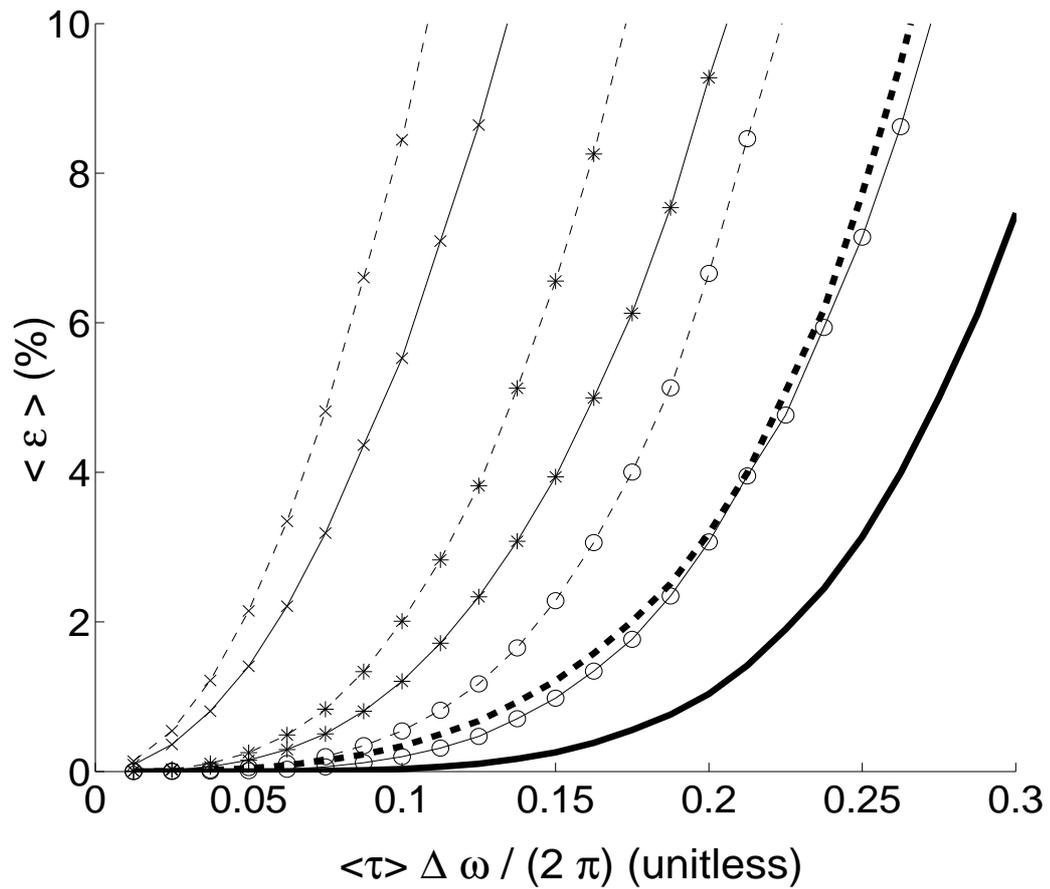


Figure 3.1: The average relative error, $\langle \epsilon \rangle$, for 5,000 fiber realizations, as a function of the normalized frequency deviation, $\langle \tau \rangle \Delta \omega / (2\pi)$, for a mean DGD $\langle \tau \rangle$. Dashed lines - power series expansion, Eq. (3.18), solid lines - Magnus expansion, Eq. (3.11). The first to fourth order results are indicated by \times , $*$, \circ and no markers, respectively.

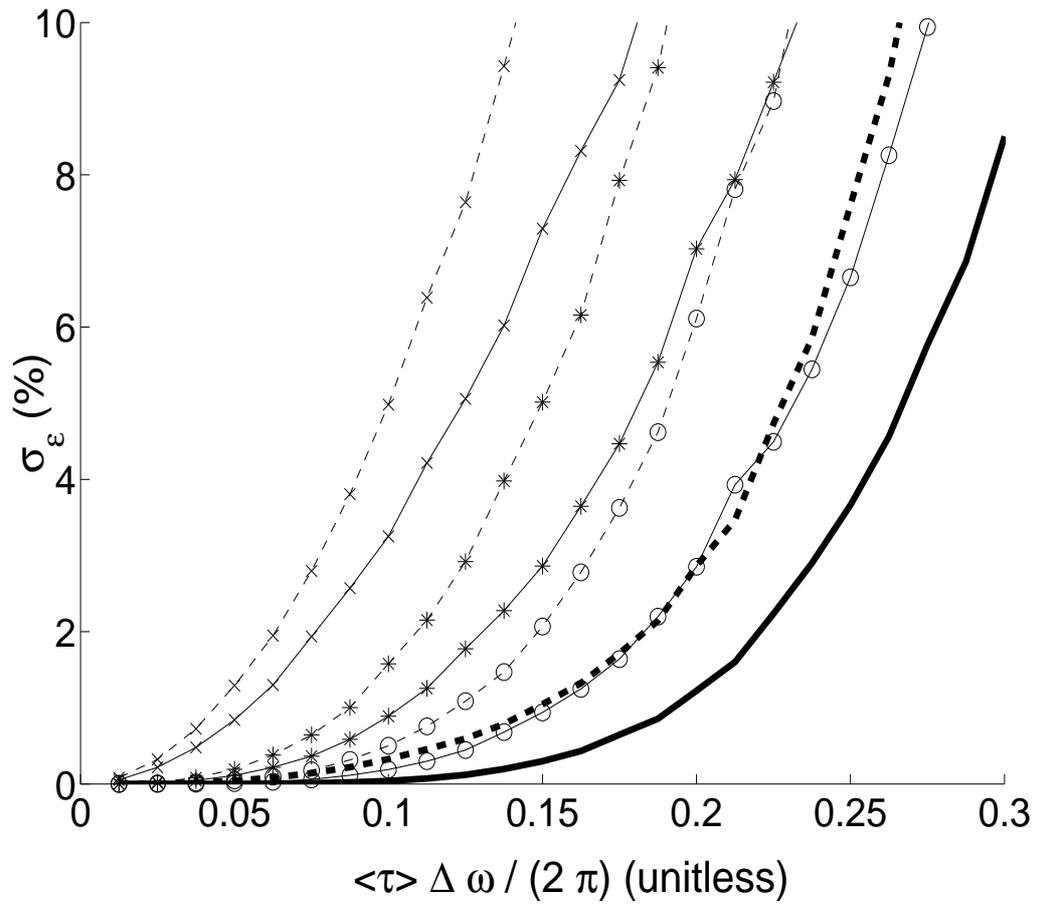


Figure 3.2: Same as Fig. 3.1, except for the standard deviation of the relative error, σ_ϵ .

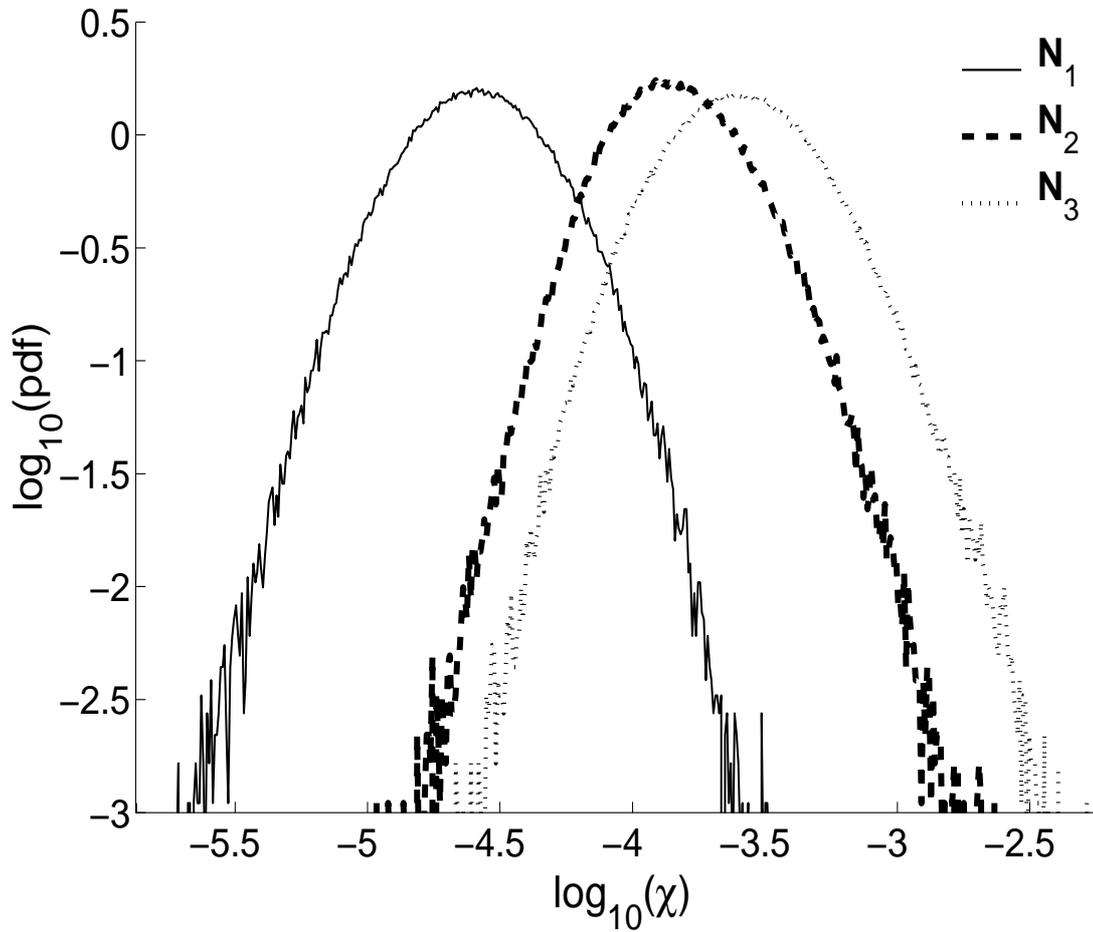


Figure 3.3: The simulated probability density function (pdf) of the relative error, χ , associated with the Taylor coefficients N_1 , N_2 , and N_3 for 2.5×10^5 fiber realizations with 30.0 ps and 10.0 dB mean PDL. The solid, dashed and dotted lines represent the relative error of N_1 , N_2 , and N_3 , respectively.

either the Magnus expansion of Eq. (3.11) with zero PDL, or, for comparison purposes, the power series of Eqs. (3.18). The “actual” Jones matrix, \mathbf{U}_{act} , instead corresponds to the Jones matrix evaluated numerically at $\Delta\omega$. We model an optical fiber link with a mean DGD $\langle\tau\rangle = 25$ ps by a set of 100 randomly oriented, linearly birefringent, polarization-maintaining (PM) fiber segments. The vectors $\vec{\Omega}_n$ are obtained by finite differencing the resulting Jones matrix.

Fig. 3.1 depicts the average relative error, $\langle\epsilon\rangle$, for 5,000 fiber realizations, as a function of the normalized frequency deviation, $\langle\tau\rangle\Delta\omega/(2\pi)$ (the curves are then independent of $\langle\tau\rangle$). Analogous results for the standard deviation of the relative error, σ_ϵ , are shown in Fig. 3.2. The solid and dashed lines in Fig. 3.1 and Fig. 3.2 were calculated with the Magnus Eq. (3.11) and power series expansions Eq. (3.18), while \times , \star , and \circ and no markers indicate the first to fourth order results in $\Delta\omega$, respectively. Clearly, the Magnus expansion, Eq. (3.11), yields more accurate results for the Jones matrix than the explicit power series of Eq. (3.18) when the optical frequency differs substantially from the center frequency.

We next verify our procedure in the presence of both PMD and PDL through experiment and numerical simulation for the first three orders $n = 1, 2, 3$ in $\Delta\omega$.

In our experiment, the Jones matrix $\mathbf{T}(\omega)$ was first determined over a 4 nm range in 0.1 nm steps. We employed an HP 81689A tunable laser source continuously calibrated with an Ando AQ6140 wavelength meter, and measured the Jones matrix from the output of an HP 8509B polarimeter. The device under test (DUT) consisted of three sections of PM fiber interleaved with two adjustable PDL elements, producing a net PDL in the range of 0 dB to 2 dB as recorded by the HP 8509B.

Next, a Jones to Mueller matrix conversion is performed through Eq. (3.1), and the matrix exponent at each frequency is extracted according to $\mathbf{N}(\omega_m - \omega_0) = \ln(\mathbf{M}(\omega_m)\mathbf{M}(\omega_0)^{-1})$, in which $m = 1, 2, \dots, M$ and M denotes the number of measured frequencies. We then derive the Taylor series coefficients \mathbf{N}_n from a polynomial least square fit of $\mathbf{N}(\Delta\omega)$. For the purpose of comparison, we estimate \mathbf{N}_n from the analytic expression, Eq. (3.7), in

Table 3.1: Measured relative error χ (%).

<i>DGD</i>	2.2 ps	2.3 ps	3.9 ps	9.0 ps	6.8 ps	10.0 ps
<i>PDL</i>	0.3 dB	1.0 dB	2.0 dB	0.3 dB	0.9 dB	2.3 dB
\mathbf{N}_1	0.21	0.33	0.54	0.41	0.48	0.69
\mathbf{N}_2	4.3	4.5	6.3	4.8	6.0	7.3
\mathbf{N}_3	6.8	9.7	17.4	19.0	14.4	—

which the quantities $\vec{\Omega}_n$ and $\vec{\Lambda}_n$ are obtained by finite-differencing the measured Jones matrices, $\mathbf{T}(\omega_m)$, according to $\vec{W} \cdot \vec{\sigma} = 2i\mathbf{T}_\omega\mathbf{T}^{-1}$. Denoting the estimates of \mathbf{N}_n from the two procedures by $\mathbf{N}_n^{(1)}$ and $\mathbf{N}_n^{(2)}$, respectively, we calculate the relative error χ as

$$\chi = \frac{\|\mathbf{N}_n^{(2)} - \mathbf{N}_n^{(1)}\|}{\|\mathbf{N}_n^{(1)}\|} \times 100\%. \quad (3.20)$$

Table 3.1 displays the error, χ , for several system configurations, arranged according to the measured DGD and PDL at the center frequency. While wavelength drift and other systematic errors in many cases precluded the accurate determination of \mathbf{N}_3 , for a 0.1 nm wavelength step size, the error in \mathbf{N}_1 and \mathbf{N}_2 does not exceed 8% for $\text{PDL} \leq 2$ dB. Accordingly, although we examined only a small number of single system realizations, the data clearly indicates that the measured frequency dependence of the Mueller matrix coincides with that predicted from Eq. (3.6) to second order and is consistent with theoretical predictions to third order.

Finally, to verify if the above measurements are characteristic of the optical fiber behavior, we calculate the probability density functions associated with the error, χ , for each coefficient \mathbf{N}_n , $n = 1, 2, 3$ [85]. Here, our calculation employs 100 randomly oriented, linearly birefringent and lossy elements to simulate a fiber with a mean DGD of 30.0 ps and an artificially large mean PDL of 10.0 dB, which is chosen to generate physically interesting “worst case” values for $\vec{\Lambda}_n$. After 2.5×10^5 fiber realizations, c.f. Fig. 3.3, the observed mean values of χ are found to be $3.1 \times 10^{-5}\%$, $1.7 \times 10^{-4}\%$, and $3.2 \times 10^{-4}\%$ for \mathbf{N}_n , $n = 1, 2, 3$, respectively. The worst case, $\chi \approx 10^{-2.5}\% \approx 3.2 \times 10^{-3}\%$, occurs for \mathbf{N}_3 ,

indicating that for these conditions the accuracy of the Magnus expansion approximation, Eq. (3.6), to third order in $\Delta\omega$, exceeds 4 digits.

3.6 Conclusions

We have employed the Magnus expansion to model the frequency dependence of the Mueller matrix in birefringent systems affected by both PMD and PDL and further verified that our formalism reproduces previous power-series expansion techniques. However, although any Mueller or Jones matrix model obtained by solving the underlying differential equation in frequency will agree with our formalism, the Magnus expansion explicitly preserves the underlying symmetries of the Mueller matrix [86, 87] and further yields physically realizable operator expansions that facilitate the design of joint PMD and PDL compensators [88, 115, 118]. In addition, other models [31, 118], can be directly obtained from the Magnus expansion through application of the Baker-Campbell-Hausdorff identity [42, 88]. Unfortunately, while the numerical computations can be performed to any desired order of accuracy, the experimental determination of third and higher order PMD vectors requires a level of precision that appears to be beyond our present capabilities. However, improved data analysis techniques, may enable such measurements.

Chapter 4

A Clifford Algebra Analysis of PMD and PDL

We overview Clifford algebra and its application to polarization mode dispersion (PMD) and polarization dependent loss (PDL). We demonstrate that the Jones matrix corresponds to a Lorentz transformation in this algebra and obtain its frequency dependent evolution through the Magnus expansion. We also comment on the application of Clifford algebra methods to partially polarized light.

4.1 Introduction

Despite the frequency with which Jones and Stokes matrices are employed in the analysis of optical polarization effects, several recent papers have considered alternate algebraic approaches to describing polarization mode dispersion (PMD) and polarization dependent loss (PDL) in single mode fibers. In particular, quaternions obey a multiplication rule that follow the algebra of the Pauli spin matrices and therefore the Jones matrices (or their $SO^+(1, 3)$ Mueller matrix covering group) in a particularly transparent fashion.

Quaternions, however, constitute only a single, four-dimensional, “subalgebra” of a Clifford algebra that extends this construct to vectors of arbitrary dimension. Mathematical objects that are isomorphic to complex numbers, linear algebra, and quaternions can be similarly modeled in the Clifford algebra formalism, and can then be rapidly manipulated with a transparent geometric interpretation.

Below we first overview the Clifford algebra formalism [9, 27, 49]. We demonstrate the equivalence between the Clifford product in three-dimensional space and physical rotations. We also show that optical elements that are affected by PMD and PDL but which preserve the coherence of the optical field are described by Lorentz transformations in the Clifford algebra formulation. We obtain a particularly simple form of the Magnus expansion solution [71, 75, 88, 118] for the frequency dependence of the polarization evolution in the Clifford algebra formulation and finally discuss the manner in which Clifford algebra techniques can be employed to reduce the abstraction and the computational requirements of problems involving partial polarization.

4.2 Summary of Clifford algebra

In an N -dimensional space, \mathcal{R}^N , spanned by orthogonal unit vectors $\{\hat{e}_1, \hat{e}_2, \dots, \hat{e}_N\}$, Clifford algebra is based upon an associative, but non-commutative, vector multiplication satisfying two fundamental identities [9, 11, 49]

$$\begin{aligned}\hat{e}_i \hat{e}_i &= |\hat{e}_i|^2 = 1, \\ \hat{e}_i \hat{e}_j &= -\hat{e}_j \hat{e}_i.\end{aligned}\tag{4.1}$$

Specializing to three-dimensions for simplicity, Eq. (4.1) implies $\vec{a}^2 = |\vec{a}|^2$ for the vector $\vec{a} = a_1 \hat{e}_1 + a_2 \hat{e}_2 + a_3 \hat{e}_3$. The products $\hat{e}_i \hat{e}_j$, form three distinct “bivectors”. These differ from vectors since $(\hat{e}_i \hat{e}_j)^2 = -\hat{e}_i (\hat{e}_j \hat{e}_j) \hat{e}_i = -1$. The single “trivector” $\hat{e}_i \hat{e}_j \hat{e}_k$ is unchanged in sign under a cyclic permutation of its indices, while $(\hat{e}_i \hat{e}_j \hat{e}_k)^2 = -1$.

Scalars, vectors, bivectors and trivectors comprise a basis for a new vector space \mathcal{Cl}_3 , such that we can create "multivector" objects as the linear combination

$$\begin{aligned}\check{p} &= p_0 + p_1\hat{e}_1 + p_2\hat{e}_2 + p_3\hat{e}_3 \\ &+ p_4\hat{e}_1\hat{e}_2 + p_5\hat{e}_1\hat{e}_3 + p_6\hat{e}_2\hat{e}_3 + p_7\hat{e}_1\hat{e}_2\hat{e}_3,\end{aligned}\tag{4.2}$$

with real coefficients p_n . The vector space \mathcal{R}^3 is therefore a subspace of \mathcal{Cl}_3 . Since the trivector $\hat{e}_1\hat{e}_2\hat{e}_3$ both commutes with all other basis elements and satisfies $(\hat{e}_1\hat{e}_2\hat{e}_3)^2 = -1$, it is often denoted by the suggestive label $I \equiv \hat{e}_1\hat{e}_2\hat{e}_3$ and is termed a pseudoscalar. Indeed, the multivectors $\check{p} = p_0 + p_7I$ are isomorphic to standard complex numbers.

Multiplication of two multivectors \check{a} and \check{b} is expressed in terms of their "Clifford product"

$$\check{a}\check{b} = \check{a} \cdot \check{b} + \check{a} \wedge \check{b},\tag{4.3}$$

which is decomposed into the symmetric "inner" product, $\check{a} \cdot \check{b} = (\check{a}\check{b} + \check{b}\check{a})/2$, and an anti-symmetric "outer" product, $\check{a} \wedge \check{b} = (\check{a}\check{b} - \check{b}\check{a})/2$. As an example, if \vec{a} and \vec{b} are the vectors $\vec{a} = a_1\hat{e}_1 + a_2\hat{e}_2 + a_3\hat{e}_3$ and $\vec{b} = b_1\hat{e}_1 + b_2\hat{e}_2 + b_3\hat{e}_3$, their product is found through repeated application of Eq. (4.1) to be

$$\begin{aligned}\vec{a}\vec{b} &= (a_1b_1 + a_2b_2 + a_3b_3) \\ &+ (a_1b_2 - a_2b_1)\hat{e}_1\hat{e}_2 \\ &+ (a_1b_3 - a_3b_1)\hat{e}_1\hat{e}_3 \\ &+ (a_2b_3 - a_3b_2)\hat{e}_2\hat{e}_3,\end{aligned}\tag{4.4}$$

with an analogous expression for $\vec{b}\vec{a}$. Accordingly, we observe that the inner product, $(\vec{a}\vec{b} + \vec{b}\vec{a})/2$, coincides with the standard vector dot product. The antisymmetric outer product instead corresponds to the expression

$$\begin{aligned}\vec{a} \wedge \vec{b} &= \frac{1}{2}(\vec{a}\vec{b} - \vec{b}\vec{a}) \\ &= (a_1b_2 - a_2b_1)\hat{e}_1\hat{e}_2 \\ &+ (a_1b_3 - a_3b_1)\hat{e}_1\hat{e}_3 \\ &+ (a_2b_3 - a_3b_2)\hat{e}_2\hat{e}_3\end{aligned}\tag{4.5}$$

and is therefore a bivector. Since, however, $\hat{e}_1\hat{e}_2 = (\hat{e}_1\hat{e}_2\hat{e}_3)\hat{e}_3 = I\hat{e}_3$, while $\hat{e}_1\hat{e}_3 = -I\hat{e}_2$ and $\hat{e}_2\hat{e}_3 = I\hat{e}_1$, the above outer product is related in three-dimensions to the standard cross-product through

$$\vec{a} \wedge \vec{b} = I\vec{a} \times \vec{b}, \quad (4.6)$$

such that

$$\vec{a}\vec{b} = \vec{a} \cdot \vec{b} + I\vec{a} \times \vec{b}. \quad (4.7)$$

Hence the magnitude of the bivector $\vec{a} \wedge \vec{b}$ equals the area of the parallelogram formed by \vec{a} and \vec{b} . This bivector is often associated with the \vec{a}, \vec{b} plane. Comparing Eq. (4.7) with quaternion multiplication [55] we additionally observe that linear combinations of scalars and bivectors of the form $\check{a} = a_0 - \vec{a}I$ generate an algebra isomorphic to that of quaternions.

If the bivectors $\hat{e}_i\hat{e}_j$ are recast as products of vectors with the pseudoscalar, I , multivectors in $\mathcal{C}\ell_3$ can be written in the form $\check{p} = p + \vec{p}$, where p and \vec{p} are "pseudocomplex" quantities; that is, expressions of the form $\zeta + \xi I$ where ζ and ξ are real scalars and vectors, respectively. The spatial inversion and the Hermitian conjugate of \check{p} are then defined as

$$\check{\check{p}} = p - \vec{p} \quad (4.8)$$

and

$$\check{p}^\dagger = p^* + \vec{p}^*, \quad (4.9)$$

with $(\zeta + \xi I)^* = \zeta - \xi I$. The inverse of a multivector then takes the form

$$\check{p}^{-1} = \check{\check{p}}/(\check{p}\check{\check{p}}) \quad (4.10)$$

enabling division by a multivector quantity. The Clifford algebra of three dimensional space, $\mathcal{C}\ell_3$, naturally represents four dimensional spacetime since the Minkowski spacetime metric is represented by the Clifford product $\check{p}\check{\check{p}}$, which yields the pseudocomplex scalar $p^2 - \vec{p}^2$.

We now establish the relationship between the exponential operator in Clifford algebra and physical rotations. Rotations of the vector \vec{v} by an angle θ about the axis \hat{n} are effected

in Clifford algebra by the symmetric product [27]

$$\vec{v}' = \check{R}\vec{v}\check{R}^\dagger, \quad (4.11)$$

where

$$\check{R} = \exp\left(-I\frac{\theta}{2}\hat{n}\right) = \cos\left(\frac{\theta}{2}\right) - I\sin\left(\frac{\theta}{2}\right)\hat{n} \quad (4.12)$$

is itself a multivector, satisfying $\check{R}\check{\bar{R}} = 1$. The second expression for \check{R} in Eq. (4.12) is formed by expanding the exponential in power series and applying $\hat{n}^2 = 1$. Finally, we note that the symmetric rotation of Eq. (4.11) remains valid for all multivectors, that is, when $\vec{v} \rightarrow \check{p}$.

Substituting $\vec{b} + I\vec{a}$ for $\theta\hat{n}$, for real three-vectors \vec{b} and \vec{a} , leads to a representation of \check{R} in terms of the Lorentz transformation [9]

$$\begin{aligned} \check{R} &= \exp\left[-\frac{I}{2}(\vec{b} + I\vec{a})\right] \\ &= \cos\left(\frac{\lambda}{2}\right) - I\sin\left(\frac{\lambda}{2}\right)\check{w}, \end{aligned} \quad (4.13)$$

where $\lambda = \sqrt{\vec{b}^2 - \vec{a}^2 + 2I\vec{a} \cdot \vec{b}}$ and $\check{w} = (\vec{b} + I\vec{a})/\lambda$. Such a form preserves the property $\check{R}\check{\bar{R}} = 1$, as well as the norm of the transformed vector $\check{p}' = \check{R}\check{p}\check{R}^\dagger$, i.e. $\check{p}\check{\bar{p}} = \check{p}'\check{\bar{p}'}$, and therefore the Minkowski metric.

4.3 PMD and PDL

In a single-mode optical waveguide supporting two (possibly quasi-)orthogonal modes, the Jones vectors of the output and the frequency-independent input electric field polarizations, $|t(\omega)\rangle$ and $|s\rangle$, are related through a complex 2×2 Jones matrix with unit determinant such that $|t(\omega)\rangle = \mathbf{T}(\omega)|s\rangle$. The coherency matrix, formed by the ensemble average of $|s\rangle\langle s|$, describes the statistics of a stochastically varying electric field. It consequently transforms according to

$$|t\rangle\langle t| = \mathbf{T}|s\rangle\langle s|\mathbf{T}^\dagger, \quad (4.14)$$

which is identical in form to Eq. (4.11). Observing that a Jones matrix \mathbf{T} with $\det(\mathbf{T}) = 1$ can be expressed in terms of two vectors \vec{a} and \vec{b} according to $\mathbf{T} = \exp\left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\sigma}\right]$, where $\vec{\sigma}$ represents the Pauli spin vector, we identify Eq. (4.14) as the Jones matrix representation of a Clifford algebra Lorentz transformation.

In the case of a frequency dependent Lorentz transformation $\check{T}(\omega)$ resulting from PMD and PDL in an optical system, the Clifford algebra analog of the Jones space operator $\mathbf{T}_\omega \mathbf{T}^{-1}$ is $\check{T}_\omega \check{T}$. Upon substituting \check{T} into $\check{T}_\omega \check{T}$, we find

$$\frac{d\check{T}}{d\omega} \check{T} = -\frac{I}{2} \left[\frac{d\lambda}{d\omega} \check{w} + \sin(\lambda) \frac{d\check{w}}{d\omega} + I(1 - \cos(\lambda)) \frac{d\check{w}}{d\omega} \check{w} \right], \quad (4.15)$$

where λ and \check{w} are given by Eq. (4.13) [45]. The quantity within square brackets is itself a multivector, which can be further decomposed into real and imaginary components according to

$$\begin{aligned} \check{W} &\equiv \vec{\Omega} + I\vec{\Lambda} \\ &= \frac{d\lambda}{d\omega} \check{w} + \sin(\lambda) \frac{d\check{w}}{d\omega} + I(1 - \cos(\lambda)) \frac{d\check{w}}{d\omega} \check{w}. \end{aligned} \quad (4.16)$$

The quantity \check{W} is then the Clifford algebra representation of the complex principal state vector [50].

To obtain the transformation \check{T} in terms of the real three-vectors $\vec{\Omega}$ and $\vec{\Lambda}$, we observe that the frequency derivative, denoted below through the subscript ω , of \check{T} always satisfies

$$\frac{d\check{T}}{d\omega} = \left(\frac{d\check{T}}{d\omega} \check{T} \right) \check{T} = \check{H}(\omega) \check{T}, \quad (4.17)$$

where $\check{T} \check{T} = 1$ identically, and $\check{H}(\omega) = -\frac{I}{2}(\vec{\Omega} + I\vec{\Lambda})$. $\check{T}(\omega)$ is then obtained as the product

$$\check{T}(\omega) = F_\omega \exp \left[\int_{\omega_0}^{\omega} \check{H}(\omega_1) d\omega_1 \right] \check{T}(\omega_0), \quad (4.18)$$

in which F_ω denotes the frequency ordering operator [42]. A more convenient representation of the general solution is provided by the Magnus expansion [71, 75], namely

$$\check{T}(\omega) = \exp \left[\sum_{n=1}^{\infty} \check{B}_n(\omega) \right] \check{T}(\omega_0). \quad (4.19)$$

The first two Magnus expansion coefficients are

$$\begin{aligned}\check{B}_1(\omega) &= \int_{\omega_0}^{\omega} d\omega_1 \check{H}(\omega_1) \\ \check{B}_2(\omega) &= \frac{1}{2} \int_{\omega_0}^{\omega} \int_{\omega_0}^{\omega_1} d\omega_2 d\omega_1 [\check{H}(\omega_1), \check{H}(\omega_2)],\end{aligned}\quad (4.20)$$

in which square brackets denote commutation, and \check{B}_n for $n > 2$ are related to those of lower order through recursion relations [75]. Taylor expanding the vectors $\vec{\Omega}$ and $\vec{\Lambda}$ in Eq. (4.20) to third order in $\Delta\omega$, and directly evaluating the Magnus coefficients, we find

$$\begin{aligned}\check{B}_1 &= \frac{1}{2}(\vec{\Lambda}_0\Delta\omega + \vec{\Lambda}_1\frac{\Delta\omega^2}{2!} + \vec{\Lambda}_2\frac{\Delta\omega^3}{3!}) \\ &\quad - \frac{I}{2}(\vec{\Omega}_0\Delta\omega + \vec{\Omega}_1\frac{\Delta\omega^2}{2!} + \vec{\Omega}_2\frac{\Delta\omega^3}{3!}) + O(\Delta\omega^4) \\ \check{B}_2 &= \frac{1}{4}\frac{\Delta\omega^3}{3!}(\vec{\Lambda}_1 \wedge \vec{\Lambda}_0 + \vec{\Omega}_0 \wedge \vec{\Omega}_1 \\ &\quad + I\vec{\Lambda}_0 \wedge \vec{\Omega}_1 + I\vec{\Omega}_0 \wedge \vec{\Lambda}_1) + O(\Delta\omega^4).\end{aligned}\quad (4.21)$$

Here, we have applied the commutator relation $[\vec{a}, \vec{b}] = \vec{a}\vec{b} - \vec{b}\vec{a} = 2\vec{a} \wedge \vec{b}$, and employed subscripts to denote frequency derivatives, i.e. $A_n = (d^n A/d\omega^n)|_{\omega_0}$ [88]. The Lorentz transformation \check{T} is generated to third order in frequency once the above expansion coefficients are inserted into Eq. (4.19), and the wedge product is replaced by the cross-product according to $\vec{a} \wedge \vec{b} = I\vec{a} \times \vec{b}$, yielding

$$\begin{aligned}\check{T} &= \exp\left[-\frac{I}{2}(\vec{b} + I\vec{a})\right] \\ \vec{a} &= \vec{\Lambda}_0\Delta\omega + \vec{\Lambda}_1\frac{\Delta\omega^2}{2!} \\ &\quad + \left(\vec{\Lambda}_2 - \frac{1}{2}\vec{\Lambda}_0 \times \vec{\Omega}_1 - \frac{1}{2}\vec{\Omega}_0 \times \vec{\Lambda}_1\right)\frac{\Delta\omega^3}{3!} + O(\Delta\omega^4) \\ \vec{b} &= \vec{\Omega}_0\Delta\omega + \vec{\Omega}_1\frac{\Delta\omega^2}{2!} \\ &\quad + \left(\vec{\Omega}_2 - \frac{1}{2}\vec{\Omega}_0 \times \vec{\Omega}_1 - \frac{1}{2}\vec{\Lambda}_1 \times \vec{\Lambda}_0\right)\frac{\Delta\omega^3}{3!} + O(\Delta\omega^4).\end{aligned}\quad (4.22)$$

When two fibers described by the individual transformations $\check{T}_{(1)}$ and $\check{T}_{(2)}$ are concatenated, the composite transformation is given by the product $\check{T} = \check{T}_{(2)}\check{T}_{(1)}$; while the corresponding multivectors \check{H} satisfy

$$\frac{d\check{T}}{d\omega}\check{T} = \check{H}_{(2)} + \check{T}_{(2)}\check{H}_{(1)}\check{T}_{(2)}. \quad (4.23)$$

The last term in the above equation does not transform in the same manner as Eq. (4.14) under \check{T} , but rather as the product $\check{t}\check{s}$, i.e. $\check{t}'\check{s}' = (\check{T}\check{t}\check{T}^\dagger)(\check{T}^\dagger\check{s}\check{T}) = \check{T}(\check{t}\check{s})\check{T}$, consistent with the properties of Lorentz transformations. From the invariance of scalars under Lorentz transformations, the real and imaginary components of the scalar quantity $\check{H}_{(1)}^2$, namely $\vec{\Omega}_{(1)}^2 - \vec{\Lambda}_{(1)}^2$ and $\vec{\Omega}_{(1)} \cdot \vec{\Lambda}_{(1)}$, are similarly unchanged in Eq. (4.23).

Another application of Clifford algebra in optics that is directly relevant to PMD is the description of the polarization coherency matrix. While a full analysis of this topic exceeds the scope of this chapter, the essential aspects of the theory are summarized below. The formalism can then be immediately adapted to the propagation of partially polarized light such as that occurring in PMD measurements performed by illuminating a single mode fiber with a LED or white light source.

We first represent the coherency matrix

$$|s\rangle\langle s| = \frac{1}{2}(s_0\sigma_0 + \vec{s} \cdot \vec{\sigma}) \quad (4.24)$$

as a linear combination of the 2×2 identity matrix, σ_0 , and the Pauli spin matrices, $\vec{\sigma}$ with coefficients given by the real valued Stokes parameters, s_0 and \vec{s} . The Stokes parameters comprise a 4 component Lorentz vector, which can be assembled into a real-valued multivector $\check{s} = s_0 + \vec{s}$. Propagation through a birefringent and lossy media is described by the transformation

$$\check{t} = \check{T}\check{s}\check{T}^\dagger \quad (4.25)$$

in which $\check{t} = t_0 + \vec{t}$. The Clifford representation of the Jones matrix for a system with both PMD and PDL is, in analogy with Eq. (4.13), $\check{T} = \exp\left[-\frac{I}{2}(\vec{b} + I\vec{a})\right]$. Partial polarization is described by replacing \check{s} in the above formulas by the expression $s_0 + (1 - f)\vec{s}$, in which

f represents the fraction of depolarized light [11]. A linear polarizer in the \hat{n} direction and a square-law power detector then correspond to the operators $(1 + \hat{n})/2$, and $(\check{s} + \bar{s})/2$ respectively [11]. Numerous additional transformations of partially polarized states can be similarly represented and manipulated with Clifford algebraic methods, providing a further compelling benefit of the formalism.

4.4 Conclusions

We have demonstrated that Clifford algebra provides a framework for the description of polarization evolution in single-mode optical fibers in the presence of PMD and PDL. In particular, the Lorentz transformation that relates the Stokes vectors at the input and output of the system can be compactly expressed as a Clifford algebra product. Since these transformations are easily manipulated, numerous physical results can be derived in a straightforward fashion. Further, the method can be simply extended to partially polarized and partially coherent light.

Chapter 5

A Least-Squares Analysis of the Mueller Matrix

In a single-mode fiber excited by light with a fixed polarization state, the output polarizations obtained at two different optical frequencies are related by a Mueller matrix. We now examine least-squares procedures for estimating this matrix from repeated measurements of the output Stokes vector for a random set of input polarization states. We then apply these methods to the determination of polarization mode dispersion and polarization dependent loss in an optical fiber. We find that a relatively simple formalism leads to results that are comparable to those of far more involved techniques.

5.1 Introduction

The Mueller matrix, $\mathbf{M}(\omega)$, of a linear single-mode optical fiber represents the linear transformation between the polarizations of the incoming and outgoing light expressed in Stokes space. Associated with $\mathbf{M}(\omega)$ is the difference transformation, $\mathbf{M}_1(\omega_1 - \omega_0) = \mathbf{M}(\omega_1)\mathbf{M}(\omega_0)^{-1}$, that relates the output polarizations at the optical frequencies ω_0 and ω_1

for a fixed input polarization. If the fiber is excited by a random set of input polarizations [77] and the output Stokes vectors are measured at ω_1 and ω_0 for each polarization in this set, $\mathbf{M}_1(\omega_1 - \omega_0)$, and subsequently the polarization-mode dispersion (PMD) and polarization-dependent loss (PDL) coefficients can be estimated. In this chapter, we examine three computational methods for determining \mathbf{M}_1 from these measurements; namely, (1) an adaptation of Ref. [104] in which the Lorentz group properties satisfied by \mathbf{M}_1 are incorporated into a least-squares formulation through Lagrange multipliers, (2) a novel simplified procedure in which the solution to the least-squares problem is factored into the product of two matrices, one of which possesses the correct Lorentz group symmetries, and (3) the numerical optimization procedure of Refs. [58] and [57]. A numerical study reveals that the second approach yields results that are nearly as accurate as those obtained with numerical optimization.

If an optical system is excited by a field described by a frequency independent input Jones vector $|s\rangle$, the associated outgoing field vector, $|t(\omega)\rangle$, is given by $|t(\omega)\rangle = \mathbf{T}(\omega)|s\rangle$ where the matrix \mathbf{T} is normalized such that $\det(\mathbf{T}) = 1$, since we consider only the differential attenuation and phase [50]. The corresponding relationship between the input and output Stokes vectors $\tilde{s} = \langle s|\tilde{\sigma}|s\rangle$ and $\tilde{t} = \langle t|\tilde{\sigma}|t\rangle$ is $\tilde{t}(\omega) = \mathbf{M}(\omega)\tilde{s}$, in which $\tilde{\sigma} = (\boldsymbol{\sigma}_0, \vec{\sigma})$, and $\boldsymbol{\sigma}_0, \vec{\sigma}$ represent the 2×2 identity matrix and the Pauli spin vector [45]. Denoting complex and Hermitian conjugation by $*$ and \dagger , respectively, the Mueller matrix can be written as the Kronecker product $\mathbf{M} = \mathbf{A}(\mathbf{T} \otimes \mathbf{T}^*)\mathbf{A}^\dagger$ for the unitary matrix

$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix}. \quad (5.1)$$

From this representation of the Mueller matrix, we can establish through direct calculation that \mathbf{M} is orthogonal with respect to the Minkowski metric tensor [65]; that is, $\mathbf{M}^+\mathbf{M} = \mathbf{I}$, with $\mathbf{M}^+ = \mathbf{g}\mathbf{M}^T\mathbf{g}$, \mathbf{I} the 4×4 identity matrix and $\mathbf{g} = \text{diag}(-1, 1, 1, 1)$, c.f. Appendix B.3.

The Jones matrix, $\mathbf{T}(\omega)$, satisfies the differential identity $(d\mathbf{T}/d\omega)\mathbf{T}^{-1} = -i\vec{W} \cdot \vec{\sigma}/2$

where $\vec{W} = \vec{\Omega} + i\vec{\Lambda}$, with $\vec{\Omega}$ and $\vec{\Lambda}$ real vectors, is termed the complex principal state. [50] Evaluating $(d\mathbf{M}/d\omega)\mathbf{M}^{-1}$ from the Kronecker product form of the Mueller matrix yields the corresponding differential equation [88]

$$\frac{d\mathbf{M}}{d\omega}\mathbf{M}^{-1} = \mathbf{H}(\omega) = \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} \quad (5.2)$$

for the frequency evolution of \mathbf{M} . A general solution of Eq. (5.2) is provided by the Magnus expansion, Eq. (3.4) [42, 71, 75], which yields to third order in $\Delta\omega$

$$\mathbf{M}(\omega) = e^{\mathbf{N}(\Delta\omega)}\mathbf{M}(\omega_0) \quad (5.3)$$

where $\mathbf{N}(\Delta\omega) = \mathbf{H}_0\Delta\omega + \mathbf{H}_1\frac{\Delta\omega^2}{2!} + (\mathbf{H}_2 - \frac{1}{2}[\mathbf{H}_0, \mathbf{H}_1])\frac{\Delta\omega^3}{3!}$, and the subscript n denotes $(d^n/d\omega^n)|_{\omega_0}$. Alternatively, Eq. (5.2) describes the Mueller matrix directly in terms of the vectors $\vec{\Omega}_n$ and $\vec{\Lambda}_n$, c.f. Eq. (3.8).

5.2 Least-squares formulation

If the output Stokes vector of a fiber is measured for several input polarizations at the frequencies ω_0 and ω_1 , the most general relationship between the two sets of values takes the form [28, 73, 94]

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{M}_1\mathbf{Y}_0 \\ \mathbf{Y}_1 &= \begin{bmatrix} \tilde{t}_{(1)}(\omega_1), & \tilde{t}_{(2)}(\omega_1), & \dots, & \tilde{t}_{(N)}(\omega_1) \end{bmatrix}, \\ \mathbf{Y}_0 &= \begin{bmatrix} \tilde{t}_{(1)}(\omega_0), & \tilde{t}_{(2)}(\omega_0), & \dots, & \tilde{t}_{(N)}(\omega_0) \end{bmatrix}. \end{aligned} \quad (5.4)$$

In the above formula, $\tilde{t}_{(n)}(\omega)$ represents the output Stokes vector for the n^{th} input polarization state so that \mathbf{Y}_1 and \mathbf{Y}_0 are $4 \times N$ matrices. The matrix \mathbf{M}_1 is then given by $\mathbf{M}_1 = \mathbf{M}(\omega_1)\mathbf{M}(\omega_0)^{-1} = \exp[\mathbf{N}(\Delta\omega)]$; c.f. Eq. (5.3). Since the exact algebraic solution for \mathbf{M}_1 that minimizes the least-squares error $(\|\mathbf{Y}_1 - \mathbf{M}_1\mathbf{Y}_0\|)^2$ subject to the constraint

$\mathbf{M}_1^+ \mathbf{M}_1 = \mathbf{I}$, to our knowledge, does not yet exist, we instead investigate two approximate solution methods and the numerical optimization procedure of Refs. [58] and [57] (to simplify the mathematics, $\|\dots\|$ is identified with the Frobenius matrix norm).

Our first procedure, advanced in Ref. [104], employs the least-squares estimate

$$\mathbf{M}_1 = \mathbf{U} \mathbf{g} \mathbf{V} \mathbf{g}. \quad (5.5)$$

As demonstrated in Ref. [89], the matrices \mathbf{U} and \mathbf{V} can always be expressed in terms of $\mathbf{Y}_1 \mathbf{Y}_0^T$ through a Minkowski space singular value decomposition (SVD) that factors $\mathbf{Y}_1 \mathbf{Y}_0^T = \mathbf{U} \mathbf{D} \mathbf{V}$ such that $\mathbf{U}^{-1} = \mathbf{U}^+$, $\mathbf{V}^{-1} = \mathbf{V}^+$, and \mathbf{D} is diagonal. Algorithms for computing \mathbf{U} , \mathbf{D} and \mathbf{V} are presented in Ref. [104]. However, Eq. (5.5) does not minimize the true error $(\|\mathbf{Y}_1 - \mathbf{M}_1 \mathbf{Y}_0\|)^2$, but rather a cost function that preserves certain symmetries associated with the Lorentz group [104]. As shown below numerically, this leads to a certain degree of inaccuracy in Eq. (5.5).

The second procedure follows from the unconstrained least-squares solution, $\mathbf{M}_{LS} = \mathbf{Y}_1 \mathbf{Y}_0^T (\mathbf{Y}_0 \mathbf{Y}_0^T)^{-1}$, of Eqs (5.4) for which, in our case, $\mathbf{M}_{LS}^+ \mathbf{M}_{LS} \neq \mathbf{I}$. To ensure that the Mueller matrix possesses the required symmetries, we decompose \mathbf{M}_{LS} into the product of a matrix \mathbf{S}_1 that is symmetric with respect to the Minkowski metric tensor, i.e., $\mathbf{S}_1^+ = \mathbf{S}_1$, and a second, orthogonal, component $\mathbf{M}_1^{-1} = \mathbf{M}_1^+$, such that

$$\mathbf{M}_{LS} = \mathbf{M}_1 \mathbf{S}_1 \quad (5.6)$$

with

$$\begin{aligned} \mathbf{S}_1 &= (\mathbf{M}_{LS}^+ \mathbf{M}_{LS})^{\frac{1}{2}} \\ \mathbf{M}_1 &= \mathbf{M}_{LS} (\mathbf{M}_{LS}^+ \mathbf{M}_{LS})^{-\frac{1}{2}}. \end{aligned} \quad (5.7)$$

The order of the factorization in Eq. (5.6) affects the value of \mathbf{S}_1 but leaves \mathbf{M}_1 invariant; that is, $\mathbf{M}_{LS} = \mathbf{M}_1 \mathbf{S}_1 = \mathbf{S}'_1 \mathbf{M}_1$ in which $\mathbf{S}'_1 \neq \mathbf{S}_1$. Accordingly, in this procedure \mathbf{M}_1 is associated with the optimum Mueller matrix of Eqs. (5.4).

An alternative algorithm, proposed in Refs. [58] and [57], minimizes the error through constrained optimization of the least-squares cost function

$$\|\mathbf{Y}_1 - \mathbf{M}_1 \mathbf{Y}_0\|^2 - \text{Tr}(\mathbf{L}(\mathbf{M}_1^+ \mathbf{M}_1 - \mathbf{I})), \quad (5.8)$$

with respect to \mathbf{M}_1 , where the elements of the matrix \mathbf{L} are Lagrange multipliers that implement the constraint $\mathbf{M}_1^+ \mathbf{M}_1 = \mathbf{I}$. The above equation for \mathbf{M}_1 is then minimized iteratively.

The PMD and PDL coefficients $\vec{\Omega}_m$ and $\vec{\Lambda}_m$ can be determined from any of the methods above after finite-differencing $\mathbf{N}(\Delta\omega) = \ln(\mathbf{M}_1)$. This yields an estimate for the m^{th} derivative of the evolution operator, denoted here by $\mathbf{H}_m^{\text{est}}$, from Eq. (5.3) [52]. The step size between adjacent frequencies, $\Delta\omega = \omega_1 - \omega_0$, however, must be sufficiently small that the matrix logarithm is confined to its principal branch. We employ a five-point finite difference approximation to the derivative operator to obtain $\mathbf{H}_0^{\text{est}}$ and $\mathbf{H}_1^{\text{est}}$ from least-squares estimates of \mathbf{M}_1 at the frequencies $\omega_0 \pm \Delta\omega$ and $\omega_0 \pm 2\Delta\omega$.

5.3 Numerical results

We now compare numerically our least-squares estimate of $\mathbf{H}_m^{\text{est}}$ for a varying number of Stokes vector measurements, N . Our fiber model consists of a sequence of 100 randomly oriented linearly birefringent but lossy segments selected such that the mean PMD and PDL of the link is 15.0 ps and 3.0 dB, respectively. The matrices \mathbf{Y}_1 and \mathbf{Y}_0 are constructed from the output polarizations obtained from N random, numerically generated input states after adding additive Gaussian noise to \mathbf{Y}_1 with a standard deviation of either 1.0% or 3.0% of $|\tilde{t}(\omega_0)|$. It should be noted, however, that standard polarization measurements exhibit a far smaller measurement error than indicated by these values.

In Figs. 5.1 and 5.2 we display the variation of both the average, $\langle\epsilon\rangle$, and the standard deviation, σ_ϵ , of the relative error,

$$\epsilon = \frac{\|\mathbf{H}_0^{\text{est}} - \mathbf{H}_0\|}{\|\mathbf{H}_0\|} \times 100\%, \quad (5.9)$$

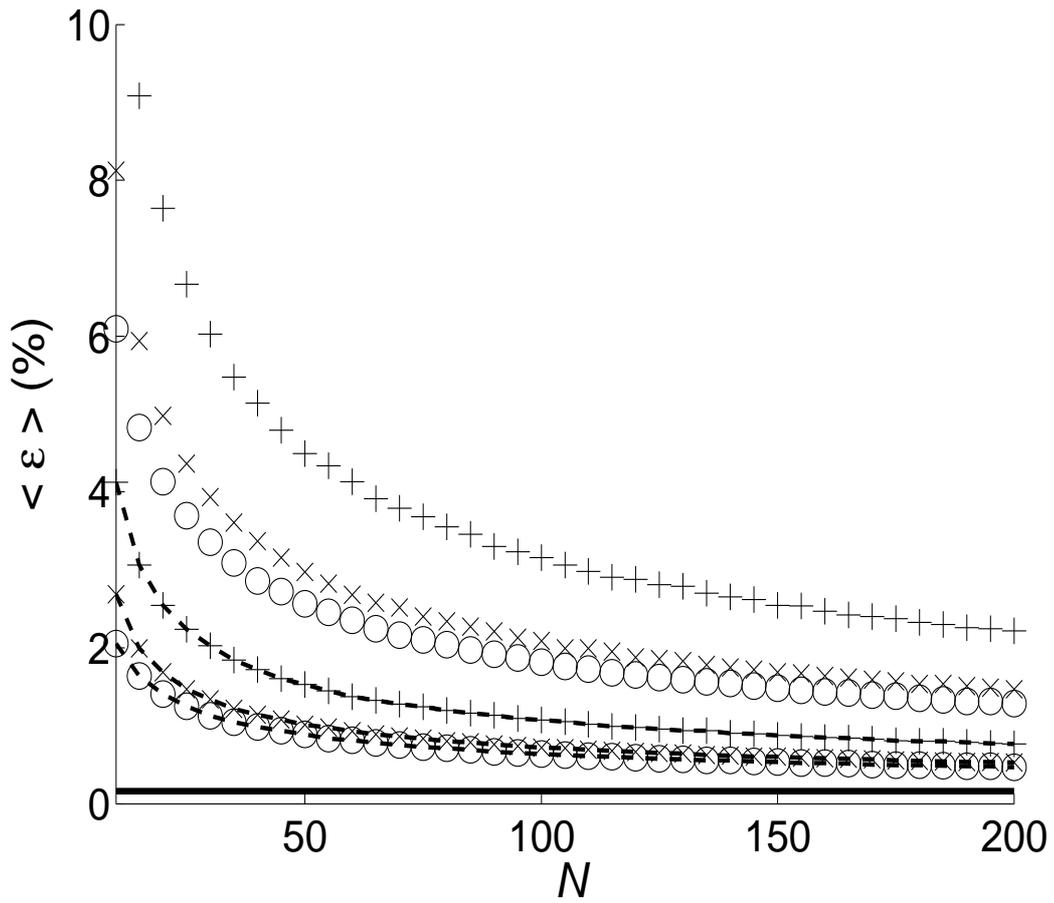


Figure 5.1: Average relative error, $\langle \epsilon \rangle$, as a function of the number of Stokes vector measurements, N . The +, \times and \circ markers denote $\langle \epsilon \rangle$ calculated using Eqs. (5.5), (5.6) and (5.8), respectively. Dashed lines - results for 1.0% additive noise, markers only - 3.0% additive noise.

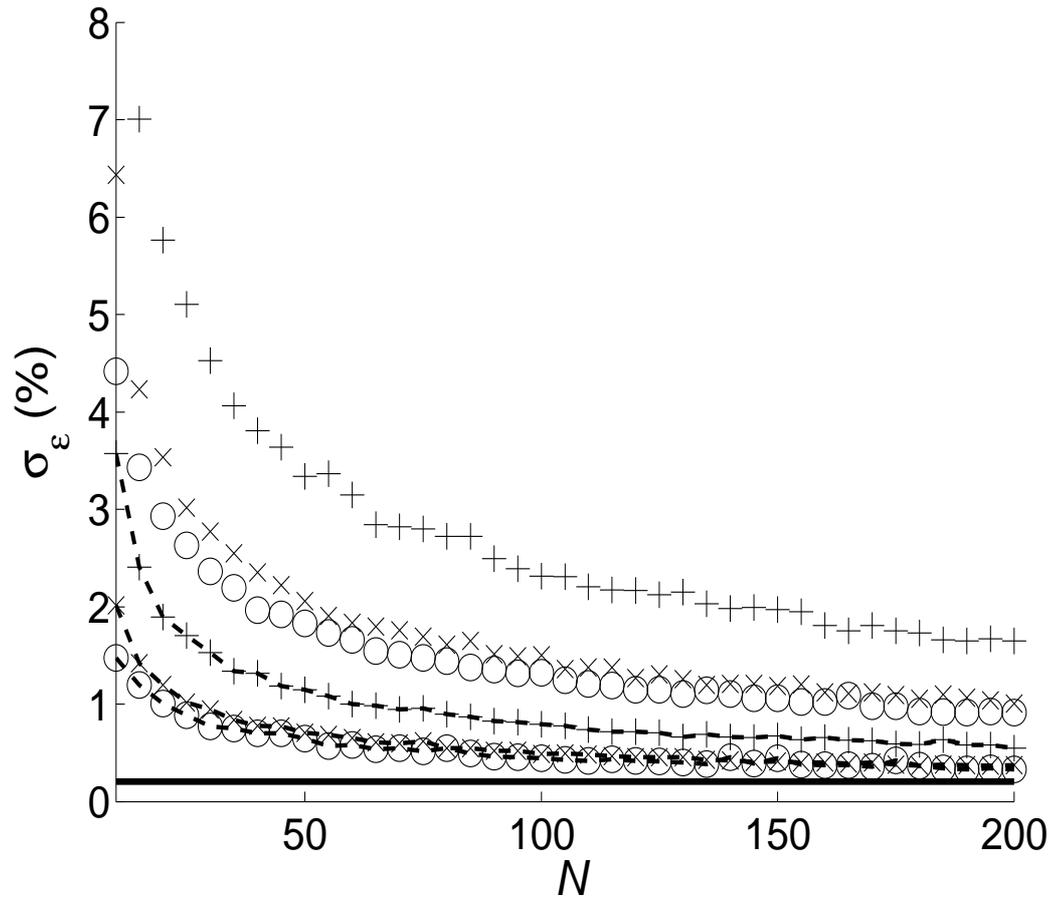


Figure 5.2: As in Fig. 5.1 but for the standard deviation of the relative error, σ_ϵ

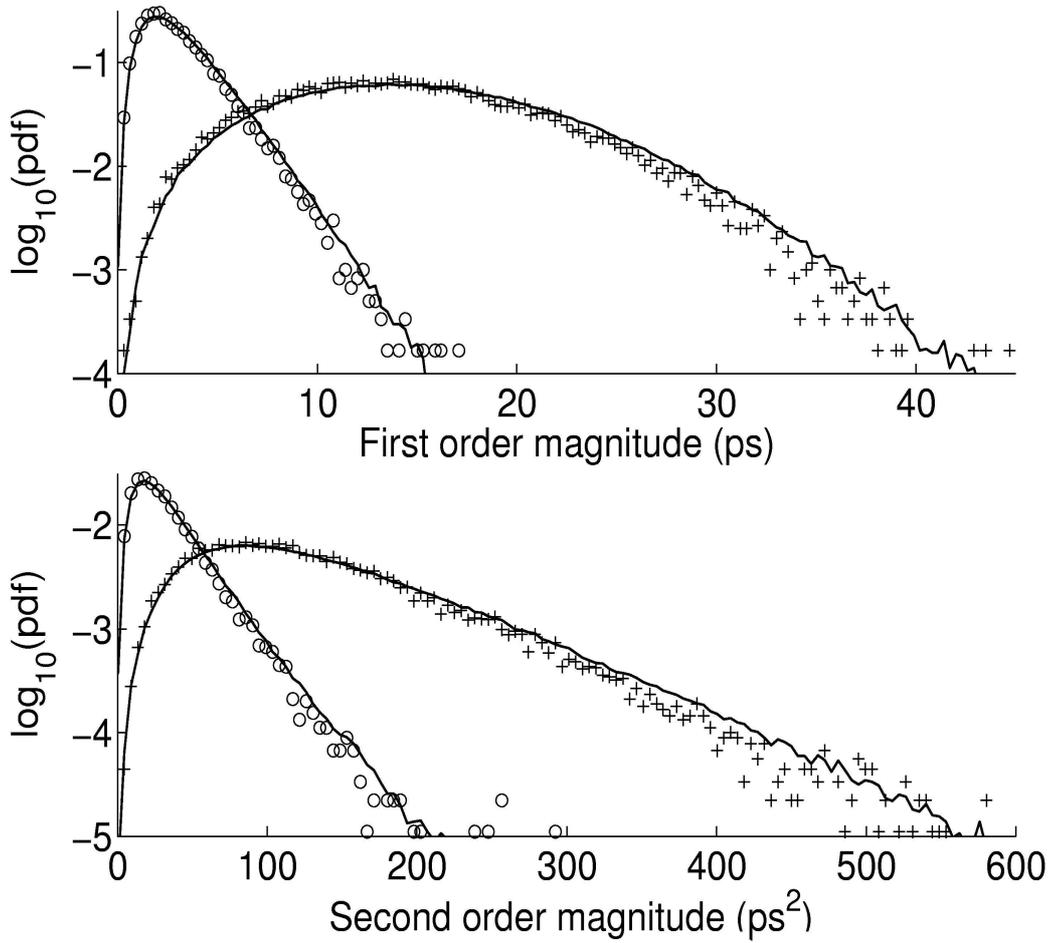


Figure 5.3: Histograms of (top) $|\vec{\Lambda}_0|$ (circles), and $|\vec{\Omega}_0|$ (crosses), and (bottom) $|\vec{\Lambda}_1|$ (circles), and $|\vec{\Omega}_1|$ (crosses) estimated from the least-squares algorithm of Eq. (5.6).

as a function of the number, N , of sampled Stokes vectors. Each statistical average, $\langle \dots \rangle$, is performed over 20,000 fiber realizations. The exact result, \mathbf{H}_0 , is evaluated by finite differencing the numerically determined Mueller matrix. The values of $\mathbf{H}_0^{\text{est}}$ obtained from Eq. (5.5), Eq. (5.6) and the minimization of Eq. (5.8), in which the finite differences were performed with a 0.05nm wavelength step size, are denoted by +, \times and \circ markers. Further, the dashed curves, isolated markers and solid curves were obtained at numerical noise levels of 1.0%, 3.0% and 0%, respectively. In the last case, for all the least-squares algorithms considered here the average and standard deviation of the relative error depend only on the frequency step size and are independent of the number of input polarizations per measurement. While all curves, as expected from statistical arguments, converge as $\langle \epsilon \rangle \sim 1/\sqrt{N}$, the average relative error of our simplified procedure, Eq. (5.6), is comparable to that of the explicit minimization method of Eq. (5.8).

Finally, we analyze the probability density functions of the PMD and PDL, obtained from our simplified algorithm, Eq. (5.6). Each simulated measurement consists of 50 numerically sampled output Stokes vector evaluations for a wavelength step size of 0.05nm, together with 1.0% additive numerical noise. After 20,000 iterations, we obtain the top graph of Fig. 5.3 for the magnitudes of the first order vectors, $|\vec{\Lambda}_0|$ (circles) and $|\vec{\Omega}_0|$ (crosses). The solid curves in the figure were instead generated by finite differencing the exact Mueller matrix of Eq. (5.3). The second graph in Fig. 5.3 presents the corresponding results for the second order vectors $|\vec{\Lambda}_1|$, and $|\vec{\Omega}_1|$. Clearly, the agreement with theory further confirms the accuracy of our least-squares estimation procedure.

5.4 Conclusions

In conclusion, we have examined several methods for estimating the Mueller matrix transformation between the output field polarizations at adjacent optical frequencies for a fixed input polarization direction from repeated measurements of output Stokes vectors. A simple computational method proved highly accurate and stable against simulated mea-

surement error. In particular, for our PMD and PDL values, the resulting estimate of the first-order complex principal state vector for measurements of ≈ 50 output polarization samples yielded an accuracy of $< 3\%$ even in the presence of 3% additive Gaussian noise. Since these measurements can be performed with fast, repeatable polarization rotators such as multi-stage lithium niobate devices coupled to high-speed polarimeters, such techniques can be readily implemented experimentally and are applicable to general optical systems as well.

Chapter 6

Efficient Jones Matrix Expansions

Next, we estimate the frequency-dependence of the Jones matrix of an optical fiber from its values at several discrete optical frequencies and illustrate our procedures through calculations of high-order polarization-mode dispersion in optical fibers.

6.1 Introduction

To simulate quantities such as the optical pulse distortion in an optical communication channel, the Jones matrix is typically evaluated at many different frequencies within the signal bandwidth, leading to large evaluation times when performing statistical averages over channel configurations. We here introduce methods for decreasing the number of required evaluation points for a given level of accuracy. Our procedures yield up to an order of magnitude reduction in computation time in studies of optical pulse distortion compared to standard techniques based on fiber emulators.

Our first procedure simultaneously interpolates both the Jones matrix and the polarization-mode dispersion (PMD) vector of the optical fiber from their calculated or measured values at several discrete optical frequencies through a unitary matrix representation [24, 56, 97].

A second set of methods instead expand the Jones matrix around the carrier frequency in either a Taylor or Chebyshev series or a Padé approximant [81, 115, 116]. We find that the bandwidth of the Chebyshev series and the Padé approximant increases linearly as additional expansion orders are included, unlike that of the Taylor series expression which displays a square-root dependence on the order number [20, 59]. For a given frequency spacing between input data points, the error of our simplified Jones matrix interpolation procedure is found to be comparable to that of a high-order Taylor expansion.

6.2 Numerical Techniques

The unitary Jones matrix of an optical fiber, $\mathbf{T}(\omega)$, and the output PMD vector, $\vec{\Omega}(\omega) \cdot \vec{\sigma} = 2i(d\mathbf{T}/d\omega)\mathbf{T}^{-1}$, with $\vec{\sigma}$ the Pauli spin vector, are normally measured or calculated at N_ω frequencies $\omega_n = \omega_0 + n\delta$, around the optical carrier, ω_0 , for any multiple n of the fundamental frequency offset δ that enters into our subsequent expressions. We here construct a unitary Jones matrix, $\mathbf{U}(\omega)$, that optimally interpolates between the discrete data values $\mathbf{T}_{(n)} \equiv \mathbf{T}(\omega_n)$ and $\vec{\Omega}_{(n)} \equiv \vec{\Omega}(\omega_n)$. While, for simplicity, this paper only considers uniformly spaced ω_n , our results can be immediately generalized to unequal frequency separations.

Our first procedure evaluates two additional Jones matrices

$$\begin{aligned} \mathbf{A}_{(n)} &= \exp\left[-\frac{i\delta}{6}\vec{\Omega}_{(n)} \cdot \vec{\sigma}\right] \mathbf{T}_{(n)} \\ \mathbf{B}_{(n)} &= \exp\left[\frac{i\delta}{6}\vec{\Omega}_{(n+1)} \cdot \vec{\sigma}\right] \mathbf{T}_{(n+1)} \end{aligned} \quad (6.1)$$

in each frequency interval $[\omega_n, \omega_{n+1}]$ [56]. In the case of zero polarization-dependent loss (PDL), the interpolating Jones matrix, $\mathbf{U}(\omega)$, is given by

$$\mathbf{U}(\omega) \equiv [\mathbf{T}_{(n+1)}\mathbf{B}_{(n)}^{-1}]^{\beta_3(t)} [\mathbf{B}_{(n)}\mathbf{A}_{(n)}^{-1}]^{\beta_2(t)} [\mathbf{A}_{(n)}\mathbf{T}_{(n)}^{-1}]^{\beta_1(t)} \mathbf{T}_{(n)} \quad (6.2)$$

with

$$\begin{aligned}
\beta_1(t) &= 1 - (1 - t)^3 \\
\beta_2(t) &= t^2(3 - 2t) \\
\beta_3(t) &= t^3,
\end{aligned} \tag{6.3}$$

for $t = (\omega - \omega_n)/\delta$ and $\omega_n \leq \omega \leq \omega_{n+1}$ such that $0 \leq t \leq 1$, see Appendix C.4. A similar interpolation of the input PMD vectors [59] can be performed by reversing the order of Eqs. (6.1) and (6.2), and replacing $\vec{\Omega}$ by $\vec{\Omega}_{\text{input}}$.

Since $\mathbf{U}(\omega)$ is represented as the product of unitary matrices in Eq. (6.2), its unitarity is guaranteed [24, 97]. Further, by substitution $\mathbf{U}(\omega_n) \equiv \mathbf{T}_{(n)}$ and $\mathbf{U}(\omega_{n+1}) = \mathbf{T}_{(n+1)}$, confirming that Eq. (6.2) interpolates the discrete Jones matrix data, while evaluating $d\mathbf{U}/d\omega$ similarly yields the input values $\vec{\Omega}_{(n)}$. That is, if the unitary matrix $\exp[-i\psi(\hat{n} \cdot \vec{\sigma})/2]$ is denoted by \mathbf{F} with ψ and \hat{n} a constant scalar and unit vector, respectively, then for any function $f(t)$,

$$\frac{d\mathbf{F}^{f(t)}}{dt} = \left(\frac{df}{dt} \right) (\ln \mathbf{F}) \mathbf{F}^{f(t)} \tag{6.4}$$

while

$$\mathbf{F}^{f(t)} = \exp \left[-\frac{i}{2} \psi f(t) (\hat{n} \cdot \vec{\sigma}) \right] \tag{6.5}$$

Here $\ln \mathbf{F} \equiv -i\psi(\hat{n} \cdot \vec{\sigma})/2$ [24]. Accordingly, differentiating $\mathbf{U}(\omega)$ yields, as required [24, 56],

$$\left[\frac{d\mathbf{U}}{d\omega} \mathbf{U}^{-1} \right] \Big|_{\omega_n} = -\frac{i}{2} \vec{\Omega}_{(n)} \cdot \vec{\sigma} \tag{6.6}$$

and

$$\left[\frac{d\mathbf{U}}{d\omega} \mathbf{U}^{-1} \right] \Big|_{\omega_{n+1}} = -\frac{i}{2} \vec{\Omega}_{(n+1)} \cdot \vec{\sigma} \tag{6.7}$$

An alternative method is obtained from the observation that the accuracy of the Taylor series expansion of the Jones matrix $\mathbf{T}(\omega_0 + \Delta\omega) = \sum_{k=0}^N \mathbf{T}_k \Delta\omega^k$, with

$$\mathbf{T}_k \equiv \frac{1}{k!} \left. \frac{d^k \mathbf{T}(\omega)}{d\omega^k} \right|_{\omega_0} \tag{6.8}$$

at large $\Delta\omega$ is considerably less than that of the $[Q, M]$ -Padé approximant [81, 116]

$$\sum_{k=0}^N \mathbf{T}_k \Delta\omega^k = \left[\sum_{q=0}^Q \mathbf{J}_q \Delta\omega^q \right] \left[\mathbf{I} + \sum_{m=1}^M \mathbf{K}_m \Delta\omega^m \right]^{-1} \quad (6.9)$$

For $Q+M \leq N$, the coefficients \mathbf{J}_q and \mathbf{K}_m are the solutions of a linear system of $Q+M+1$ equations involving the \mathbf{T}_k . These can be generated by equating equal orders of $\Delta\omega$ after multiplying both sides of Eq. (6.9) by the denominator of the Padé approximant [81]. This yields

$$\begin{bmatrix} \mathbf{K}_1 \\ \vdots \\ \mathbf{K}_M \end{bmatrix} = - \begin{bmatrix} \mathbf{C}_{11} & \dots & \mathbf{C}_{1M} \\ \vdots & & \vdots \\ \mathbf{C}_{M1} & \dots & \mathbf{C}_{MM} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{T}_{Q+1} \\ \vdots \\ \mathbf{T}_{Q+M} \end{bmatrix} \quad (6.10)$$

The elements of the block matrix \mathbf{C} are 2×2 matrices given for $n, m \in [1, \dots, M]$ by $\mathbf{C}_{nm} \equiv \mathbf{T}_{Q+n-m}$ if $Q+n-m \geq 0$ and by the 2×2 null matrix $\mathbf{0}$ otherwise. Additionally, $\mathbf{J}_0 = \mathbf{T}_0$ while

$$\begin{bmatrix} \mathbf{J}_1 \\ \vdots \\ \mathbf{J}_Q \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1 \\ \vdots \\ \mathbf{T}_Q \end{bmatrix} + \begin{bmatrix} \mathbf{D}_{11} & \dots & \mathbf{D}_{1M} \\ \vdots & & \vdots \\ \mathbf{D}_{Q1} & \dots & \mathbf{D}_{QM} \end{bmatrix} \begin{bmatrix} \mathbf{K}_1 \\ \vdots \\ \mathbf{K}_M \end{bmatrix} \quad (6.11)$$

in which $\mathbf{D}_{qm} \equiv \mathbf{T}_{q-m}$ and $\mathbf{D}_{qm} = \mathbf{0}$ for $q \geq m$ and $q < m$, respectively, with $q \in [1, \dots, Q]$, $m \in [1, \dots, M]$.

A third procedure evaluates the Jones matrix, $\mathbf{T}(\omega)$, at the zeros

$$\omega_{(k)} = \omega_0 + \Delta\omega \cos \left[\frac{\pi}{2N} (2k+1) \right] \quad (6.12)$$

$k = 0, \dots, N-1$, of the N^{th} order Chebyshev polynomial of the first kind in the interval $|\omega - \omega_0| \leq \Delta\omega$. Chebyshev series expansion coefficients are then derived from standard orthogonality conditions yielding the approximation [115]

$$\mathbf{T}(\omega) \approx -\frac{1}{2} \mathbf{N}_0 + \sum_{j=0}^{N-1} \mathbf{N}_j T_j \left(\frac{\omega - \omega_0}{\Delta\omega} \right) \quad (6.13)$$

with

$$\mathbf{N}_j = \frac{2}{N} \sum_{k=0}^{N-1} \mathbf{T}(\omega_{(k)}) \cos\left(\frac{\pi j}{2N}(2k+1)\right) \quad (6.14)$$

Here $j = 0, \dots, N-1$, and $T_j(x)$ denotes the Chebyshev polynomial of the first kind of order j with $-1 \leq x \leq 1$ [81].

To implement our procedures numerically, c.f. Sec. 6.3, we here introduce exact formulas for the derivatives of the Jones matrix, \mathbf{T}_k , associated with an N_{seg} section fiber emulator. We start with the recursive expression for the Jones matrix after m linearly birefringent sections, $\mathbf{T}^{(m)}(\omega) = [\mathbf{R}^{(m)}]^\omega \mathbf{T}^{(m-1)}(\omega)$, where $\mathbf{R}^{(m)} \equiv \exp[-i\tau^{(m)}(\hat{\mathbf{n}}^{(m)} \cdot \vec{\sigma})/2]$ and $[\mathbf{R}^{(m)}]^\omega = \exp[-i\omega\tau^{(m)}(\hat{\mathbf{n}}^{(m)} \cdot \vec{\sigma})/2]$, in which the differential group delay (DGD) of the m^{th} section, $\tau^{(m)}$, and the unit vector $\hat{\mathbf{n}}^{(m)}$ are both independent of the optical frequency ω . For each $k = 0, \dots, N$ we differentiate $\mathbf{T}^{(m)}(\omega)$, with $m = 1, \dots, N_{\text{seg}}$, k times, apply

$$\frac{d[\mathbf{R}^{(m)}]^\omega}{d\omega} = [\ln \mathbf{R}^{(m)}][\mathbf{R}^{(m)}]^\omega, \quad (6.15)$$

and finally eliminate terms of the form $[\mathbf{R}^{(m)}]^\omega [d^{k-1}\mathbf{T}^{(m-1)}/d\omega^{k-1}]$ by substituting for $\mathbf{T}_{k-1}^{(m)}$. With the initial conditions $\mathbf{T}_0^{(0)} = \mathbf{I}$ and $\mathbf{T}_{k \geq 1}^{(0)} = \mathbf{0}$ this yields, c.f. Appendix A.4,

$$\mathbf{T}_k^{(m)} = [\mathbf{R}^{(m)}]^\omega \mathbf{T}_k^{(m-1)} - \sum_{q=1}^k \frac{1}{q!} [-\ln \mathbf{R}^{(m)}]^q \mathbf{T}_{k-q}^{(m)} \quad (6.16)$$

Accordingly, we can conveniently obtain $\mathbf{T}_k^{(m)}$ without first determining $k-1$ derivatives of the PMD vector. Note that if this method is instead applied to the PMD concatenation rule [59] in which the PMD vector $\tau^{(m)}\hat{\mathbf{n}}^{(m)}$ of the m^{th} segment is taken as frequency-independent, we generate exact recursive formulas for k^{th} order PMD vectors.

6.3 Numerical Results

We now compare the procedures of Sec. 6.2 through an analysis of an $N_{\text{seg}} = 100$ section fiber emulator with an overall mean DGD, $\tau_{\text{avg}} = 20$ ps, so that $\tau^{(m)} = \tau_{\text{avg}}\sqrt{3\pi/8N_{\text{seg}}} =$

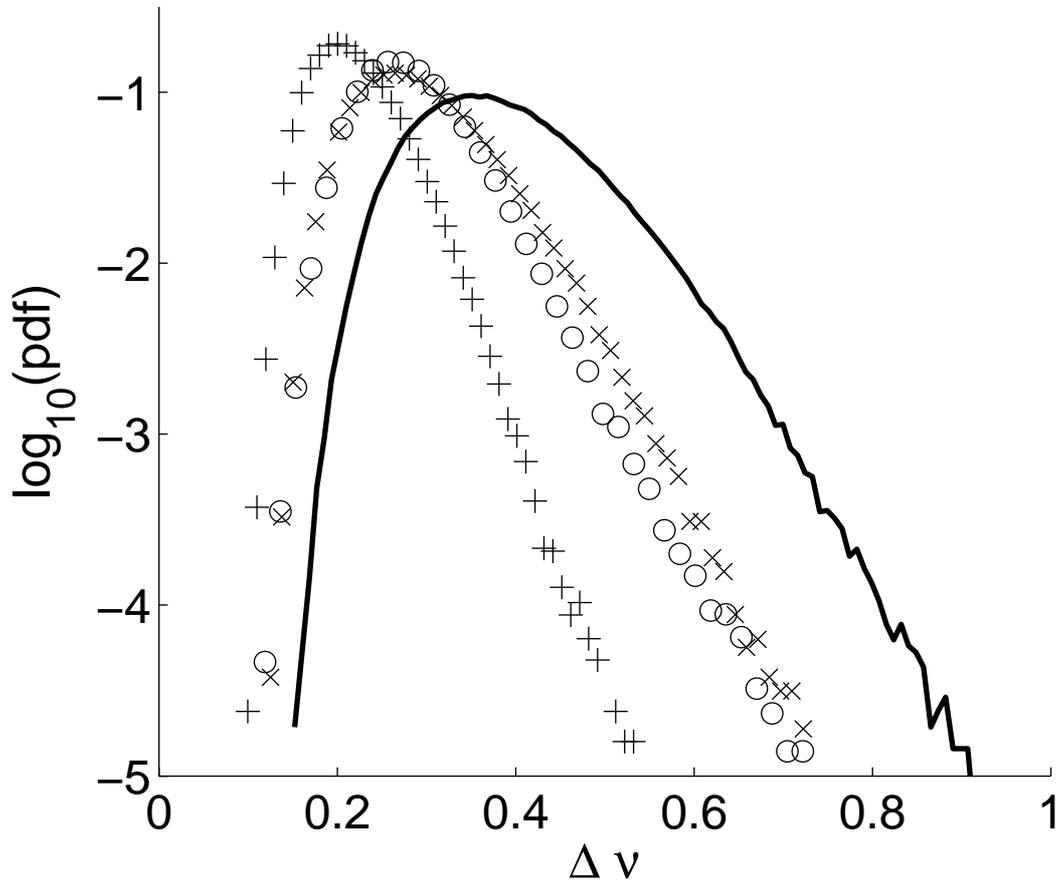


Figure 6.1: The pdf of the normalized bandwidth, $\Delta\nu = \tau_{avg}\Delta\omega_\alpha/2\pi$, where $\epsilon_\alpha(\Delta\omega_\alpha) = \epsilon_{max} = 2\%$ for each configuration, $\alpha = 1, 2, \dots, (N_\alpha = 5 \times 10^5)$, of a $N_{seg} = 100$ fiber emulator as calculated with the fourth order Chebyshev (\times), and Taylor series ($+$) expansions as well as the $[2, 2]$ -Padé approximant (\circ) and the interpolation procedure of Eq. (6.2) (solid line).

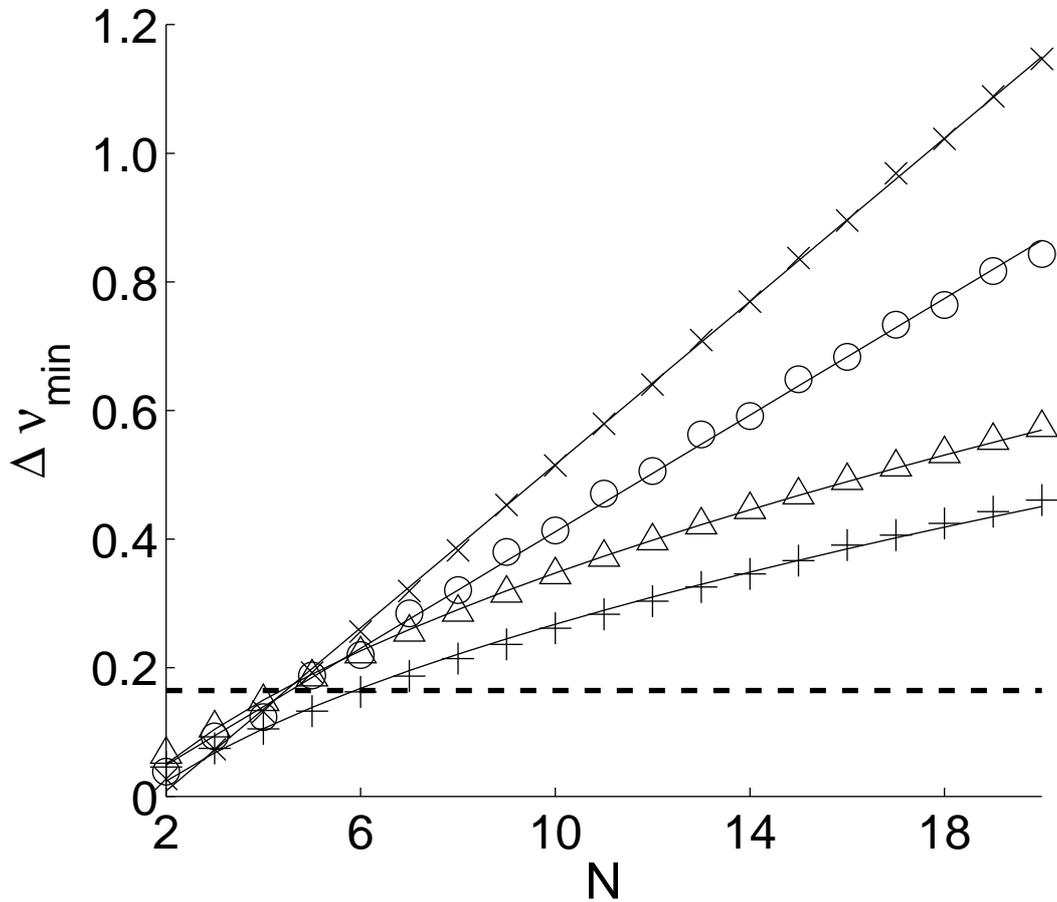


Figure 6.2: The minimum value, $\Delta\nu_{\min}$, of $\Delta\nu$ that limits the maximum recorded error, Eq. (6.17), $< 2\%$ for 99.99% of the randomly generated samples graphed as a function of expansion order, N . The \times , $+$ and \circ markers represent Chebyshev, Taylor and Padé approximant expansions, respectively, while the horizontal dashed line represents $\Delta\nu_{\min}$ from Fig. 6.1, of Eq. (6.2). The Δ markers correspond to the $\Delta\nu_{\min}$ obtained if the Taylor series error is limited to $< 2\%$ instead for 95% of the samples. The superimposed solid lines are optimized linear and square-root function approximations.

2.17 ps for each $m = 1, \dots, N_{\text{seg}}$ while the $\hat{n}^{(m)}$ are chosen from a uniform statistical distribution over the Poincaré sphere. The maximum error, $\epsilon_\alpha(\Delta\omega)$, between the estimated, $\mathbf{T}_{\text{est}}(\omega)$, and exact Jones matrices, $\mathbf{T}(\omega)$, over the frequency interval $0 \leq \omega - \omega_0 \leq \Delta\omega$ is given for each realization, α , of the emulator by

$$\epsilon_\alpha(\Delta\omega) = \max_{0 \leq \omega - \omega_0 \leq \Delta\omega} \left[\frac{\|\mathbf{T}_{\text{est}}(\omega) - \mathbf{T}(\omega)\|}{\|\mathbf{T}(\omega)\|} \right] \cdot 100\%, \quad (6.17)$$

where $\|\mathbf{T}\| \equiv \sqrt{\text{Tr}(\mathbf{T}^\dagger \mathbf{T})}$ is the Frobenius matrix norm, \dagger represents Hermitian conjugation and $\text{Tr}(\dots)$ indicates a trace. Next, for each $\alpha = 1, 2, \dots, (N_\alpha = 5 \times 10^5)$ a bisection method [81] is employed to determine the bandwidth, $\Delta\omega_\alpha$, such that $\epsilon_\alpha(\Delta\omega_\alpha) = \epsilon_{\text{max}}$ for a maximum specified error, ϵ_{max} . The results of N_α such calculations for each of the numerical techniques of Sec. 6.2 yield corresponding estimates of the probability density function (pdf) of $\Delta\omega_\alpha$. The advantage of such an approach as opposed to, e.g. calculation of the system penalty function, Q , is that Eq. (6.17) is independent of a specific hardware implementation and therefore provides an unambiguous measure of the relative accuracy of our procedures.

Our results are typified by the pdf of $\Delta\omega_\alpha$ for $\epsilon_{\text{max}} = 2\%$ as a function of the normalized frequency $\Delta\nu = \tau_{\text{avg}} \Delta\omega_\alpha / 2\pi$ as displayed in Fig. 6.1. In this figure \mathbf{T}_{est} is approximated by a fourth order Chebyshev polynomial (\times), a fourth order Taylor expansion ($+$), a [2, 2]-Padé approximant (\circ), and the interpolation method of Eq. (6.2) (solid line) for which $\delta \rightarrow \Delta\omega$ is the spacing between adjacent frequency values. Clearly Eq. (6.2), the [2, 2]-Padé approximant and the fourth order Chebyshev series significantly increase the frequency range over which the Jones matrix can be accurately estimated compared to the fourth order Taylor series procedure.

Next, to determine the dependence of the range of accuracy of each of the methods of Sec. 6.2 on the expansion order N , for $N = 1, 2, \dots, 20$, we determine as in the calculation of Fig. 6.1 the normalized frequency, $\Delta\nu_{\text{min}}$, such that the probability of $\Delta\nu$ exceeding $\Delta\nu_{\text{min}}$, $P(\Delta\nu \geq \Delta\nu_{\text{min}})$, is 99.99%. Fig. 6.2 then displays the variation of $\Delta\nu_{\text{min}}$ with N in which the \circ , $+$ and \times markers indicate the Padé approximant, Taylor series and

Chebyshev expansions, respectively, while the horizontal dashed line corresponds to $\Delta\nu_{\min}$ from Fig. 6.1 with the Jones matrix interpolation of Eq. (6.2). Here we have employed a Padé approximant with $Q = M = N/2$ for even N and $Q = (N - 1)/2$ and $M = (N + 1)/2$ for odd N (our results were found to be relatively independent of the values chosen for Q and $M = N - Q$).

Although $\Delta\nu < \Delta\nu_{\min}$ with a probability $1 - P(\Delta\nu \geq \Delta\nu_{\min})$, the qualitative behavior of the variation of $\Delta\nu_{\min}$ with expansion order, N , is independent of $P(\Delta\nu \geq \Delta\nu_{\min})$. To illustrate, the Δ markers of Fig. 6.2 represent the Taylor series expansion method results with $\Delta\nu_{\min}$ instead determined by $P(\Delta\nu \geq \Delta\nu_{\min}) = 95\%$. Despite our inability to determine a universal value of $\Delta\nu_{\min}$ such that $\Delta\nu \geq \Delta\nu_{\min}$ for all possible emulator realizations, the rapid decrease of the pdf near $\Delta\nu \approx 0.15$ in Fig. 6.1, indicates that a small reduction in $\Delta\nu_{\min}$ substantially decreases the probability that the estimation error exceeds ϵ_{\max} .

Further, we note from the solid lines in Fig. 6.2, which are optimal fits to linear and square-root functions, that $\Delta\nu_{\min}$ increases as $\sim\sqrt{N}$ for the Taylor series expansion [20], but varies instead linearly with N for both the Chebyshev expansion and Padé approximant up to $N = 20$ (although deviations from linearity are evident in the Padé approximant results for large N). The error associated with the Taylor series additionally exceeds that of the Padé approximant for $N \geq 4$. Further, the $\Delta\nu_{\min}$ corresponding to Eq. (6.2) is comparable to that of a sixth order Taylor series, fifth order Chebyshev or a [2, 3]-Padé approximant. Accordingly, Eq. (6.2) provides a highly accurate method for estimating the frequency dependence of the Jones matrix from the values of the Jones matrix and the PMD vector at discrete optical frequencies.

It should be noted as well that many other expansion and interpolation techniques exist that could also be successfully employed in specific contexts. These include most notably the Padé-Chebyshev [116], the natural cubic spline [81] and quaternion spline [97] interpolation procedures. However, for the numerical example of this paper, the Padé-Chebyshev expansion, which results from an expansion in Chebyshev polynomials of the first kind

as in Eq. (6.9), provided only minor improvement over the standard Chebyshev method of Sec. 6.2 at the cost of significant additional computational complexity. Natural cubic spline interpolation, which separately interpolates the two complex values that uniquely specify the unitary matrix $\mathbf{T}(\omega)$ [81], similarly yielded an error comparable to that of the third order Taylor series and therefore would be relevant when the Jones matrix inputs are only available at discrete optical frequencies. Finally, while Ref. [97] provides algorithms for interpolating unit-quaternions, the error in our example exceeds that obtained even with standard cubic spline procedures.

6.4 Conclusions

While the frequency behavior of the Jones matrix is normally characterized by PMD coefficients, other parameterizations can be preferable in the presence of large high-order PMD. For example, we demonstrated above that the bandwidth of the Padé approximant varies linearly with expansion order, unlike the square-root dependence of the standard PMD expansion. We also presented a highly-accurate quaternion procedure [24, 56] that interpolates the Jones matrix from its values together with the corresponding PMD vectors at a given discrete set of optical frequencies. The increased numerical accuracy and programming simplicity afforded by such methods should yield improved simulation methods for high bit-rate single or multiple channel optical systems.

Chapter 7

Alignment Methods for Biased Multicanonical Sampling

The efficiency of the multicanonical procedure can be significantly improved by applying an additional bias to the numerically generated sample space. However, results obtained by biasing in different sampling regions cannot in general be accurately combined since their relative normalization coefficients are not known precisely. In this chapter, we demonstrate that for overlapping biasing regions a simple iterative procedure can be employed to determine the required coefficients.

7.1 Introduction

Communication system studies normally involve the estimation of small error probabilities that are associated with physically unlikely system configurations. Numerical system simulations have therefore recently employed the multicanonical method [15, 112–114] to increase the likelihood of physically interesting sampling events. In our initial studies of the pdf of the differential group delay (DGD) of an optical fiber [37, 45, 54] we found

that multicanonical procedures reduced the number of samples required to predict or measure the low-probability region of the pdf by orders of magnitude. However, the increase in computational efficiency, while large, can still be insufficient for experimental applications [69, 70], in which orders of magnitude fewer system configurations can be sampled compared to numerical simulations. Accordingly several improvements to multicanonical sampling have been proposed [68, 117], including biased multicanonical approaches that concentrate the statistical samples into restricted regions of the space of system observables. With appropriate bias functions, we performed measurements of the pdf of the DGD to an unparalleled level of precision with standard equipment [69].

However, if our method is applied repeatedly to the same problem but with the sample space biased in different regions, the associated histograms cannot be combined to predict the global system behavior since their relative normalization is undetermined. Here we present a procedure for calculating the required normalization constants when the sampling regions of successive biased multicanonical iterations (or separate multicanonical calculations) overlap. In particular, the relative normalization constants are obtained from an iterative procedure in which the converged solution minimizes the mean square error (MSE) of the pdf estimate in the overlapping regions. Our method yields a substantial increase in computational efficiency with modest additional programming effort.

Methods similar to ours have been previously applied to the Monte Carlo method and to the canonical distribution [5] generated by the Metropolis algorithm. These include successive umbrella sampling [102, 103, 108], bridge sampling [12], importance sampling and the weighted histogram analysis method (WHAM) of Refs. [33, 61, 101]. However, the procedure of this chapter is, to our knowledge, the first implementation of such techniques to the multicanonical ensemble [13, 14, 100, 117].

7.2 Multicanonical sampling

The multicanonical method is an iterative, biased statistical sampling method that presupposes a complete absence of prior information about the system properties. The procedure generates the value of any function of system variables, $p'(\vec{E})$, weighted by its probability of occurrence which is collectively denoted by $p(\vec{E})$. The $\vec{E}(\vec{\alpha})$ are a vector of N_E system observables that are dependent on N_α stochastic parameters, $\vec{\alpha}$. In the calculations of this chapter, the observables correspond to the DGD of an optical fiber system, the randomly varying parameters, $\vec{\alpha}$ are the angles of various polarization controllers [69, 70], and $p(\vec{E})$ is the pdf associated with finding the system in a given state of the observable. However, it should be emphasized that the choice of these quantities is effectively arbitrary (in the more general case, which is discussed and illustrated in detail in Ref. [117], the histograms below must be updated by a different value than unity).

To implement the multicanonical method, the physically relevant region of the output values (solution space) \vec{E} is divided into N histogram bins with volume ΔV centered at \vec{E}_i with $i = 1, 2, \dots, N$. A histogram that here corresponds to the current estimate of the (unnormalized) pdf, which we denote p_0 , and a second histogram, $H_1(\vec{E})$, that stores temporary values, are first set to unity, consistent with an absence of prior information. We then generate a set of random system variables, $\vec{\alpha}^{cur}$, and compute the associated observables \vec{E}^{cur} . A transition is then generated from $\vec{\alpha}^{new} = \vec{\alpha}^{cur} + \Delta\vec{\alpha}$ where $\Delta\vec{\alpha}$ represents a small random perturbation (which can be chosen from an effectively arbitrary probability distribution, as verified in Ref. [117]), leading to new observables \vec{E}^{new} . This transition is accepted with probability $\min[1, p(\vec{E}^{cur})/p(\vec{E}^{new})]$, in which case the system variables $\vec{\alpha}^{cur}$ are replaced by $\vec{\alpha}^{new}$; otherwise a self-transition occurs and $\vec{\alpha}^{new}$ is set to $\vec{\alpha}^{cur}$ for the following step [112]. In either case, the value of the histogram bin $H_1(\vec{E})$ corresponding to the updated \vec{E}^{cur} is incremented. After N_s transitions, a new, rebased, estimate of $p \equiv p_1$ is generated from $p_1(\vec{E}) = cp_0(\vec{E})H_1(\vec{E})$, where the normalization constant c is normally chosen according to the criterion that p_1 corresponds to a probability distribution. While this procedure initially leads to the Monte Carlo result, in the subsequent iteration, $p_0(\vec{E})$

is replaced by $p_1(\vec{E})$ and the above steps are repeated. The acceptance rule then increases the sampling probability of states with small $p(\vec{E})$ in such a manner that the resulting histogram becomes increasingly independent of \vec{E} as the calculation progresses.

7.3 Biased multicanonical sampling

Biasing methods increase the multicanonical sampling probability in a limited region \mathcal{R}_k of the system observables. One procedure multiplies the current estimate of $p(\vec{E})$ by a bias enhancement function $F(\vec{E})$ that is constant within \mathcal{R}_k , and rapidly increases away from this region, i.e. $\tilde{p}_{m-1}(\vec{E}) = p_{m-1}(\vec{E})F(\vec{E})$ at the beginning of each iteration [68]. The resulting modified transition rule augments the probability of transitions into \mathcal{R}_k ; however, at the end of the iteration loop, an unbiased estimate of p is recovered from $p_m(\vec{E}) = \tilde{p}_{m-1}(\vec{E})H_m(\vec{E})$ as in the importance sampling technique. In the one-dimensional calculations of this chapter, \vec{E} corresponds to a single variable E while in the n^{th} multicanonical iteration we employ

$$F(E) = \begin{cases} \exp\left[\frac{1}{2\sigma^2}(E - E_n^{\text{L}})^2\right], & E < E_n^{\text{L}} \\ \exp\left[\frac{1}{2\sigma^2}(E - E_n^{\text{R}})^2\right], & E > E_n^{\text{R}} \\ 1, & E_n^{\text{L}} \leq E \leq E_n^{\text{R}} \end{cases} \quad (7.1)$$

in which E_n^{L} and E_n^{R} are the left and right limits of the region over which $F(E)$ is constant.

If the relative normalization is known, we can combine the estimates, $\hat{p}_i(\vec{E})$, $i = 1, \dots, (m - 1)$, of $p(\vec{E})$, obtained from either $m - 1$ successive iterations or the converged results of $m - 1$ biased multicanonical calculations with different bias functions, $F_i(\vec{E})$, for which the corresponding regions \mathcal{R}_i differ. For example, according to the standard importance sampling prescription, the result in the k^{th} histogram bin is obtained by weighing each contribution according to the relative number of samples, $H_i(\vec{E}_k)$, in this bin; that

is [107, 117]

$$p(\vec{E}_k) = \frac{\sum_{i=1}^{m-1} H_i(\vec{E}_k) \hat{p}_i(\vec{E}_k)}{\sum_{i=1}^{m-1} H_i(\vec{E}_k)}. \quad (7.2)$$

Unfortunately, the relative normalization of the individual histograms is generally undetermined unless the biasing is sufficiently weak that all calculations yield acceptable estimates of the large-amplitude regions of the pdf. However, if the regions \mathcal{R}_i and \mathcal{R}_{i-1} partially overlap, the respective pdfs \hat{p}_i and \hat{p}_{i-1} , differ in the common region by at most a normalization constant if statistical fluctuations are neglected. Below, we present a slight modification of the procedure of Refs. [33, 61, 101]. The resulting algorithm determines the ratio of the unknown normalization constants by minimizing the mean squared error (MSE) of the pdf estimate.

We accordingly consider a slightly more general expression,

$$\hat{p}(\vec{E}_k) = \sum_{i=1}^{m-1} A_i w_i(\vec{E}_k) \hat{p}_i(\vec{E}_k) \quad (7.3)$$

for the pdf estimate at the start of the m^{th} multicanonical iteration, formed by the weighted sum

$$\sum_{i=1}^{m-1} w_i(\vec{E}_k) = 1 \quad (7.4)$$

of $m - 1$ biased estimates, $\hat{p}_i(\vec{E}_k)$, for all k . The normalization constant, A_i , is defined by

$$p_i(\vec{E}) = \begin{cases} p(\vec{E})/A_i, & \vec{E} \in \mathcal{R}_i \\ 0, & \text{otherwise} \end{cases} \quad (7.5)$$

in which $p(\vec{E})$ and $p_i(\vec{E})$ are the exact and the biased pdf:s, respectively. That is, A_i denotes the integral of $p(\vec{E})$ over the i^{th} biasing region, \mathcal{R}_i . The optimal weight functions

can therefore be determined by minimizing the MSE [96] of $\hat{p}(\vec{E}_k)$,

$$\begin{aligned} \text{MSE}\{\hat{p}(\vec{E}_k)\} &\equiv \text{E}\{[\hat{p}(\vec{E}_k) - p(\vec{E}_k)]^2\} \\ &= \sum_{i=1}^{m-1} A_i^2 w_i^2(\vec{E}_k) \text{Var}\{\hat{p}_i(\vec{E}_k)\} \\ &\quad + [\text{E}\{\hat{p}(\vec{E}_k)\} - p(\vec{E}_k)]^2, \end{aligned} \quad (7.6)$$

with respect to $w_i(\vec{E}_k)$ while simultaneously imposing the normalization condition of Eq. (7.4) [2, 61, 101, 107]. Here $\text{E}\{\dots\}$ and $\text{Var}\{\dots\}$ represent the expectation and variance over a large number of multicanonical simulations.

To minimize Eq. (7.6), the constraint of Eq. (7.4) is introduced through a Lagrange multiplier, λ , and each partial derivative of

$$\hat{f}(\vec{E}_k) = \text{MSE}\{\hat{p}(\vec{E}_k)\} - \lambda \left\{ \sum_{i=1}^{m-1} w_i(\vec{E}_k) - 1 \right\}, \quad (7.7)$$

with respect to w_j is subsequently equated to zero. This yields for the optimum weight functions

$$w_i(\vec{E}_k) = \frac{I_i^{(k)} / [A_i^2 \text{Var}\{\hat{p}_i(\vec{E}_k)\}]}{\sum_{j=1}^{m-1} I_j^{(k)} / [A_j^2 \text{Var}\{\hat{p}_j(\vec{E}_k)\}]} \quad (7.8)$$

The expression $I_i^{(k)}$ is zero if $\hat{p}_i(\vec{E}_k) = 0$ and one otherwise, which ensures that the pdf estimate is correctly weighted in the absence of histogram samples.

The variance appearing in Eq. (7.8) is given in terms of the n_i samples recorded in H_i by [34, 96, 101]

$$\text{Var}\{\hat{p}_i(\vec{E}_k)\} = \frac{1}{n_i \Delta V} g_i(\vec{E}_k) p_i(\vec{E}_k). \quad (7.9)$$

Here $g_i(\vec{E}_k)$ should incorporate the covariance between successive samples in the k^{th} histogram bin during the i^{th} multicanonical iteration [101]. For simplicity, however, we approximate $g_i(\vec{E}_k)$ by a constant for all i [101], and further replace $p_i(\vec{E}_k) \rightarrow p(\vec{E}_k)/A_i$

according to Eq. (7.5). This yields the approximation

$$\hat{p}(\vec{E}_k) \approx \frac{\sum_{i=1}^{m-1} n_i \hat{p}_i(\vec{E}_k)}{\sum_{j=1}^{m-1} I_j^{(k)} n_j / A_j}. \quad (7.10)$$

We proceed as in Refs. [61, 101] by observing that

$$A_n = \int_{\mathcal{R}_n} p(\vec{E}) d\vec{E} \approx \sum_{\{k_n | \vec{E}_{k_n} \in \mathcal{R}_n\}} \hat{p}(\vec{E}_{k_n}) \Delta V. \quad (7.11)$$

Applying Eq. (7.10) then yields

$$A_n \propto \sum_{\{k_n | \vec{E}_{k_n} \in \mathcal{R}_n\}} \frac{\sum_{i=1}^{m-1} n_i \hat{p}_i(\vec{E}_{k_n})}{\sum_{j=1}^{m-1} I_j^{(k_n)} n_j / A_j}. \quad (7.12)$$

To solve the above expression, all normalization constants A_n are initialized to unity and then inserted into Eq. (7.12), which provides new estimates of the A_n . The procedure is then iterated until the relative change in A_n is less than a specified value, c.f. Appendix C.6. To avoid numerical divergences, A_n is typically replaced with cA_n after each iteration. In our calculations, the constant c is determined from the condition that the components of the pdf \hat{p} , Eq. (7.10), sum to unity.

If the intersection of \mathcal{R}_{n-1} and \mathcal{R}_n is disjoint from any other biasing region for every n , the converged A_n/A_{n-1} equals the ratio of the sums of the histogram values within the overlapping region. If, on the other hand, more than one successive biasing region overlaps, the full iterative solution of Eq. (7.12) is required.

The number of iterations can be considerably reduced if the initial estimates for A_n in Eq. (7.12) are chosen to minimize the vertical offset between \hat{p}_{n-1} and \hat{p}_n . That is, if $A_n^{(0)}$

represents the initial value of A_n , we minimize the weighted least-squares error [5] with weighting constants w_k^{ls} , namely,

$$\sum_{k=1}^N w_k^{ls} [\hat{p}_{n-1}(\vec{E}_k) - x_n \hat{p}_n(\vec{E}_k)]^2, \quad (7.13)$$

with respect to $x_n \equiv A_n^{(0)}/A_{n-1}^{(0)}$, $n = 2, \dots, (m-1)$. This yields

$$\frac{A_n^{(0)}}{A_{n-1}^{(0)}} = \frac{\sum_{k=1}^N w_k^{ls} \hat{p}_{n-1}(\vec{E}_k) \hat{p}_n(\vec{E}_k)}{\sum_{k=1}^N w_k^{ls} [\hat{p}_n(\vec{E}_k)]^2}. \quad (7.14)$$

Similarly, $A_n^{(0)}$ can be recursively determined from Eq. (7.14) after setting $A_1^{(0)} = 1$. The choice of weights, w_k^{ls} , is not unique; in our calculations, however, we set w_k^{ls} to the product $(H_n(\vec{E}_k) - 1)(H_{n-1}(\vec{E}_k) - 1)$ (recall that all histogram bins are initialized to unity).

Accordingly, to determine the pdf of a observable, E , we first compute a Monte Carlo estimate of the pdf, \hat{p}_1 , and select as the initial biasing region an interval centered on the last point, E_0 , in the calculation for which ten or more samples have been recorded. This criterion reflects the empirical observation that with our computational parameters, these points are normally surrounded by a region of histogram bins that contain at least one sample. The width of the biasing range, ΔE , is then selected for every iteration according to

$$\Delta E = 2 \frac{E_{\max} - E_0}{1 + 2\gamma(N_I - 1)}. \quad (7.15)$$

where E_{\max} and N_I denote the maximum value of interest of the system observable and the number of biased multicanonical iterations, respectively and γ is the fraction of the window width by which the biasing range is shifted to the right for each multicanonical iteration, i.e.

$$\begin{aligned} E_n^L &= E_0 + \Delta E [(n-1)\gamma - 1/2] \\ E_n^R &= E_0 + \Delta E [(n-1)\gamma + 1/2] \end{aligned} \quad (7.16)$$

for $n = 1, \dots, N_I$. That is, $\gamma = 0.75$ corresponds to a 25% overlap of the new biasing region with the previous region. Employing N_s samples and $F(E)$ given by Eq. (7.1), a biased multicanonical histogram, H_2 is calculated as in Sec. 7.2 leading to a new pdf estimate \hat{p}_2 . The ratio of the normalization constants A_1 and A_2 for this first iteration is then determined through Eq. (7.12), after which the updated pdf is computed from Eq. (7.10). The biasing region is then shifted to the right according to Eq. (7.16) and the procedure repeated such that the unnormalized pdf estimates $\hat{p}_1, \dots, \hat{p}_{m-1}$, after $m - 1$ biased multicanonical iterations are combined in an analogous fashion to yield the improved pdf estimate for the m^{th} iteration.

Note that our method differs substantially from, for example, umbrella sampling, in which intermediate results are only combined at the end of the calculation. Further, as noted in [67], our approach enables regions of state space with physically interesting properties to be located and examined dynamically by positioning the biasing region according to the results of previous iterations.

7.4 Numerical results

Before establishing the experimental significance of our formalism, we first demonstrate its accuracy through numerical simulations. These will analyze the established test case for which the system output variable $E \equiv \tau(\vec{\alpha})$ is the DGD, τ , of a simulated optical fiber emulator. In the calculations of this chapter, the emulator consists of $N_{sec} = 100$ sections of randomly oriented, polarization maintaining (PM) fiber [26]. The DGD of a single fiber section, τ_{sec} , is determined by the requirement that the average DGD $\tau_{avg} = \tau_{sec} \sqrt{8N_{sec}/3\pi} = 25$ ps. The system parameters, $\vec{\alpha}$, then correspond to the coupling angles between the PM fiber segments [112]. Our calculations further employ ten 2×10^4 -sample iterations with $\gamma = 0.3$. Events are recorded in 100 equal size histogram bins in the interval $0 < \tau/\tau_{avg} < 10$.

We first display the rate of convergence afforded by Eq. (7.12) in Fig. 7.1, which displays

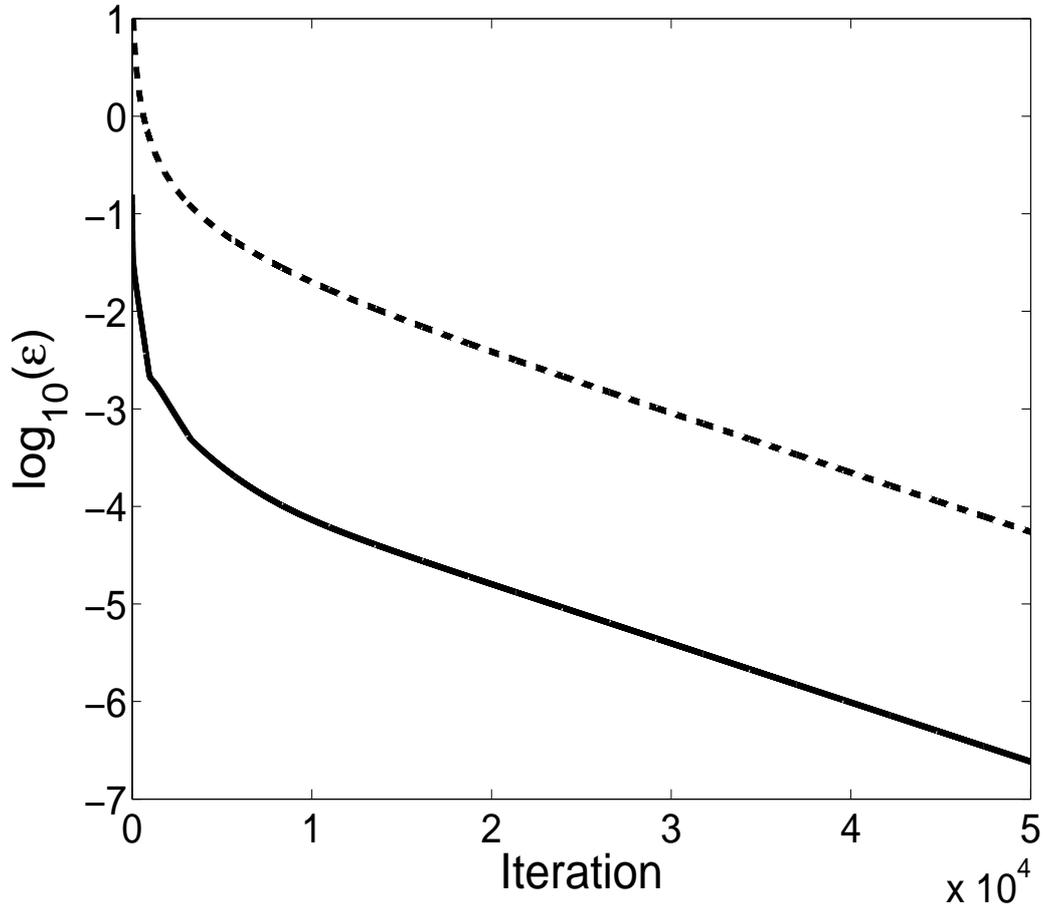


Figure 7.1: The maximum relative error, Eq. (7.17), for the initial estimate of the normalization constants of Eq. (7.14) (solid line) and for the iterated results $A_n^{(0)} = 1$, $n = 1, \dots, N_I$ for these constants (dotted line). Results are illustrated for a 100 segment fiber emulator with $\tau_{avg} = 25$ ps and 10 biased multicanonical iterations with 2×10^4 points per iteration with $\gamma = 0.3$.

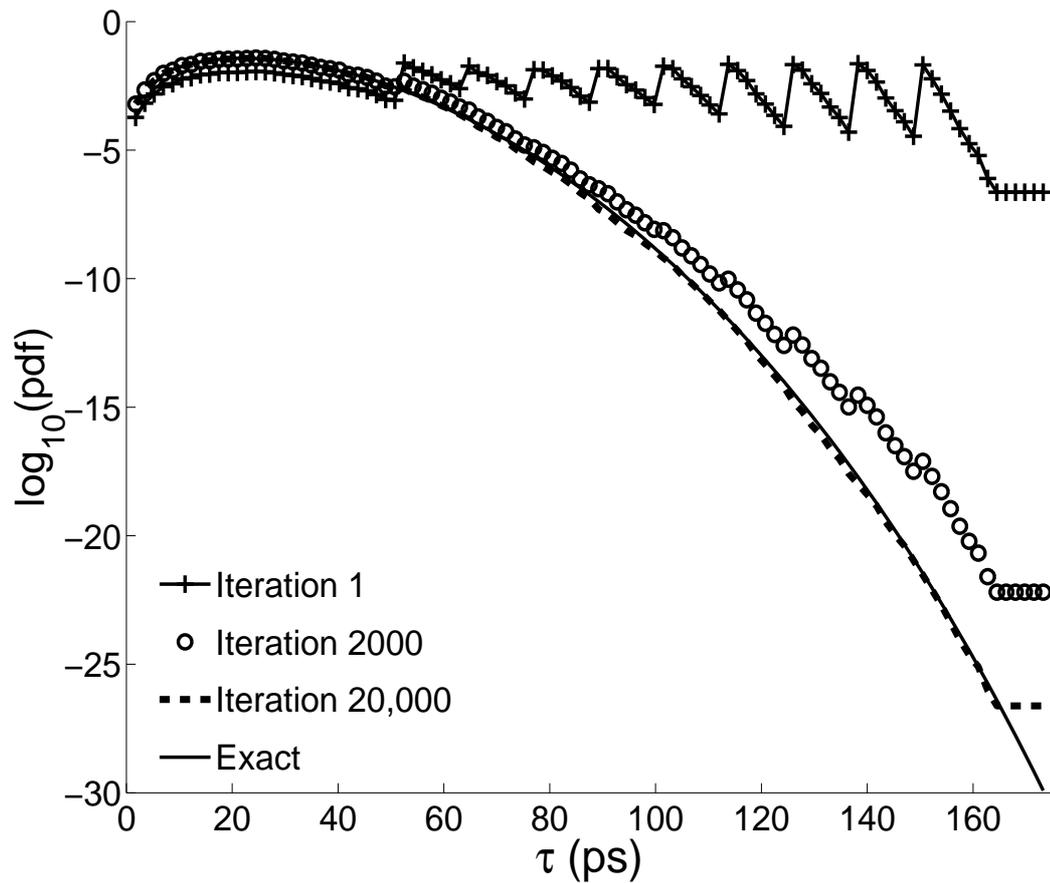


Figure 7.2: The pdf, Eq. (7.10), obtained from the calculation of Fig. 7.1 after joining the results of 10 different biased calculations after 1 (crosses), 2000 (circles) and 20,000 (dotted-line) iterations of Eq. (7.12), The solid line is the exact result of Eq. (1.17) [54]. The normalization constants are here initialized to $A_n^{(0)} = 1$, $n = 1, \dots, N_I$.

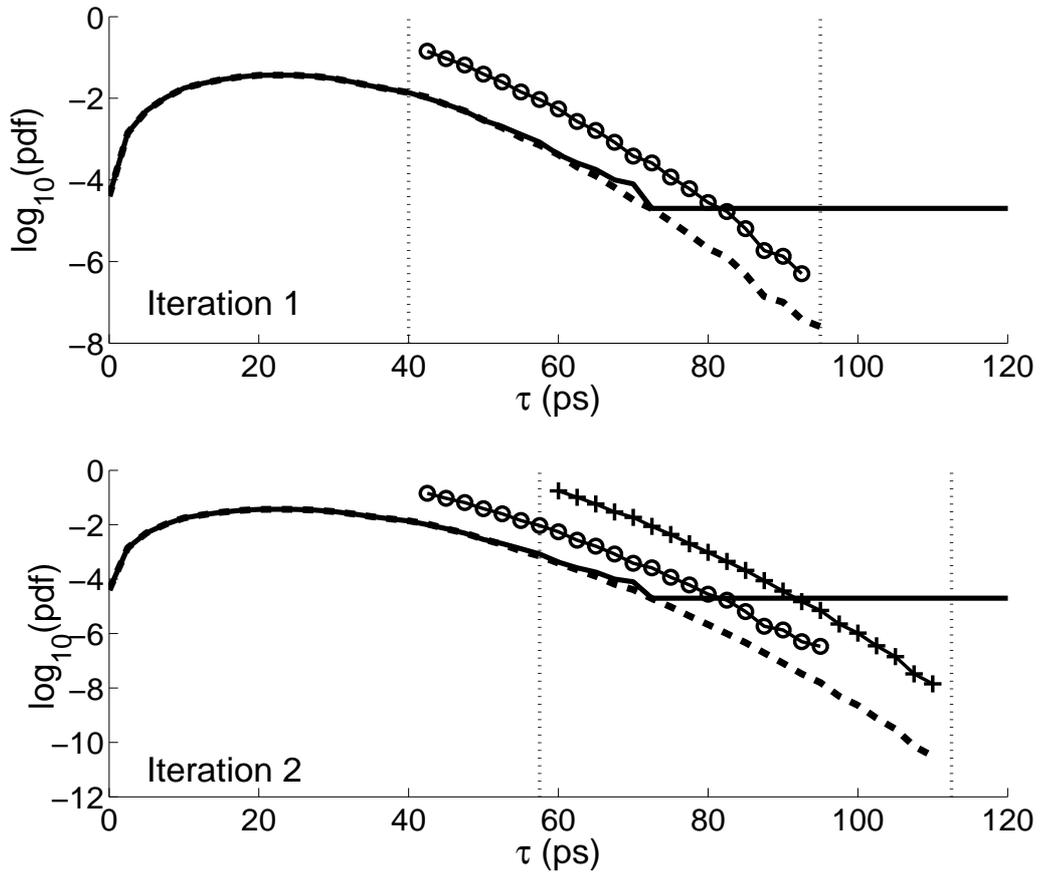


Figure 7.3: (a) The biased pdf estimate \hat{p}_2 before (circles) and after (dashed line) combination with the initial Monte Carlo estimate \hat{p}_1 (solid line) according to Eq. (7.10) with $\gamma = 0.3$. (b) Analogous results for the second biasing iteration in which \hat{p}_1 , \hat{p}_2 and \hat{p}_3 (crosses) are combined according to Eq. (7.10) (solid line). The vertical lines indicate the biasing region.

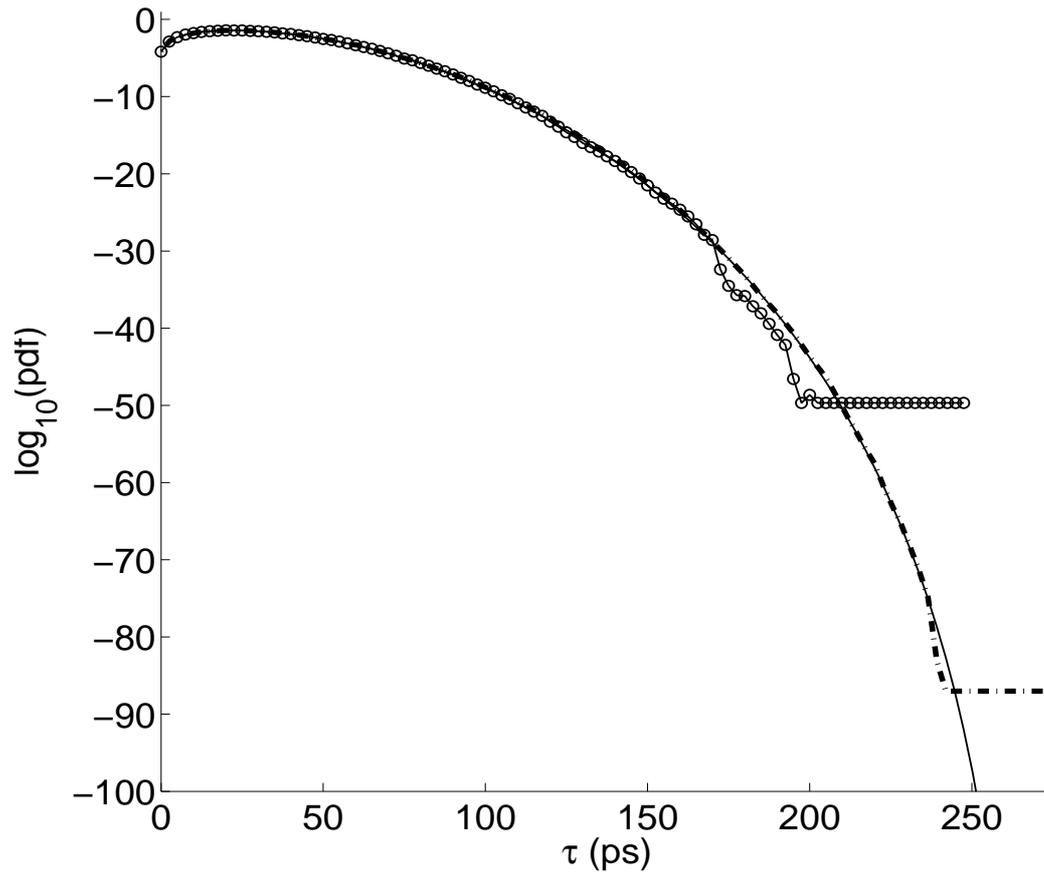


Figure 7.4: The pdf of the DGD of a 100 section fiber emulator calculated with fifty 5×10^4 -sample iterations with the iterative biased multicanonical method (dashed-dotted line) and the standard multicanonical procedure (\circ) together with the analytic result (solid line).

the maximum relative error

$$\epsilon = \max_n \left[\frac{A_n^{(i)} - A_n^{(i-1)}}{A_n^{(i-1)}} \right] \times 100\% \quad (7.17)$$

of the normalization constants, after the i^{th} iteration, $i = 2, 3, \dots, 50000$. The initial estimates $A_n^{(0)} = 1$ yield the dotted line in Fig. 7.1, while the initial estimates of Eq. (7.14) instead yield the solid line in the figure. The combined pdf estimates obtained from Eq. (7.10) after 1 (+), 2000 (\circ) and 20,000 (dotted line) iterations of Eq. (7.12) for the case in which all A_n were initialized to unity are presented in Fig. 7.2.

Next, in Fig. 7.3 we illustrate the combination of piecewise biased estimates of the pdf. The solid curve in the top graph of Fig. 7.3 is the initial pdf estimate, \hat{p}_1 obtained from a Monte Carlo simulation, while the subsequent biased estimates, \hat{p}_2 , before and after the application of Eq. (7.12) are indicated by circular markers and dashed lines, respectively. The quantity \hat{p}_2 is displayed only within region of constant bias, as the histogram values outside this region does not affect the updated pdf estimate in our procedure. The second graph in Fig. 7.3 depicts the corresponding results for the second iteration for which the dotted line is now generated by combining the three results \hat{p}_1 , \hat{p}_2 and \hat{p}_3 according to Eq. (7.10).

After fifty 5×10^4 -sample iterations with $\gamma = 0.25$, we generate the dashed-dotted line in Fig. 7.4 with our biased multicanonical method, while the solid line in the figure is the analytic result, Eq. (1.17) [54]. The result of the standard multicanonical procedure with fifteen 1.67×10^5 -sample iterations, which was found empirically to yield near ideal results for the pdf in the low-probability tail region, is shown with \circ markers in Fig. 7.4. Evidently, even without careful optimization of the computational parameters, our biased multicanonical method yields a considerable increase in computational efficiency compared to the standard multicanonical procedure with the same number of samples.

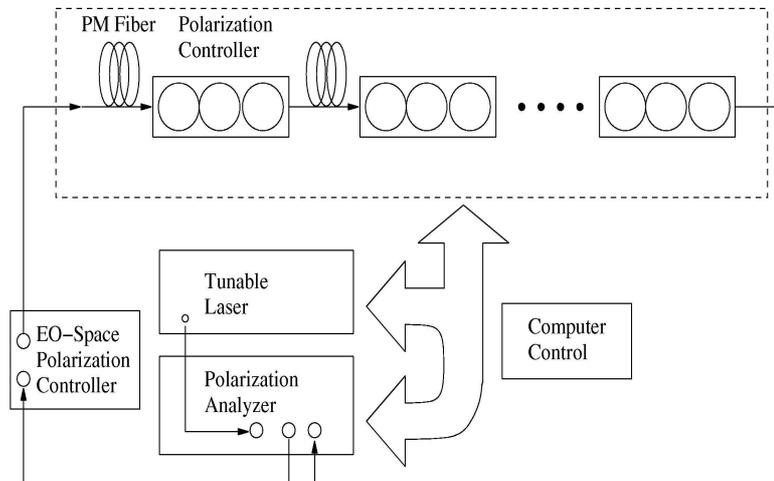


Figure 7.5: The biased multicanonical experimental setup.

7.5 Experimental measurements

To demonstrate that our procedure can be applied experimentally, we measured the pdf of the DGD, τ , of an 8 stage optical fiber emulator following the procedure of Ref. [69]. The experimental setup, c.f. Fig. 7.5, employed eight General Photonics PolarRITE II three axis polarization controllers separated by unequal length PM fibers. As in Sec. 7.4, $E \equiv \tau(\vec{\alpha})$ represents the relevant system observable, namely the DGD, while $\vec{\alpha}$, corresponds to the 24 input voltages applied to the 8 polarization controllers. The DGD of the fiber emulator was measured by applying the Jones matrix eigenanalysis (JME) procedure to the output of an HP 8509B polarization analyzer combined with a Tunicus PRI tunable laser [47]. The mean and maximum DGD for the system are approximately 11.7 ps and 37.4 ps, respectively.

In our experiment, the Jones matrix was measured at two wavelengths displaced by 0.1nm by cycling an EO-Space 8-stage polarization controller through 5 random input polarization states and reading the corresponding output Stokes vectors from the analog outputs of the HP 8509B with a National Instruments PCI-6221 input card. The Jones matrix was computed from the output values according to the least-squares method of Refs. [86] and [94], at a rate five times faster than the native HP 8509B JME measure-

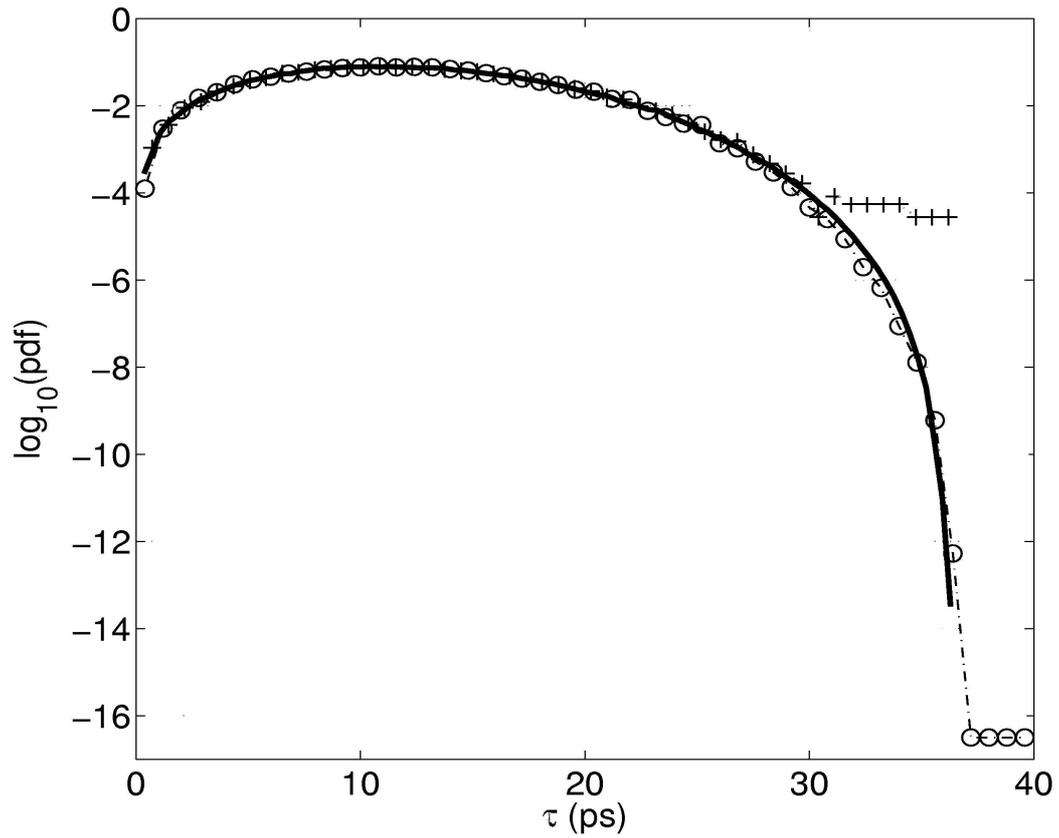


Figure 7.6: The experimentally determined pdf of the DGD of an 8 section fiber emulator for 45,000 samples measured with the piecewise biased multicanonical method (circles) and the standard Monte Carlo procedure (crosses). The solid line indicates the corresponding numerical results for three 5×10^5 sample iterations in the standard multicanonical method.

ment algorithms, although further improvements could be attained with relatively minor component modifications.

Employing 10,000 Monte Carlo measurements followed by seven, 5,000-sample biased multicanonical iterations with $\gamma = 0.5$, the piecewise procedure of Sec. 7.3 yielded the \circ markers in Fig. 7.6 for the measured pdf of the DGD, τ , of the fiber emulator. For comparison, the points marked $+$ in the figure were generated with a standard 45,000-point Monte Carlo measurement. Finally, the solid line of Fig. 7.6 indicates the simulated pdf for three 5×10^5 -sample iterations of the standard multicanonical algorithm in which the DGD for each PM fiber segment corresponded to the values measured in the experimental setup. Although the number of multicanonical iterations and sampling events were chosen to produce probabilities of occurrence comparable to the experimental results, the agreement between the calculated and experimentally determined pdfs is still notable [26, 37]. Evidently, our aligned piecewise biasing procedure computes the low probability regions of the pdf with considerably greater efficiency than the standard multicanonical method.

7.6 Conclusions

We have employed an iterative method to combine the intermediate results of strongly biased multicanonical calculations. Our procedure can be immediately adapted to any biasing formalism, as for example, directly calculating other quantities than the pdf [114], raising the intermediate pdf or histogram variable to a power (power method), dynamically biasing in multidimensional space to locate regions with physically interesting properties [67] and employing highly confining bias functions to strictly limit the statistical events to a small region of parameter space (barrier method) [117]. Since in each case, the sampling region is limited to a far smaller region of parameter space than in standard multicanonical calculations, considerable increases in computational efficiency are possible, especially in high-dimensional problems, which is of great significance in experimental applications.

Chapter 8

Multicanonical Analyses of System Penalties

We finally employ biased multicanonical sampling to generate system configurations with specific probabilities of eye-closure either before or after compensation (equalization). In an optical communications context, such states can be analyzed by their PMD coefficients. We find that the PMD coefficients of states with equal eye-closure probabilities vary significantly for different compensator structures, an observation which has implications for system characterization.

8.1 Introduction

In communications theory, the properties of system configurations with specified eye-closure penalties must often be examined. However, the probability of such configurations can be very low; further, the penalty may depend on numerous system parameters. For our model 10 Gb/s optical communication system application these include the modulation format, mean differential group delay (DGD), and presence of optical or electrical PMD compensation [6, 19, 98]. In this chapter, we however demonstrate that the biased

multicanonical method [83,117] provides an efficient means of sampling states with a given penalty metric that can then be characterized in terms of a few significant global system variables, here the lowest-order PMD coefficients.

The organization of this chapter is as follows. First, we review our formulation of PMD compensation in terms of products of non-commuting exponential operators in the special case of zero polarization-dependent loss (PDL). We also introduce an eye-closure penalty metric for a simplified optical system model. Biased multicanonical sampling is then employed to sample emulator configurations with approximately equal eye-closure penalties at a low probability of occurrence. The density of these configurations is analyzed in terms of the first, second and third order PMD coefficients after optical compensation.

8.2 Compensator Jones matrix

The frequency-dependence of the Jones matrix, $\mathbf{T}(\omega)$, of an optical fiber with zero PDL is described by the differential equation $[\mathrm{d}\mathbf{T}(\omega)/\mathrm{d}\omega] \mathbf{T}(\omega)^{-1} = -i\vec{\Omega}(\omega) \cdot \vec{\sigma}/2$ in which $\vec{\Omega}(\omega)$ is a real three-component vector and ω denotes the optical frequency [59]. A convenient form of the solution to the above equation in terms of Taylor orders of $\vec{\Omega}(\omega)$ is given by the Magnus expansion, which to third order in the frequency deviation, $\Delta\omega = \omega - \omega_0$, from the optical carrier frequency, ω_0 , yields

$$\begin{aligned} \mathbf{T}(\omega)\mathbf{T}(\omega_0)^{-1} = & \exp\left[-\frac{i}{2}\left(\vec{\Omega}_0\Delta\omega + \vec{\Omega}_1\frac{\Delta\omega^2}{2!} \right. \right. \\ & \left. \left. + \left(\vec{\Omega}_2 - \frac{1}{2}\vec{\Omega}_0 \times \vec{\Omega}_1\right)\frac{\Delta\omega^3}{3!} + \dots\right) \cdot \vec{\sigma}\right]. \end{aligned} \quad (8.1)$$

in which $\vec{\Omega}_n \equiv (\mathrm{d}^n\vec{\Omega}/\mathrm{d}\omega^n)|_{\omega_0}$ [88]. Truncating the series in the exponential yields an expression that can be expressed as a product of non-commuting exponential operators, although the exponentials must be properly ordered to preserve the level of accuracy in $\Delta\omega$ [118]. The inverse of each of these exponential operators can, at least in principle, be implemented experimentally as a particular compensator element. In this manner, we

obtain expressions for the Jones matrices of the form $\mathbf{T}_C^{(n)}(\omega) = \mathbf{C}^{(n)}(\Delta\omega)\mathbf{T}(\omega)$ that yield a minimum residual polarization dependence after n^{th} order PMD compensation. To first, second and third order, respectively, these are [60, 88, 118]

$$\mathbf{C}^{(1)}(\Delta\omega) = \exp\left[\frac{i\Delta\omega}{2}\vec{\Omega}_0 \cdot \vec{\sigma}\right], \quad (8.2)$$

$$\begin{aligned} \mathbf{C}^{(2)}(\Delta\omega) = & \exp\left[\frac{i\Delta\omega}{4}\vec{\Omega}_0 \cdot \vec{\sigma}\right] \exp\left[\frac{i\Delta\omega^2}{4}\vec{\Omega}_1 \cdot \vec{\sigma}\right] \\ & \exp\left[\frac{i\Delta\omega}{4}\vec{\Omega}_0 \cdot \vec{\sigma}\right], \end{aligned} \quad (8.3)$$

$$\begin{aligned} \mathbf{C}^{(3)}(\Delta\omega) = & \exp\left[\frac{i\Delta\omega^3}{12}\vec{\Omega}_2 \cdot \vec{\sigma}\right] \exp\left[\frac{i\Delta\omega}{3}\vec{\Omega}_0 \cdot \vec{\sigma}\right] \\ & \exp\left[\frac{i\Delta\omega^2}{4}\vec{\Omega}_1 \cdot \vec{\sigma}\right] \exp\left[\frac{i\Delta\omega}{6}\vec{\Omega}_0 \cdot \vec{\sigma}\right] \end{aligned} \quad (8.4)$$

Unfortunately, however, because the PMD coefficients are obtained through a Taylor expansion the third order compensator of Eq. (8.4) yields improved performance compared to Eq. (8.3), for $\Delta\omega$ smaller than a certain correlation frequency, while for larger frequency deviations the error in the third-order result tends to increase rapidly, degrading the eye-closure penalty [112].

8.3 Numerical results

We now calculate the joint conditional probability $P(|\vec{\Omega}_0|, |\vec{\Omega}_1|, |\vec{\Omega}_2| | \Delta q_{\text{wc}})$, that the communication system will display certain values of the first-, second- and third-order PMD vector magnitudes. However, we restrict our attention to states with a given ‘‘worst-case’’ value of the eye-closure penalty, Δq_{wc} ; that is, system configurations for which the probability density of an eye-closure is 10^{-5} and, in particular, demonstrate that this quantity can be efficiently determined with biased multicanonical sampling [83, 117].

We accordingly model $\mathbf{T}(\omega)$ of an optical communication system with a mean DGD, τ_{avg} , as a concatenation of $N_{\text{sec}} = 100$ polarization-maintaining fiber sections, each of which possesses a DGD $\tau_{\text{sec}} = \tau_{\text{avg}}\sqrt{3\pi/8N_{\text{sec}}}$. Then, for each emulator realization, the

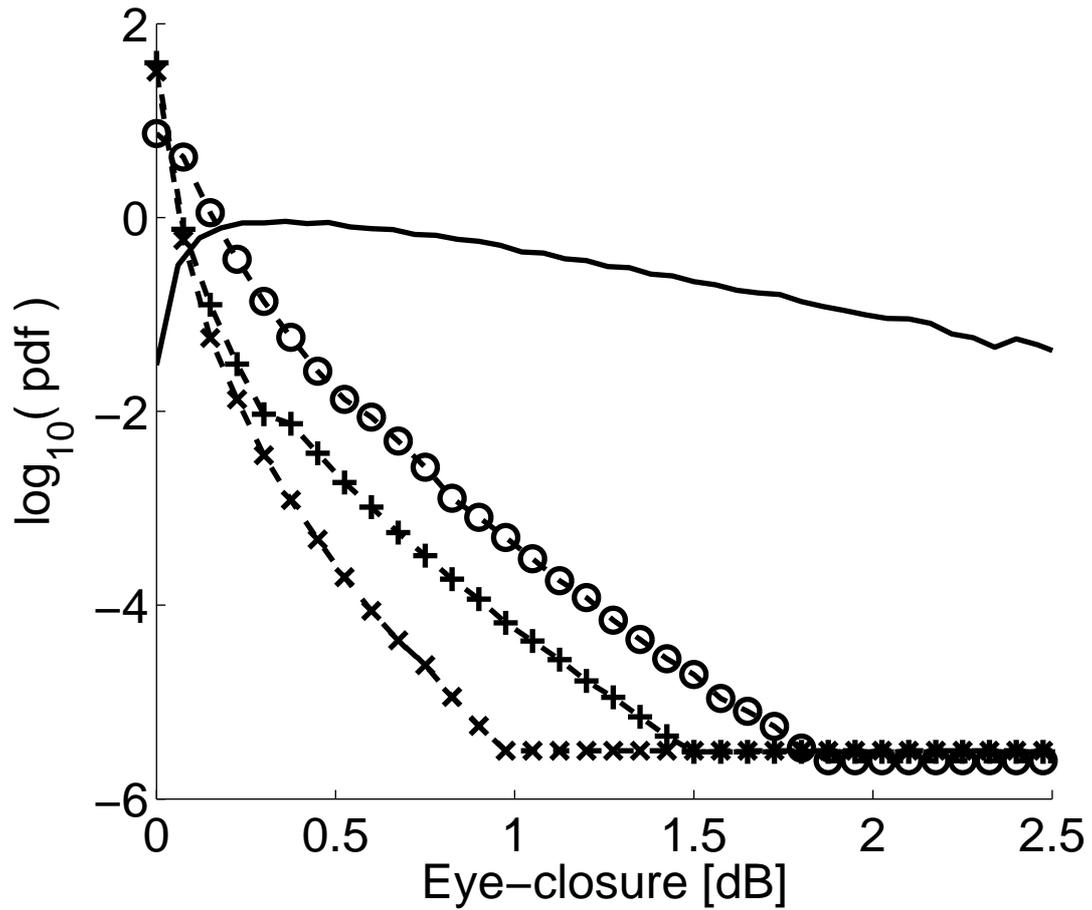


Figure 8.1: Probability density function of the eye-closure penalty for $\tau_{avg} = 30$ ps without compensation (solid line), and after the compensators described by the Jones matrices of Eq. (8.2) (\circ), Eq. (8.3) (\times), and Eq. (8.4) ($+$).

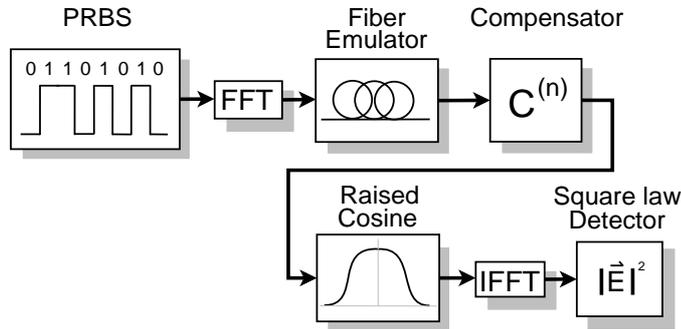


Figure 8.2: Simulation configuration.

propagation through the fiber of a 10 Gb/s non-return to zero (NRZ) x -polarized optical waveform, obtained by sampling a 32 bit pseudo-random bit sequence (PRBS) at $T/16$, is simulated. Optical PMD compensation is then included through the Jones matrices of Eqs. (8.2)-(8.4) in which the $\vec{\Omega}_n$ are obtained by finite-differencing the numerically determined $\mathbf{T}(\omega)$. The resulting output pulse was filtered with a raised-cosine filter with a roll-off factor of 1.0 and detected with a simulated square-law detector as depicted in Fig. 8.2.

The eye-closure penalty, $\Delta q = \log_{10}(q_{\text{b2b}}/q_{\text{out}})$, where q_{out} and q_{b2b} respectively denote the minimum eye opening at the output of the emulator and for a back-to-back transmitter/receiver pair, was then estimated for each fiber realization with an inner-eye measure algorithm based on the minimum vertical eye opening within a 20 ps jitter window centered on the optimum sampling phase [95, 98]. This window interval approximately reproduces the penalty resulting from timing phase jitter in the clock recovery circuit.

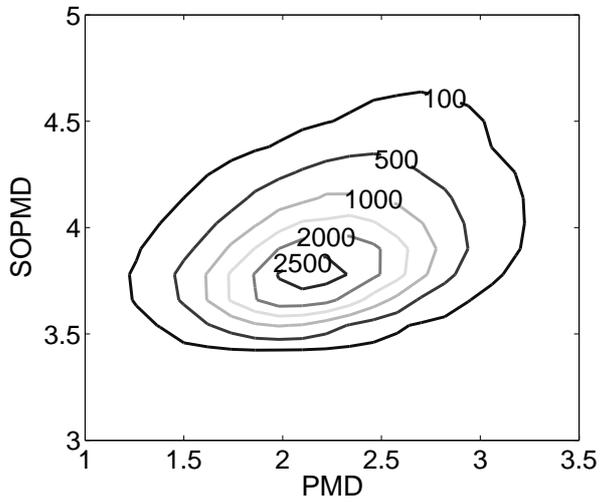
To simulate the pdf of the eye-closure penalty we adapted the biased multicanonical technique of Refs. [83, 117]. Here we first subdivided the $[0, 3]$ dB range of eye-closure penalty into 10 overlapping regions, $\mathcal{R}_1, \mathcal{R}_2 \dots \mathcal{R}_{10}$, such that adjacent biasing regions overlap by 30%. We denote the left and right boundaries of the m^{th} region by Δq_m^{L} and Δq_m^{R} . In the m^{th} multicanonical iteration ($m = 1, \dots, 10$), an auxiliary bias function is applied that is constant within \mathcal{R}_m and increases exponentially away from the region

boundaries [83, 117]. While the resulting modified multicanonical transition rule [16] substantially augments the probability of transitions into the biasing region $\Delta q_m^L \leq \Delta q \leq \Delta q_m^R$, an unbiased estimate of the pdf is subsequently recovered by multiplying the previous pdf estimate by the histogram of visited states [117]. However since the relative normalization of the results obtained in different biasing regions is undetermined, we determine the full pdf by iteratively minimizing the mean squared variance of a linear combination of the individual distributions [83].

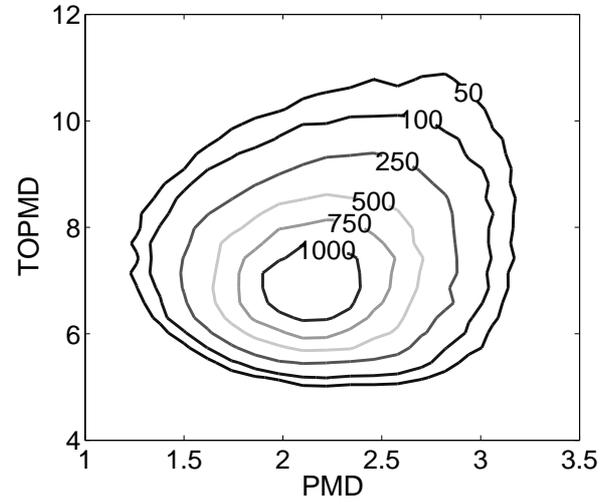
Accordingly, we first generated 5×10^4 fiber realizations biased within \mathcal{R}_m for each biased multicanonical iteration, m . Once an \mathcal{R}_m was obtained such that included events contribute to the eye closure penalty with a probability density after compensation of 10^{-5} , the bias function was narrowed to concentrate the final set of 10^5 samples within the restricted region $\Delta q \in \Delta q_{wc} \pm 0.002$ dB.

In Fig. 8.1 we display the pdf of the eye-closure penalty for an optical fiber characterized by $\tau_{avg} = 30$ ps, followed by PMD compensation corresponding to Eq. (8.2) (\circ markers), Eq. (8.3) (\times markers), Eq. (8.4) ($+$ markers) as well as without PMD compensation (solid line). That third order compensation, Eq. (8.4), yields a greater eye-closure penalty than second order compensation, Eq. (8.3), at low-probabilities of occurrence indicates that these events are associated with large PMD values for which the third order compensator considerably augments fourth and higher order effects that dominate the eye-closure penalty at large frequency offsets.

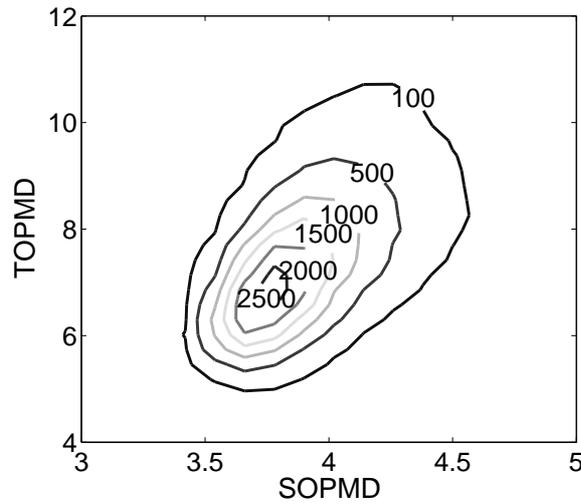
Next, Figs. 8.3-8.5 displays histograms for the joint marginal distribution, $P(|\vec{\Omega}_0|, |\vec{\Omega}_1|, |\vec{\Omega}_2| | \Delta q_{wc})$, obtained from the iso-penalty ensemble of emulator states for each of the compensators of (Fig. 8.3) Eq. (8.2), (Fig. 8.4) Eq. (8.3) and (Fig. 8.5) Eq. (8.4). Here we normalize the first- (PMD), second- (SOPMD) and third-order (TOPMD) PMD with respect to τ_{avg} and employ 50^3 bins within the region bounded by the planes in Figs. 8.3-8.5. The subplots of this figure display the projections of the three dimensional histogram onto the PMD/SOPMD, PMD/TOPMD and SOPMD/TOPMD surfaces. Diagrams analogous to Figs. 8.3-8.5 for $\Omega_{n,x}$, $\Omega_{n,y}$, and $\Omega_{n,z}$, the x , y and z components of $\vec{\Omega}_n$, displayed in the



(a) PMD and SOPMD.

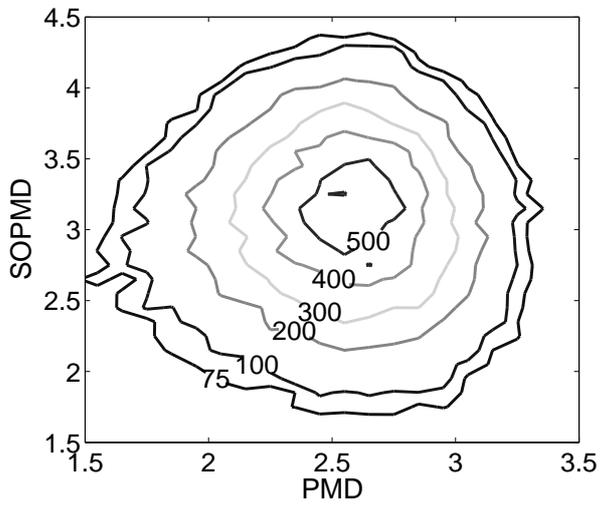


(b) PMD and TOPMD.

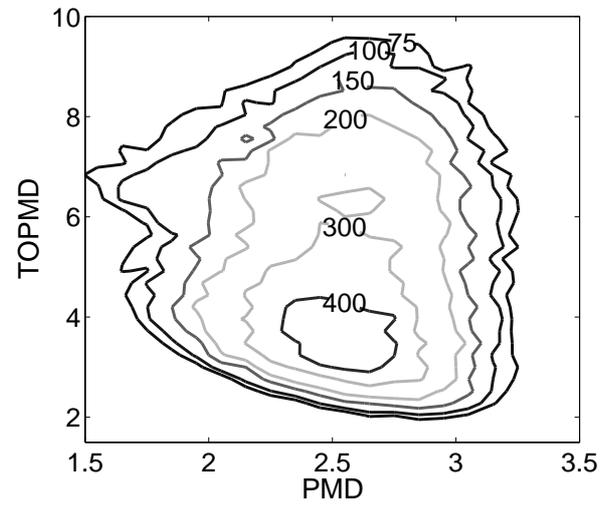


(c) SOPMD and TOPMD.

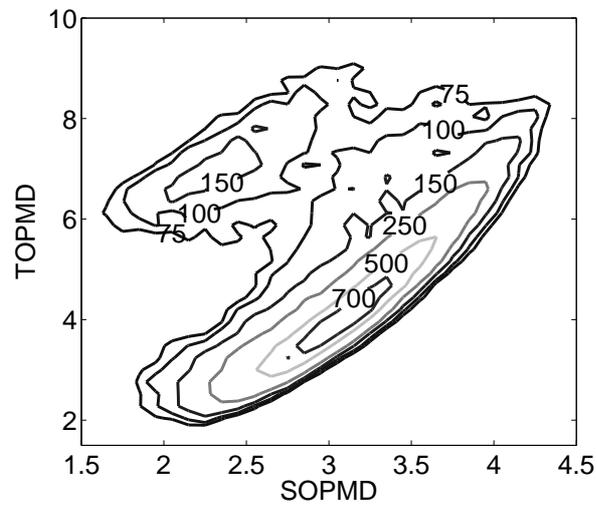
Figure 8.3: Projections of the histogram of the states of our system model that contribute to the eye-closure penalty with a 10^{-5} probability density after the compensation of Eq. (8.2) onto normalized (a) PMD/SOPMD, (b) PMD/TOPMD and (c) SOPMD/TOPMD planes.



(a) PMD and SOPMD.

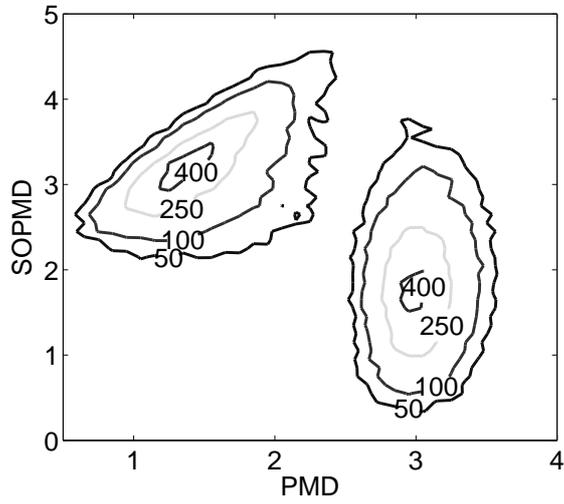


(b) PMD and TOPMD.

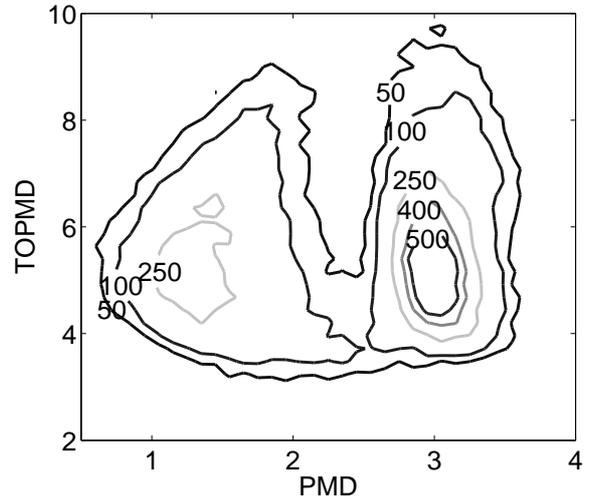


(c) SOPMD and TOPMD.

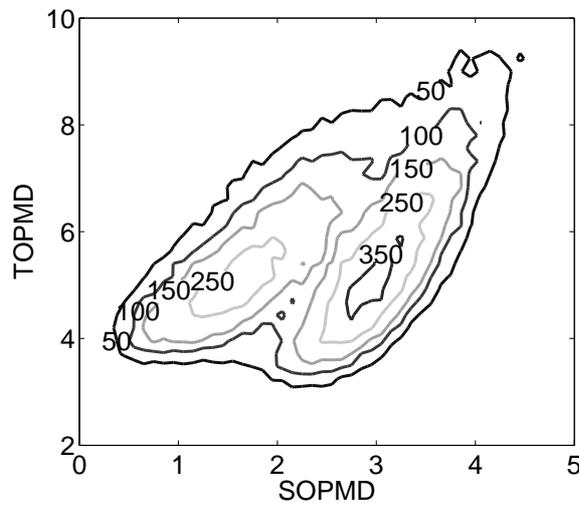
Figure 8.4: Same as Fig. 8.4 except for the compensator Jones matrix of Eq. (8.3).



(a) PMD and SOPMD.



(b) PMD and TOPMD.



(c) SOPMD and TOPMD.

Figure 8.5: Same as Fig. 8.4 except for the compensator Jones matrix of Eq. (8.4).

planes bounded by the x , y and z axes similarly yield the most probable orientation of $\vec{\Omega}_n$ relative to the input state of polarization.

Clearly, the worst-case states before compensation result from localized regions of this PMD variable space. However, a central result of this chapter is that the geometry and location of these regions depend significantly on the PMD compensator structure. As well, the properties vary with the mean DGD of the fiber emulator, although only a single value of τ_{avg} is examined here for space reasons. Therefore, the most probable system configuration at a given error probability cannot be identified with a single set of values of the system observables that can be applied across different PMD compensators and mean DGD values. However, for a given compensator structure, knowledge of the location of the most probable state could enable rapid numerical simulation and experimental characterization of system performance.

8.4 Conclusions

We have demonstrated that the biased multicanonical technique [83, 117] can be employed to generate ensembles of fiber emulator states with equal eye-closure penalty after optical PMD compensation. Our calculations indicate that large system penalties result from restricted regions of the space of all system configurations, that we have visualized by projecting each iso-penalty emulator ensemble onto a low-dimensional space of system observables. In certain cases, dimensional reduction techniques such as linear or non-linear principal component analysis [1], could be employed to further delineate the geometrical structure of these regions. Significantly, the location of the worst-case states for PMD compensated systems depends on the compensation technique. Knowledge of the position of the worst case state for an uncompensated system cannot therefore be employed to estimate the performance of a particular compensator. Despite this, the agreement between our theoretical predictions and experimental measurements could be used to analyze compensator performance.

Chapter 9

Conclusions

We have developed a general formalism for the Mueller matrix, $(d\mathbf{M}/d\omega)\mathbf{M}^{-1}$, in the presence of both PMD and PDL. Our approach yields recursive methods based on the Magnus expansion for evaluating successive orders of the frequency variation of the Mueller matrix, $\mathbf{M}(\omega)$. Although the results of previous authors [31, 48, 118] can be derived from our formalism through application of the Baker-Campbell-Hausdorff identity, the Magnus expansion explicitly preserves the underlying symmetries of the Mueller matrix and further yields physically realizable operator expansions that facilitate the design of novel joint PMD and PDL compensators.

Next, we reformulated the Lorentz transformation relating the input and output Stokes vectors of an optical system with PMD and PDL with Clifford algebra. Several methods for estimating the Mueller matrix linking the output field polarizations at adjacent optical frequencies that preserved this Lorentz group symmetry were then investigated and a simple computational method proved highly accurate and stable against simulated measurement error. When implemented experimentally with a fast multi-stage lithium niobate polarization rotator coupled to a high-speed polarimeter, significant improvements in experimental accuracy and an order of magnitude reduction in experimental run-time were realized.

While the frequency behaviour of the Jones matrix has typically been characterized by the PMD and PDL coefficients, we demonstrated that other parameterizations are more desirable when higher-order effects are significant. In this context, we presented a highly-accurate quaternion interpolation procedure that can be employed to interpolate the Jones matrix from its values together with the corresponding PMD vectors for a given set of optical frequencies. The increased numerical accuracy and programming efficiency afforded by our method should facilitate improved “all-order” PMD/PDL simulation methods for high bit-rate single or multiple channel optical systems.

Finally, we demonstrated that biased multicanonical sampling [83,117] can be employed to generate ensembles of fiber emulator states with equal eye-closure penalty after optical PMD compensation. Restricted regions of the space of all system configurations were found to induce large system penalties. This result may have implications in the design and testing of optical compensation devices.

Appendix A

PMD/PDL Frequency Evolution

A.1 Principal states of polarization for systems with PMD and PDL

We first derive the Stokes space principal states of polarization (PSPs) for systems with nonzero PMD and PDL, providing a simplified alternative to the methods of Ref. [50]. In terms of the Jones matrix, $\mathbf{T}(\omega)$, c.f. Chapter 2,

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} = -\frac{i}{2}\vec{W} \cdot \vec{\sigma} \quad (\text{A.1})$$

where $\vec{W}(\omega) = \vec{\Omega}(\omega) + i\vec{\Lambda}(\omega)$, with $\vec{\Omega}$, $\vec{\Lambda}$ real vectors. The Jones space PSP, $|p\rangle$, by definition must satisfy the eigenvalue equation

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1}|p\rangle = -\frac{i}{2}\chi|p\rangle \quad (\text{A.2})$$

where $\chi = \pm\sqrt{\vec{W} \cdot \vec{W}} = \pm\sqrt{|\vec{\Omega}|^2 - |\vec{\Lambda}|^2 + 2i\vec{\Omega} \cdot \vec{\Lambda}}$ are complex eigenvalues of $\vec{W} \cdot \vec{\sigma}$. The Stokes space PSP $\hat{p} = \langle p|\vec{\sigma}|p\rangle$ can therefore be expressed after multiplication by $|\chi|^2$ as

$$\begin{aligned} |\chi|^2\hat{p} &= \langle p|\chi^*\vec{\sigma}\chi|p\rangle \\ &= \langle p|(\vec{W} \cdot \vec{\sigma})^\dagger\vec{\sigma}(\vec{W} \cdot \vec{\sigma})|p\rangle \\ &= \langle p|(\vec{W}^* \cdot \vec{\sigma})\vec{\sigma}(\vec{W} \cdot \vec{\sigma})|p\rangle \end{aligned} \quad (\text{A.3})$$

since by Eqs. (A.1)-(A.2), $(\vec{W} \cdot \vec{\sigma})|p\rangle = \chi|p\rangle$. Inserting Eq. (B.15) into Eq. (A.3) yields

$$|\chi|^2 \hat{p} = \vec{W} \langle p | (\vec{W}^* \cdot \vec{\sigma}) | p \rangle + \vec{W}^* \langle p | (\vec{W} \cdot \vec{\sigma}) | p \rangle - |\vec{W}|^2 \langle p | \vec{\sigma} | p \rangle - i(\vec{W}^* \times \vec{W}) \langle p | \sigma_0 | p \rangle \quad (\text{A.4})$$

or, after applying the $(\vec{W} \cdot \vec{\sigma})|p\rangle = \chi|p\rangle$ eigenvalue relation,

$$(|\vec{W}|^2 + |\chi|^2) \hat{p} = \vec{W} \chi^* + \vec{W}^* \chi + i(\vec{W} \times \vec{W}^*) \quad (\text{A.5})$$

Decomposing into real and imaginary components $\chi = \tau + i\eta$ and $\vec{W} = \vec{\Omega} + i\vec{\Lambda}$ generates our central result

$$\hat{p}_\pm = \pm \frac{2}{|\chi|^2 + |\vec{W}|^2} \left(\tau \vec{\Omega} + \eta \vec{\Lambda} \pm \vec{\Omega} \times \vec{\Lambda} \right) \quad (\text{A.6})$$

where \pm correspond to the positive and negative eigenvalues, $\chi = \pm \sqrt{\vec{W} \cdot \vec{W}}$, respectively. In the absence of PDL, $\vec{W} \rightarrow \vec{\Omega}$, and $\hat{p}_\pm = \pm \vec{\Omega} / |\vec{\Omega}|$ as expected.

A.2 The Magnus and Dyson series

The frequency evolution of the Mueller matrix, $\mathbf{M}(\omega)$, is described by the differential equation, c.f. Chapter 2,

$$\frac{d\mathbf{M}}{d\omega} = \mathbf{H}(\omega) \mathbf{M} \quad (\text{A.7})$$

with

$$\mathbf{H}(\omega) \equiv \begin{bmatrix} 0 & \vec{\Lambda}^T \\ \vec{\Lambda} & \vec{\Omega} \times \end{bmatrix} \quad (\text{A.8})$$

Here, $\vec{\Omega}$, $\vec{\Lambda}$ represent the real and imaginary components of the complex principal state vector, respectively [50, 88]. The solution for \mathbf{M} can be found by integrating Eq. (A.7),

$$\mathbf{M}(\omega) = \mathbf{M}(\omega_0) + \int_{\omega_0}^{\omega} d\omega_1 \mathbf{H}(\omega_1) \mathbf{M}(\omega_1) \quad (\text{A.9})$$

so that

$$\mathbf{M}(\omega_1) = \mathbf{M}(\omega_0) + \int_{\omega_0}^{\omega_1} d\omega_2 \mathbf{H}(\omega_2) \mathbf{M}(\omega_2) \quad (\text{A.10})$$

Eq. (A.9) can now be rewritten as the Dyson series solution [71, 75, 90, 118]

$$\begin{aligned} \mathbf{M}(\omega) &= \left[\mathbf{I}_4 + \int_{\omega_0}^{\omega} d\omega_1 \mathbf{H}(\omega_1) + \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) + \dots \right] \mathbf{M}(\omega_0) \\ &\equiv [\mathbf{I}_4 + \mathbf{J}_1 + \mathbf{J}_2 + \mathbf{J}_3 + \mathbf{J}_4 + \dots] \mathbf{M}(\omega_0) \end{aligned} \quad (\text{A.11})$$

with \mathbf{I}_4 the 4×4 identity matrix and

$$\mathbf{J}_n(\omega) \equiv \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \cdots \int_{\omega_0}^{\omega_{n-1}} d\omega_n \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) \cdots \mathbf{H}(\omega_n) \quad (\text{A.12})$$

The Magnus expansion, on the other hand, presupposes a solution of Eq. (A.7) of the general form

$$\mathbf{M}(\omega) = e^{\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3 + \mathbf{B}_4 \cdots} \mathbf{M}(\omega_0) \quad (\text{A.13})$$

Expanding the matrix exponential in power series and equating to Eq. (A.11) by grouping common orders of \mathbf{H} we find [90–92]

$$\mathbf{B}_1 = \mathbf{J}_1 \quad (\text{A.14})$$

$$\mathbf{B}_2 = \mathbf{J}_2 - \frac{\mathbf{J}_1^2}{2} \quad (\text{A.15})$$

$$\mathbf{B}_3 = \mathbf{J}_3 - \frac{1}{2}(\mathbf{J}_1 \mathbf{J}_2 + \mathbf{J}_2 \mathbf{J}_1) + \frac{\mathbf{J}_1^3}{3} \quad (\text{A.16})$$

$$\begin{aligned} \mathbf{B}_4 &= \mathbf{J}_4 - \frac{1}{2}(\mathbf{J}_1 \mathbf{J}_3 + \mathbf{J}_3 \mathbf{J}_1 + \mathbf{J}_2 \mathbf{J}_2) \\ &\quad + \frac{1}{3}(\mathbf{J}_1 \mathbf{J}_1 \mathbf{J}_2 + \mathbf{J}_1 \mathbf{J}_2 \mathbf{J}_1 + \mathbf{J}_2 \mathbf{J}_1 \mathbf{J}_1) - \frac{\mathbf{J}_1^4}{4} \end{aligned} \quad (\text{A.17})$$

with an analogous pattern for higher-order \mathbf{B}_n [90]. Rewriting each Magnus coefficient directly in terms of \mathbf{H} generates as required the equations of Chapter 2. In the case of \mathbf{B}_2 , for example, we have by Eq. (A.15)

$$\begin{aligned} \mathbf{B}_2 &= \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) - \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) \\ &= \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) - \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_1}^{\omega} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) \\ &= \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) - \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_2 \int_{\omega_0}^{\omega_2} d\omega_1 \mathbf{H}(\omega_1) \mathbf{H}(\omega_2) \\ &= \frac{1}{2} \int_{\omega_0}^{\omega} d\omega_1 \int_{\omega_0}^{\omega_1} d\omega_2 [\mathbf{H}(\omega_1), \mathbf{H}(\omega_2)] \end{aligned} \quad (\text{A.18})$$

where in the third expression the order of ω_1 and ω_2 integration was interchanged, while in the fourth ω_1 and ω_2 are integration variables.

A.3 Exponential operator expansions of the Mueller matrix

The frequency evolution of the Mueller matrix, as demonstrated in Chapter 3, is given by the Magnus expansion, which to fifth order in the frequency deviation, $\Delta\omega$, relative to the optical carrier, ω_0 , is

$$\begin{aligned}
\mathbf{M}(\omega)\mathbf{M}(\omega_0)^{-1} &= e^{\mathbf{N}(\Delta\omega)} \\
\mathbf{N}(\Delta\omega) &= \mathbf{H}_0\Delta\omega + \mathbf{H}_1\frac{\Delta\omega^2}{2!} \\
&\quad + (\mathbf{H}_2 - \frac{1}{2}[\mathbf{H}_0, \mathbf{H}_1])\frac{\Delta\omega^3}{3!} \\
&\quad + (\mathbf{H}_3 - [\mathbf{H}_0, \mathbf{H}_2])\frac{\Delta\omega^4}{4!} \\
&\quad + \left(\mathbf{H}_4 - [\mathbf{H}_1, \mathbf{H}_2] - \frac{3}{2}[\mathbf{H}_0, \mathbf{H}_3] \right. \\
&\quad + \frac{1}{6}[\mathbf{H}_0, [\mathbf{H}_0, \mathbf{H}_2]] + \frac{1}{2}[\mathbf{H}_1, [\mathbf{H}_1, \mathbf{H}_0]] \\
&\quad \left. + \frac{1}{6}[\mathbf{H}_0, [\mathbf{H}_0, [\mathbf{H}_0, \mathbf{H}_1]]]\right)\frac{\Delta\omega^5}{5!} + O(\Delta\omega^6).
\end{aligned} \tag{A.19}$$

In the above expressions, $\mathbf{H}_n = (d^n\mathbf{H}/d\omega^n)|_{\omega_0}$, while $[\dots]$ denotes matrix commutation. To express Eq. (A.19) as the product of exponential operators, we first write the desired form of the solution

$$\mathbf{M}(\omega)\mathbf{M}(\omega_0)^{-1} = e^{A_0\mathbf{H}_0\Delta\omega}e^{A_1\mathbf{H}_1\Delta\omega^2}e^{A_2\mathbf{H}_0\Delta\omega}e^{A_3\mathbf{H}_2\Delta\omega^3}\dots \tag{A.20}$$

with A_n unknown coefficients. Repeated application of the BCH identity

$$\begin{aligned}
\exp[\mathbf{F}]\exp[\mathbf{G}] &= \exp\left\{\mathbf{F} + \mathbf{G} + \frac{1}{2}[\mathbf{F}, \mathbf{G}] + \frac{1}{12}[\mathbf{F}, [\mathbf{F}, \mathbf{G}]] + \frac{1}{12}[\mathbf{G}, [\mathbf{G}, \mathbf{F}]] \right. \\
&\quad \left. - \frac{1}{48}[\mathbf{F}, [\mathbf{G}, [\mathbf{F}, \mathbf{G}]]] + \frac{1}{48}[\mathbf{G}, [\mathbf{F}, [\mathbf{G}, \mathbf{F}]]] + \dots\right\}
\end{aligned} \tag{A.21}$$

for two matrices \mathbf{F} and \mathbf{G} generates to $O(\Delta\omega^4)$

$$\begin{aligned} \mathbf{M}(\omega)\mathbf{M}(\omega_0)^{-1} &= \exp\left[(A_0 + A_2)\mathbf{H}_0\Delta\omega + A_1\mathbf{H}_1\Delta\omega^2 \right. \\ &\quad \left. + \frac{1}{2}A_1(A_0 - A_2)[\mathbf{H}_0, \mathbf{H}_1]\Delta\omega^3 + A_3\mathbf{H}_2\Delta\omega^3\right] \end{aligned} \quad (\text{A.22})$$

Comparing this expression with the Magnus solution, Eq. (A.19), yields the system of equations for the coefficients, A_n ,

$$\begin{aligned} A_1 &= \frac{1}{2} \\ A_3 &= \frac{1}{6} \\ A_0 + A_2 &= 1 \\ \frac{1}{2}A_1(A_0 - A_2) &= -\frac{1}{12} \end{aligned} \quad (\text{A.23})$$

Solving Eqs. (A.23),

$$A_0 = \frac{1}{3}, \quad A_1 = \frac{1}{2}, \quad A_2 = \frac{2}{3}, \quad A_3 = \frac{1}{6} \quad (\text{A.24})$$

and

$$\mathbf{M}(\omega)\mathbf{M}(\omega_0)^{-1} = e^{\frac{1}{3}\mathbf{H}_0\Delta\omega} e^{\frac{1}{2}\mathbf{H}_1\Delta\omega^2} e^{\frac{2}{3}\mathbf{H}_0\Delta\omega} e^{\frac{1}{6}\mathbf{H}_2\Delta\omega^3} \dots \quad (\text{A.25})$$

We observe that the number of exponential operators in Eq. (A.20) must be sufficient to enable the cancellation of high-order commutators generated by the BCH identity.

A.4 Higher-order PMD vectors in the concatenated segment model

In this appendix, we derive the exact recursion relation for the higher-order PMD vectors in a cascade of N_{seg} randomly oriented, linearly birefringent segments characterized by the Jones matrix $\mathbf{T}(\omega) = \mathbf{T}_{(N_{seg})}(\omega) \cdots \mathbf{T}_{(2)}(\omega)\mathbf{T}_{(1)}(\omega)$, where $m = 1, \dots, N_{seg}$, and

$$\mathbf{T}_{(m)}(\omega) = \exp\left[-\frac{i}{2}\omega (\vec{\tau}_{(m)} \cdot \vec{\sigma})\right] \quad (\text{A.26})$$

Here we have assumed that the DGD, $\tau_{(m)}$, and slow axis, $\hat{p}_{(m)}$, of the m^{th} segment are both independent of optical frequency, ω , such that $\vec{\tau}_{(m)} \equiv \tau_{(m)}\hat{p}_{(m)}$ is the segment's PMD vector. Each $\vec{\tau}_{(m)}$ further satisfies $(d\mathbf{R}_{(m)}/d\omega) \mathbf{R}_{(m)}^{-1} = \vec{\tau}_{(m)} \times$, where

$$\mathbf{R}_{(m)}(\omega) = \exp[\omega\vec{\tau}_{(m)} \times] \quad (\text{A.27})$$

is the Stokes rotation matrix associated with Eq. (A.26) (see Appendix B.4). The total PMD vector after m such segments, $\vec{\Omega}_{(m)}(\omega)$, can be obtained from the recursion relation [45]

$$\vec{\Omega}_{(m)}(\omega) = \vec{\tau}_{(m)} + \mathbf{R}_{(m)}(\omega)\vec{\Omega}_{(m-1)}(\omega) \quad (\text{A.28})$$

Differentiating Eq. (A.28) yields

$$\begin{aligned} \frac{d\vec{\Omega}_{(m)}}{d\omega} &= \left(\frac{d\mathbf{R}_{(m)}}{d\omega} \mathbf{R}_{(m)}^{-1} \right) \left(\mathbf{R}_{(m)} \vec{\Omega}_{(m-1)} \right) + \mathbf{R}_{(m)} \frac{d\vec{\Omega}_{(m-1)}}{d\omega} \\ &= (\vec{\tau}_{(m)} \times) \left(\vec{\Omega}_{(m)} - \vec{\tau}_{(m)} \right) + \mathbf{R}_{(m)} \frac{d\vec{\Omega}_{(m-1)}}{d\omega} \\ &= (\vec{\tau}_{(m)} \times) \vec{\Omega}_{(m)} + \mathbf{R}_{(m)} \frac{d\vec{\Omega}_{(m-1)}}{d\omega} \end{aligned} \quad (\text{A.29})$$

and

$$\begin{aligned} \frac{d^2\vec{\Omega}_{(m)}}{d\omega^2} &= (\vec{\tau}_{(m)} \times) \frac{d\vec{\Omega}_{(m)}}{d\omega} + \left(\frac{d\mathbf{R}_{(m)}}{d\omega} \mathbf{R}_{(m)}^{-1} \right) \left(\mathbf{R}_{(m)} \frac{d\vec{\Omega}_{(m-1)}}{d\omega} \right) + \mathbf{R}_{(m)} \frac{d^2\vec{\Omega}_{(m-1)}}{d\omega^2} \\ &= (\vec{\tau}_{(m)} \times) \frac{d\vec{\Omega}_{(m)}}{d\omega} + (\vec{\tau}_{(m)} \times) \left(\frac{d\vec{\Omega}_{(m)}}{d\omega} - (\vec{\tau}_{(m)} \times) \vec{\Omega}_{(m)} \right) + \mathbf{R}_{(m)} \frac{d^2\vec{\Omega}_{(m-1)}}{d\omega^2} \\ &= 2(\vec{\tau}_{(m)} \times) \frac{d\vec{\Omega}_{(m)}}{d\omega} - (\vec{\tau}_{(m)} \times)^2 \vec{\Omega}_{(m)} + \mathbf{R}_{(m)} \frac{d^2\vec{\Omega}_{(m-1)}}{d\omega^2} \end{aligned} \quad (\text{A.30})$$

Continuing this procedure generates for the n^{th} derivative, $n \geq 1$, after m birefringent segments,

$$\frac{d^n \vec{\Omega}_{(m)}}{d\omega^n} = \mathbf{R}_{(m)} \frac{d^n \vec{\Omega}_{(m-1)}}{d\omega^n} - \sum_{k=1}^n \binom{n}{k} (-\vec{\tau}_{(m)} \times)^k \frac{d^{n-k} \vec{\Omega}_{(m)}}{d\omega^{n-k}} \quad (\text{A.31})$$

with the initial condition $\vec{\Omega}_{(0)} \equiv 0$. While seemingly complex, Eq. (A.31) has a straightforward numerical implementation in i.e. Matlab, c.f. Appendix C.3. The algorithm is as follows: with $m = 1$, $\vec{\Omega}_{(1)} = \vec{\tau}_{(1)}$, and $d^n \vec{\Omega}_{(1)}/d\omega^n = 0$ for all $n \geq 1$. For $m = 2$, we calculate sequentially each $\vec{\Omega}_{(2)}$, $d\vec{\Omega}_{(2)}/d\omega$, $d^2\vec{\Omega}_{(2)}/d\omega^2$, \dots , $d^n \vec{\Omega}_{(2)}/d\omega^n$ from repeated application of Eq. (A.31). Once all the $d^n \vec{\Omega}_{(2)}/d\omega^n$ are determined, we calculate in order $\vec{\Omega}_{(3)}$, $d\vec{\Omega}_{(3)}/d\omega$, $d^2\vec{\Omega}_{(3)}/d\omega^2$, \dots , $d^n \vec{\Omega}_{(3)}/d\omega^n$. These steps are then performed N_{seg} times.

Appendix C.2 presents a similar algorithm for evaluating derivatives of the Jones matrix.

A.5 Rotating PMD vector approximation

Ref. [74] presents a model of the Jones matrix in which higher-order frequency variation is attributed to a rotation of the PMD vector in the equatorial plane of the Poincaré sphere. In this appendix we establish a generalized result describing the rotation of the PMD vector with constant angular velocity about an arbitrary Stokes space axis. The resulting, simplified Jones matrix is applicable to systems in which polarization-induced chromatic dispersion (PCD) is negligible relative to PSP depolarization.

To proceed, we assume that the PMD vector, $\tau\hat{p}$, rotates with constant angular velocity, k , in Stokes space about the unit vector \hat{u} ,

$$\vec{\Omega}(\omega) \equiv e^{k\omega\hat{u}\times}\tau\hat{p} \quad (\text{A.32})$$

or equivalently in Jones space, c.f. Eq. (B.53),

$$\vec{\Omega}(\omega) \cdot \vec{\sigma} = e^{-\frac{i\omega}{2}k\hat{u}\cdot\vec{\sigma}}(\tau\hat{p} \cdot \vec{\sigma})e^{\frac{i\omega}{2}k\hat{u}\cdot\vec{\sigma}} \quad (\text{A.33})$$

Accordingly, we seek the Jones matrix, $\mathbf{T}(\omega)$, satisfying the matrix differential equation [45, 88]

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} = -\frac{i}{2}\vec{\Omega} \cdot \vec{\sigma} = -\frac{i}{2}e^{-\frac{i\omega}{2}k\hat{u}\cdot\vec{\sigma}}(\tau\hat{p} \cdot \vec{\sigma})e^{\frac{i\omega}{2}k\hat{u}\cdot\vec{\sigma}} \quad (\text{A.34})$$

Eq. (A.34) has an exact solution of the form $\mathbf{T}(\omega) = \mathbf{X}(\omega)\mathbf{Y}(\omega)$ for which $\mathbf{X}(\omega) \equiv \exp[-i\omega k\hat{u} \cdot \vec{\sigma}/2]$ [7]. This is established by inserting $\mathbf{T} = \mathbf{X}\mathbf{Y}$ into Eq. (A.34),

$$\frac{d\mathbf{T}}{d\omega}\mathbf{T}^{-1} = -\frac{i}{2}k(\hat{u} \cdot \vec{\sigma}) + \mathbf{X}\frac{d\mathbf{Y}}{d\omega}\mathbf{Y}^{-1}\mathbf{X}^{-1} \equiv -\frac{i}{2}\mathbf{X}(\tau\hat{p} \cdot \vec{\sigma})\mathbf{X}^{-1} \quad (\text{A.35})$$

or

$$\begin{aligned} \frac{d\mathbf{Y}}{d\omega}\mathbf{Y}^{-1} &= \frac{i}{2}k\mathbf{X}^{-1}(\hat{u} \cdot \vec{\sigma})\mathbf{X} - \frac{i}{2}\tau(\hat{p} \cdot \vec{\sigma}) \\ &= \frac{i}{2}k(\hat{u} \cdot \vec{\sigma}) - \frac{i}{2}\tau(\hat{p} \cdot \vec{\sigma}) \end{aligned} \quad (\text{A.36})$$

in which we have employed $[\mathbf{X}, (\hat{u} \cdot \vec{\sigma})] = 0$ to derive the final expression. Since \hat{p} , \hat{u} , k and τ are assumed to be frequency independent, we can integrate Eq. (A.36), yielding

$$\mathbf{Y}(\omega) = e^{-\frac{i\omega}{2}(\tau\hat{p}-k\hat{u})\cdot\vec{\sigma}} \quad (\text{A.37})$$

The exact Jones matrix solution to the differential equation, Eq. (A.34), is then simply

$$\mathbf{T}(\omega) = e^{-\frac{i\omega}{2}k\hat{u}\cdot\vec{\sigma}}e^{-\frac{i\omega}{2}(\tau\hat{p}-k\hat{u})\cdot\vec{\sigma}} \quad (\text{A.38})$$

Appendix B

Matrix Identities

B.1 Properties of the Pauli spin vector

The following identities, established through direct calculation, are valid for an arbitrary polarization Jones vector, $|t\rangle$, real vectors $\vec{a} = (a_x, a_y, a_z)$ and $\vec{b} = (b_x, b_y, b_z)$, with the corresponding four-vectors \tilde{a} and \tilde{b} defined according to $\tilde{a} = (|\vec{a}|, \vec{a})$ and $\tilde{b} = (|\vec{b}|, \vec{b})$, respectively [8, 38, 45, 50, 59, 88].

$$\begin{aligned}\sigma_0 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \sigma_1 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \sigma_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \sigma_3 &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\end{aligned}\tag{B.1}$$

$$\vec{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)\tag{B.2}$$

$$\tilde{\sigma} \equiv (\sigma_0, \sigma_1, \sigma_2, \sigma_3)\tag{B.3}$$

The matrices σ_n , $n = 0, 1, 2, 3$, satisfy the orthogonality relation, for δ_{nm} the Kronecker delta,

$$\text{Tr}(\sigma_n \sigma_m) = 2\delta_{nm}\tag{B.4}$$

Consequently, σ_n form an orthogonal basis for the space of 2×2 matrices. Accordingly, an arbitrary Jones matrix \mathbf{T} can be written in terms of the complex numbers, t_n , $n = 0, 1, 2, 3$,

as

$$\mathbf{T} = t_0\boldsymbol{\sigma}_0 + t_1\boldsymbol{\sigma}_1 + t_2\boldsymbol{\sigma}_2 + t_3\boldsymbol{\sigma}_3 \quad (\text{B.5})$$

with $t_n = \text{Tr}(\boldsymbol{\sigma}_n \mathbf{T})/2$ [45].

B.1.1 Dot product identities

$$\begin{aligned} \vec{a} \cdot \vec{\sigma} &\equiv (a_x, a_y, a_z) \cdot (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3) \\ &= a_x\boldsymbol{\sigma}_1 + a_y\boldsymbol{\sigma}_2 + a_z\boldsymbol{\sigma}_3 \\ &= \begin{bmatrix} a_x & a_y - ia_z \\ a_y + ia_z & -a_x \end{bmatrix} \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned} \tilde{a} \cdot \tilde{\sigma} &\equiv (|\vec{a}|, a_x, a_y, a_z) \cdot (\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3) \\ &= |\vec{a}|\boldsymbol{\sigma}_0 + a_x\boldsymbol{\sigma}_1 + a_y\boldsymbol{\sigma}_2 + a_z\boldsymbol{\sigma}_3 \\ &= \begin{bmatrix} |\vec{a}| + a_x & a_y - ia_z \\ a_y + ia_z & |\vec{a}| - a_x \end{bmatrix} \end{aligned} \quad (\text{B.7})$$

$$(\vec{a} \cdot \vec{\sigma})^2 = |\vec{a}|^2\boldsymbol{\sigma}_0 \quad (\text{B.8})$$

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b})\boldsymbol{\sigma}_0 + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (\text{B.9})$$

$$(\vec{b} \cdot \vec{\sigma})(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \left[2(\vec{a} \cdot \vec{b})\vec{b} - |\vec{b}|^2\vec{a} \right] \cdot \vec{\sigma} \quad (\text{B.10})$$

$$\left[(\vec{a} \cdot \vec{\sigma}), (\vec{b} \cdot \vec{\sigma}) \right] = 2i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \quad (\text{B.11})$$

In the following, $\vec{w} = \vec{a} + i\vec{b}$ is a complex three-dimensional vector and $|\vec{w}|^2 \equiv \vec{w}^* \cdot \vec{w}$.

$$(\vec{w} \cdot \vec{\sigma})^\dagger = \vec{w}^* \cdot \vec{\sigma} \quad (\text{B.12})$$

$$\vec{\sigma}(\vec{w} \cdot \vec{\sigma}) = \vec{w}\boldsymbol{\sigma}_0 + i\vec{w} \times \vec{\sigma} \quad (\text{B.13})$$

$$(\vec{w}^* \cdot \vec{\sigma})\vec{\sigma} = \vec{w}^*\boldsymbol{\sigma}_0 - i\vec{w}^* \times \vec{\sigma} \quad (\text{B.14})$$

$$(\vec{w}^* \cdot \vec{\sigma})\vec{\sigma}(\vec{w} \cdot \vec{\sigma}) = \vec{w}(\vec{w}^* \cdot \vec{\sigma}) + \vec{w}^*(\vec{w} \cdot \vec{\sigma}) - |\vec{w}|^2\boldsymbol{\sigma}_0 - i(\vec{w}^* \times \vec{w})\boldsymbol{\sigma}_0 \quad (\text{B.15})$$

Eq. (B.15) represents the three 2×2 matrices, with $\vec{w} = (w_x, w_y, w_z)$,

$$(\vec{w}^* \cdot \vec{\sigma})\boldsymbol{\sigma}_1(\vec{w} \cdot \vec{\sigma}) = w_x(\vec{w}^* \cdot \vec{\sigma}) + w_x^*(\vec{w} \cdot \vec{\sigma}) - |\vec{w}|^2\boldsymbol{\sigma}_1 - i(\vec{w}^* \times \vec{w})_x\boldsymbol{\sigma}_0$$

$$(\vec{w}^* \cdot \vec{\sigma})\boldsymbol{\sigma}_2(\vec{w} \cdot \vec{\sigma}) = w_y(\vec{w}^* \cdot \vec{\sigma}) + w_y^*(\vec{w} \cdot \vec{\sigma}) - |\vec{w}|^2\boldsymbol{\sigma}_2 - i(\vec{w}^* \times \vec{w})_y\boldsymbol{\sigma}_0$$

$$(\vec{w}^* \cdot \vec{\sigma})\boldsymbol{\sigma}_3(\vec{w} \cdot \vec{\sigma}) = w_z(\vec{w}^* \cdot \vec{\sigma}) + w_z^*(\vec{w} \cdot \vec{\sigma}) - |\vec{w}|^2\boldsymbol{\sigma}_3 - i(\vec{w}^* \times \vec{w})_z\boldsymbol{\sigma}_0$$

B.1.2 Jones and Stokes vectors

$$|t\rangle = \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (\text{B.16})$$

$$\langle t| = [t_x^*, t_y^*] \quad (\text{B.17})$$

$$\vec{t} \equiv \langle t|\vec{\sigma}|t\rangle \equiv \begin{bmatrix} \langle t|\sigma_1|t\rangle \\ \langle t|\sigma_2|t\rangle \\ \langle t|\sigma_3|t\rangle \end{bmatrix} = \begin{bmatrix} |t_x|^2 - |t_y|^2 \\ t_x t_y^* + t_x^* t_y \\ i(t_x t_y^* - t_x^* t_y) \end{bmatrix} \quad (\text{B.18})$$

$$\tilde{t} \equiv \langle t|\tilde{\sigma}|t\rangle \equiv \begin{bmatrix} \langle t|\sigma_0|t\rangle \\ \langle t|\sigma_1|t\rangle \\ \langle t|\sigma_2|t\rangle \\ \langle t|\sigma_3|t\rangle \end{bmatrix} = \begin{bmatrix} |t_x|^2 + |t_y|^2 \\ |t_x|^2 - |t_y|^2 \\ t_x t_y^* + t_x^* t_y \\ i(t_x t_y^* - t_x^* t_y) \end{bmatrix} \quad (\text{B.19})$$

$$\langle t|t\rangle = |\vec{t}| \quad (\text{B.20})$$

$$|t\rangle\langle t| = \begin{bmatrix} |t_x|^2 & t_x t_y^* \\ t_x^* t_y & |t_y|^2 \end{bmatrix} = \frac{1}{2}(\tilde{t} \cdot \tilde{\sigma}) \quad (\text{B.21})$$

$$\langle t|(\vec{a} \cdot \vec{\sigma})|t\rangle = \vec{a} \cdot \vec{t} \quad (\text{B.22})$$

$$\langle t|(\vec{a} \times \vec{\sigma})|t\rangle \equiv \begin{bmatrix} \langle t|(a_y \sigma_3 - a_z \sigma_2)|t\rangle \\ \langle t|(a_z \sigma_1 - a_x \sigma_3)|t\rangle \\ \langle t|(a_x \sigma_2 - a_y \sigma_1)|t\rangle \end{bmatrix} = \vec{a} \times \vec{t} \quad (\text{B.23})$$

B.2 The Kronecker product

If \mathbf{F} and \mathbf{G} are 2×2 Jones matrices with components f_{ij} and g_{ij} , $i, j = 0, 1$, respectively, the explicit form of the Kronecker product is, c.f. Chapter 2,

$$\begin{aligned} \mathbf{F} \otimes \mathbf{G} &= \begin{bmatrix} f_{00} & f_{01} \\ f_{10} & f_{11} \end{bmatrix} \otimes \begin{bmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{bmatrix} \\ &= \begin{bmatrix} f_{00}\mathbf{G} & f_{01}\mathbf{G} \\ f_{10}\mathbf{G} & f_{11}\mathbf{G} \end{bmatrix} \\ &= \begin{bmatrix} f_{00}g_{00} & f_{00}g_{01} & f_{01}g_{00} & f_{01}g_{01} \\ f_{00}g_{10} & f_{00}g_{11} & f_{01}g_{10} & f_{01}g_{11} \\ f_{10}g_{00} & f_{10}g_{01} & f_{11}g_{00} & f_{11}g_{01} \\ f_{10}g_{10} & f_{10}g_{11} & f_{11}g_{10} & f_{11}g_{11} \end{bmatrix}. \end{aligned} \quad (\text{B.24})$$

Representing $N_X \times N_X$ and $N_Y \times N_Y$ square matrices by \mathbf{X} and \mathbf{Y} , we have

$$(\mathbf{X} \otimes \mathbf{Y})(\mathbf{J} \otimes \mathbf{K}) = (\mathbf{XJ}) \otimes (\mathbf{YK}) \quad (\text{B.25})$$

$$(\mathbf{X} \otimes \mathbf{Y})^{-1} = \mathbf{X}^{-1} \otimes \mathbf{Y}^{-1} \quad (\text{B.26})$$

$$\text{Tr}(\mathbf{X} \otimes \mathbf{Y}) = \text{Tr}(\mathbf{X})\text{Tr}(\mathbf{Y}) \quad (\text{B.27})$$

$$\det(\mathbf{X} \otimes \mathbf{Y}) = (\det \mathbf{X})^{N_Y} (\det \mathbf{Y})^{N_X} \quad (\text{B.28})$$

$$(\mathbf{X} \otimes \mathbf{Y})^n = (\mathbf{X}^n) \otimes (\mathbf{Y}^n) \quad (\text{B.29})$$

Above, \mathbf{J} and \mathbf{K} are matrices with dimensions $N_X \times N_J$ and $N_Y \times N_K$, respectively. Further, denoting the $N \times N$ identity matrix as \mathbf{I}_N ,

$$\vec{a} \times \equiv \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \quad (\text{B.30})$$

$$\begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & \vec{b} \times \end{bmatrix} \equiv \begin{bmatrix} 0 & a_x & a_y & a_z \\ a_x & 0 & -b_z & b_y \\ a_y & b_z & 0 & -b_x \\ a_z & -b_y & b_x & 0 \end{bmatrix} \quad (\text{B.31})$$

$$\mathbf{A} \equiv \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \quad (\text{B.32})$$

$$\det \mathbf{A} = -\det \mathbf{A}^\dagger = i \quad (\text{B.33})$$

$$\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}_4 \quad (\text{B.34})$$

$$\mathbf{A}((\vec{a} \cdot \vec{\sigma}) \otimes \sigma_0)\mathbf{A}^\dagger = \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & i\vec{a} \times \end{bmatrix} \quad (\text{B.35})$$

$$\mathbf{A}(\sigma_0 \otimes (\vec{a} \cdot \vec{\sigma})^*)\mathbf{A}^\dagger = \begin{bmatrix} 0 & \vec{a}^T \\ \vec{a} & -i\vec{a} \times \end{bmatrix} \quad (\text{B.36})$$

$$\mathbf{A}(\vec{a} \cdot \vec{\sigma}) \otimes (\vec{b} \cdot \vec{\sigma})^* \mathbf{A}^\dagger = \begin{bmatrix} \vec{a} \cdot \vec{b} & -i(\vec{a} \times \vec{b})^T \\ i(\vec{a} \times \vec{b}) & \vec{a}\vec{b}^T + \vec{b}\vec{a}^T - (\vec{a} \cdot \vec{b})\mathbf{I}_3 \end{bmatrix} \quad (\text{B.37})$$

B.3 The Mueller matrix

B.3.1 Relation to the Jones matrix

Let \mathbf{T} be a complex 2×2 Jones matrix such that the input and output Jones vectors, $|s\rangle$ and $|t\rangle$, respectively, are related through $|t\rangle = \mathbf{T}|s\rangle$. The 4×1 Stokes vector, \tilde{t} , is given by, c.f. Eq. (B.19),

$$\tilde{t} \equiv \langle t | \tilde{\sigma} | t \rangle = \langle s | \mathbf{T}^\dagger \tilde{\sigma} \mathbf{T} | s \rangle \quad (\text{B.38})$$

Noting that $\mathbf{T}^\dagger \sigma_i \mathbf{T}$ is Hermitian for each $i = 0, 1, 2, 3$, the components m_{ij} in the expansion $\mathbf{T}^\dagger \sigma_i \mathbf{T} \equiv m_{i0} \sigma_0 + m_{i1} \sigma_1 + m_{i2} \sigma_2 + m_{i3} \sigma_3$ are real, so that

$$\begin{bmatrix} \mathbf{T}^\dagger \sigma_0 \mathbf{T} \\ \mathbf{T}^\dagger \sigma_1 \mathbf{T} \\ \mathbf{T}^\dagger \sigma_2 \mathbf{T} \\ \mathbf{T}^\dagger \sigma_3 \mathbf{T} \end{bmatrix} = \begin{bmatrix} m_{00} \sigma_0 + m_{01} \sigma_1 + m_{02} \sigma_2 + m_{03} \sigma_3 \\ m_{10} \sigma_0 + m_{11} \sigma_1 + m_{12} \sigma_2 + m_{13} \sigma_3 \\ m_{20} \sigma_0 + m_{21} \sigma_1 + m_{22} \sigma_2 + m_{23} \sigma_3 \\ m_{30} \sigma_0 + m_{31} \sigma_1 + m_{32} \sigma_2 + m_{33} \sigma_3 \end{bmatrix} \quad (\text{B.39})$$

From Eq. (B.4) [45, 50] we have

$$\begin{aligned} m_{ij} &= \frac{1}{2} \text{Tr}(\sigma_j \mathbf{T}^\dagger \sigma_i \mathbf{T}) \\ &= \frac{1}{2} \text{Tr}(\sigma_i \mathbf{T} \sigma_j \mathbf{T}^\dagger) \end{aligned} \quad (\text{B.40})$$

The equivalence between Eq. (B.40) and the Kronecker product $\mathbf{M} = \mathbf{A}(\mathbf{T} \otimes \mathbf{T}^*)\mathbf{A}^\dagger$, for \mathbf{M} a 4×4 matrix with components m_{ij} , $i, j = 0, 1, 2, 3$, was established in Chapter 2.

In terms of the Mueller matrix, \mathbf{M} , Eq. (B.39) reduces to $\mathbf{T}^\dagger \tilde{\sigma} \mathbf{T} \equiv \mathbf{M} \tilde{\sigma}$. Inserting this expression into Eq. (B.38), we obtain as a result of the linearity of Eq. (B.39), $\tilde{t} = \langle s | \mathbf{M} \tilde{\sigma} | s \rangle = \mathbf{M} \langle s | \tilde{\sigma} | s \rangle = \mathbf{M} \tilde{s}$, that is, if the input and output Jones vectors are related through $|t\rangle = \mathbf{T}|s\rangle$, the corresponding input and output Stokes vectors are given by $\tilde{t} = \mathbf{M} \tilde{s}$ [18, 45].

B.3.2 General properties

If \mathbf{K} is an arbitrary 2×2 complex matrix [8],

$$m_{ij} = \frac{1}{2} \text{Tr}(\sigma_i \mathbf{T} \sigma_j \mathbf{T}^\dagger) \iff \mathbf{M} = \mathbf{A}(\mathbf{T} \otimes \mathbf{T}^*) \mathbf{A}^\dagger \quad (\text{B.41})$$

$$\mathbf{T} = \exp[\mathbf{K}] \iff \mathbf{M} = \exp[\mathbf{A}(\mathbf{K} \otimes \boldsymbol{\sigma}_0 + \boldsymbol{\sigma}_0 \otimes \mathbf{K}^*) \mathbf{A}^\dagger] \quad (\text{B.42})$$

$$\mathbf{T} = \exp\left[-\frac{i}{2}(\vec{b} + i\vec{a}) \cdot \vec{\sigma}\right] \iff \mathbf{M} = \exp\begin{bmatrix} 0 & \vec{a}^\top \\ \vec{a} & \vec{b} \times \end{bmatrix} \quad (\text{B.43})$$

$$\mathbf{g} \equiv \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{B.44})$$

$$\det \mathbf{M} = |\det \mathbf{T}|^4 \quad (\text{B.45})$$

$$\mathbf{g} \mathbf{M}^\top \mathbf{g} \mathbf{M} = \mathbf{I}_4 \sqrt{\det \mathbf{M}} \quad (\text{B.46})$$

$$\sqrt{\det \mathbf{M}} = \frac{1}{2} \text{Tr}(\mathbf{M}^\top \mathbf{g} \mathbf{M}) \quad (\text{B.47})$$

$$\mathbf{M}^{-1} = 2\mathbf{g} \mathbf{M}^\top \mathbf{g} / \text{Tr}(\mathbf{M}^\top \mathbf{g} \mathbf{M}) \quad (\text{B.48})$$

For partially polarized light, with $\tilde{s} = (s_0, \vec{s})$ and $s_0 \neq |\vec{s}|$, $\tilde{s}^\top \mathbf{g} \tilde{s} = |\vec{s}|^2 - s_0^2$ transforms as, c.f. Eq. (B.46),

$$\begin{aligned} \tilde{t}^\top \mathbf{g} \tilde{t} &= \tilde{s}^\top (\mathbf{M}^\top \mathbf{g} \mathbf{M}) \tilde{s} \\ &= \sqrt{\det \mathbf{M}} (\tilde{s}^\top \mathbf{g} \tilde{s}) \end{aligned} \quad (\text{B.49})$$

and is therefore invariant if $\det \mathbf{M} = 1$.

B.3.3 Coherency matrix transformations

Let $\tilde{t} = \langle t | \vec{\sigma} | t \rangle$ be the 4×1 Stokes vector corresponding to the Jones vector $|t\rangle = \mathbf{T}|s\rangle$. Defining the coherency matrix [18] as $|t\rangle\langle t|$, we have through Eq. (B.21)

$$|t\rangle\langle t| = \mathbf{T}|s\rangle\langle s|\mathbf{T}^\dagger = \frac{1}{2} \mathbf{T}(\tilde{s} \cdot \vec{\sigma}) \mathbf{T}^\dagger \quad (\text{B.50})$$

Alternatively, from Sec. B.3.1,

$$|t\rangle\langle t| = \frac{1}{2}(\tilde{t} \cdot \tilde{\boldsymbol{\sigma}}) = \frac{1}{2}(\mathbf{M}\tilde{\mathbf{s}}) \cdot \tilde{\boldsymbol{\sigma}} \quad (\text{B.51})$$

which implies

$$\mathbf{T}(\tilde{\mathbf{s}} \cdot \tilde{\boldsymbol{\sigma}})\mathbf{T}^\dagger = (\mathbf{M}\tilde{\mathbf{s}}) \cdot \tilde{\boldsymbol{\sigma}} \quad (\text{B.52})$$

Specializing to the case of unitary \mathbf{T} , a similar derivation yields

$$\mathbf{T}(\vec{\mathbf{s}} \cdot \vec{\boldsymbol{\sigma}})\mathbf{T}^\dagger = (\mathbf{R}\vec{\mathbf{s}}) \cdot \vec{\boldsymbol{\sigma}} \quad (\text{B.53})$$

The Stokes rotation matrix, \mathbf{R} , is the 3×3 submatrix in the lower right-hand corner of \mathbf{M} , c.f. Appendix B.4.

B.4 Representations of unitary Jones matrices

Consider a unitary 2×2 complex matrix, \mathbf{U} , satisfying $\det \mathbf{U} \equiv 1$. Defining \dagger and $*$ to be Hermitian and complex conjugation, respectively, the unitary condition $\mathbf{U}^\dagger \mathbf{U} = \boldsymbol{\sigma}_0$ implies that \mathbf{U} can be written

$$\mathbf{U} = \begin{bmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{bmatrix} \quad (\text{B.54})$$

The complex elements α and β , satisfying $\det \mathbf{U} = |\alpha|^2 + |\beta|^2 \equiv 1$, can be further decomposed into real and imaginary components, $\alpha = u + iv$ and $\beta = x + iy$, so that

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} u + iv & -x + iy \\ x + iy & u - iv \end{bmatrix} \\ &= u \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + iv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + iy \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - ix \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\ &\equiv u\boldsymbol{\sigma}_0 + iv\boldsymbol{\sigma}_1 + iy\boldsymbol{\sigma}_2 - ix\boldsymbol{\sigma}_3 \end{aligned} \quad (\text{B.55})$$

Defining $\vec{\mathbf{r}} \equiv (-v, -y, x)$, Eq. (B.55) reduces to

$$\begin{aligned} \mathbf{U} &= u\boldsymbol{\sigma}_0 + iv\boldsymbol{\sigma}_1 + iy\boldsymbol{\sigma}_2 - ix\boldsymbol{\sigma}_3 \\ &= u\boldsymbol{\sigma}_0 - i(\vec{\mathbf{r}} \cdot \vec{\boldsymbol{\sigma}}) \\ &= u\boldsymbol{\sigma}_0 - ir(\hat{\mathbf{r}} \cdot \vec{\boldsymbol{\sigma}}) \end{aligned} \quad (\text{B.56})$$

with $r = \sqrt{v^2 + y^2 + x^2}$, and $\hat{r} = \vec{r}/r$. Writing $\tan(\psi/2) \equiv r/u$ then yields

$$\begin{aligned} \mathbf{U} &= \cos\left(\frac{\psi}{2}\right) \boldsymbol{\sigma}_0 - i \sin\left(\frac{\psi}{2}\right) (\hat{r} \cdot \vec{\boldsymbol{\sigma}}) \\ &= \left[1 - \frac{1}{2!} \left(\frac{\psi}{2}\right)^2 + \frac{1}{4!} \left(\frac{\psi}{2}\right)^4 - \dots\right] \boldsymbol{\sigma}_0 \\ &\quad - i \left[\left(\frac{\psi}{2}\right) - \frac{1}{3!} \left(\frac{\psi}{2}\right)^3 + \frac{1}{5!} \left(\frac{\psi}{2}\right)^5 - \dots\right] (\hat{r} \cdot \vec{\boldsymbol{\sigma}}) \end{aligned} \quad (\text{B.57})$$

or, upon rearranging terms according to Eq. (B.8),

$$\begin{aligned} \mathbf{U} &= \boldsymbol{\sigma}_0 - \frac{i\psi}{2} (\hat{r} \cdot \vec{\boldsymbol{\sigma}}) + \frac{1}{2!} \left(-\frac{i\psi}{2}\right)^2 (\hat{r} \cdot \vec{\boldsymbol{\sigma}})^2 \\ &\quad + \frac{1}{3!} \left(-\frac{i\psi}{2}\right)^3 (\hat{r} \cdot \vec{\boldsymbol{\sigma}})^3 + \frac{1}{4!} \left(-\frac{i\psi}{2}\right)^4 (\hat{r} \cdot \vec{\boldsymbol{\sigma}})^4 + \dots \\ &\equiv \exp\left[-\frac{i}{2}\psi(\hat{r} \cdot \vec{\boldsymbol{\sigma}})\right] \end{aligned} \quad (\text{B.58})$$

Observe that the eigenvalues and Stokes space eigenvectors of \mathbf{U} enter Eq. (B.58) explicitly through $\exp(\pm i\psi/2)$ and $\pm \hat{r}$, respectively.

The Mueller matrix, Eq. (B.43), for unitary \mathbf{U} , is specified by its lower right-hand 3×3 submatrix, denoted here by \mathbf{R} ,

$$\mathbf{U} = e^{-\frac{i}{2}\psi(\hat{r} \cdot \vec{\boldsymbol{\sigma}})} \iff \mathbf{R} = e^{\psi \hat{r} \times} \quad (\text{B.59})$$

Expanding $\mathbf{R} = \exp[\psi \hat{r} \times]$ in power series and applying $(\hat{r} \times)^2 = \hat{r} \hat{r}^T - \mathbf{I}_3$ yields [45]

$$e^{\psi \hat{r} \times} = \cos(\psi) \mathbf{I}_3 + [1 - \cos(\psi)] \hat{r} \hat{r}^T + \sin(\psi) \hat{r} \times \quad (\text{B.60})$$

Further, Eq. (B.60) implies

$$\text{Tr}(\mathbf{R}) = 1 + 2 \cos(\psi) \quad (\text{B.61})$$

$$\mathbf{R} - \mathbf{R}^T = 2 \sin(\psi) (\hat{r} \times) \quad (\text{B.62})$$

Eqs. (B.61)-(B.62) specify ψ and \hat{r} if \mathbf{R} is known from experimental measurement or numerical simulation.

Appendix C

MATLAB Function Implementation

C.1 Mueller matrix SVD in Minkowski space

```
%—  
% MinkowskiSVD:  
% Returns the singular value decomposition of the  
% Mueller matrix M in Minkowski space. More  
% specifically, at output, M is decomposed according to:  
% M = USV  
% Where  $\mathbf{gU}^T\mathbf{gU} = \mathbf{I}$   
%  $\mathbf{gV}^T\mathbf{gV} = \mathbf{I}$   
% S = diagonal matrix  
% g = diag(-1,1,1,1)  
% The basic algorithm is adapted from:  
% D. Tweed, ‘‘Estimating rigid motions via the conformal model  
% of Euclidean space’’, Proc. Intern.  
% Conf. on Patt. Rec., vol. 2, pp. 171–174, 2004  
%—  
function [U,S,V] = MinkowskiSVD(M)  
g = diag([-1 1 1 1]);  
[E,D] = eig(g * transpose(M) * g * M);  
Ep = E * (g * transpose(E) * g * E)^(-0.5);  
N = g * transpose(Ep) * g * M * Ep;  
Q = N * D^(-0.5);  
U = Ep * Q;  
S = D^(0.5);
```

```
V = g * transpose(Ep) * g;
return;
```

C.2 Exact evaluation of Jones matrix derivatives

```
%—
% EmulateJonesMatrix:
% Expands the Jones matrix of an  $N_{sec}$  fiber emulator in
% the Taylor series  $\mathbf{T}(\omega) = \sum_{n=0}^N \mathbf{T}_n \Delta\omega^n$ 
% where  $\mathbf{T}_n \equiv \frac{1}{n!} \frac{d^n \mathbf{T}}{d\omega^n}$ .
% Input Variables:
%   rotAxis =  $3 \times N_{sec}$  matrix. Each column contains a unit
%   vector uniformly distributed on the Poincare sphere
%   (see RandomUnitVector)
%   dgd =  $1 \times N_{sec}$  array containing the
%   differential group delay of each emulator segment [ns].
%   norder = Order of the Taylor expansion.
%   w0 = Optical carrier frequency [Grad/s]
% Output Variables:
%   T =  $2 \times 2 \times (1 + norder)$  matrix.
%   T(:, :, 1) =  $\mathbf{T}_0$ 
%   T(:, :, 2) =  $\mathbf{T}_1$ 
%   T(:, :, 3) =  $\mathbf{T}_2$ 
%   :
%—
function [T] = EmulateJonesMatrix(rotAxis, dgd, norder, w0)
% Useful constants
I = eye(2,2);
Z22 = zeros(2,2);
c = cos(0.5 .* dgd .* w0);
is = i.*sin(0.5 .* dgd .* w0);

% Recursively calculate all Jones matrix derivatives
T = zeros(2,2,1 + norder);
T(:, :, 1) = I;

% Loop over emulator segments
for segIx = 1:length(rotAxis);
    % This segment's Jones matrix
```

```

M = VecDotPauli(rotAxis(:, segIx)); %  $\mathbf{M} = \hat{n} \cdot \vec{\sigma}$ 
U = c(segIx)*I - is(segIx) * M; %  $\mathbf{U} = \exp[-i\omega_0\tau(\hat{n} \cdot \vec{\sigma})/2]$ 
M = 0.5 * i * dgd(segIx) * M; %  $\mathbf{M} = i\tau\hat{n} \cdot \vec{\sigma}/2$ 

% Implements derivative recursion relation
T(:, :, 1) = U * T(:, :, 1); % Jones matrix at  $\omega_0$ 
for n = 1:norder;
    S = Z22;
    for k = 1:n; S = S - binomial(n, k)*M^k*T(:, :, 1 + n-k); end;
    T(:, :, 1 + n) = S + U*T(:, :, 1 + n); %  $n^{\text{th}}$  derivative
end
end
for n = 2:norder
    T(:, :, 1 + n) = T(:, :, 1 + n) / factorial(n);
end
return

```

C.3 Exact evaluation of PMD vector derivatives

```

%—
% EmulatePMDVector:
% Same as EmulateJonesMatrix, except that this
% function evaluates all orders of the PMD vector
% at the carrier frequency.
% The PMD vector is defined as  $\vec{\Omega}(\omega) \cdot \vec{\sigma} \equiv 2i \frac{d\mathbf{T}}{d\omega} \mathbf{T}^{-1}$ 
% The  $n^{\text{th}}$  order PMD vector is defined as  $\vec{\Omega}_n \equiv \frac{d^n \vec{\Omega}(\omega)}{d\omega^n}$ 
% Input Variables:
% Same as EmulateJonesMatrix
% Output Variables:
%  $\text{PMD} = 3 \times (1 + \text{norder})$  matrix. The  $n^{\text{th}}$  order
% PMD vector has units of  $[\text{ns}]^n$ .
%  $\text{PMD}(:, 1) = \vec{\Omega}_0$ 
%  $\text{PMD}(:, 2) = \vec{\Omega}_1$ 
%  $\text{PMD}(:, 3) = \vec{\Omega}_2$ 
%
%
%—
function [PMD] = EmulatePMDVector(rotAxis, dgd, norder, w0)
% Useful constants

```

```

Z31 = zeros(3,1);

% Initialize recursion
PMD = zeros(3, 1+norder);

% Loop over all segments
for segIx = 1:length(rotAxis);
    segpmd = dgd(segIx) .* rotAxis(:,segIx); % Segment PMD vector:  $\tau \hat{n}$ 
    M = MatrixCross(-segpmd); %  $\mathbf{M} = -\tau \hat{n} \times$ 
    R = StokesRotation(w0.*dgd(segIx), rotAxis(:, segIx)); %  $\mathbf{R} = \exp[\omega_0 \tau \hat{n} \times]$ 

    % Recursion for PMD vectors
    PMD(:,1) = segpmd + R*PMD(:,1); % First order PMD
    for n = 1:norder
        S = Z31;
        for k = 1:n; S = S - binomial(n,k)*M^k*PMD(:,1 + n-k); end;
        PMD(:,1 + n) = S + R*PMD(:,1 + n);
    end
end
return

```

C.4 Jones matrix interpolation

```

%—
% EqualSpacedPMD:
% Evaluates the Jones matrix and PMD vector at a series
% of  $N_{grid}$  optical frequencies equally spaced by  $\delta$ .
% Input Variables:
%   rotAxis =  $3 \times N_{sec}$  matrix. Each column contains a unit
%   vector uniformly distributed on the Poincare sphere
%   (see RandomUnitVector)
%   dgd =  $N_{sec} \times 1$  array containing the DGD of each emulator segment [ns].
%   w0 = Optical carrier frequency [Grad].
%   wgrid =  $N_{grid} \times 1$  array of equally spaced optical frequencies [Grad].
% Output Variables:
%   T =  $2 \times 2 \times N_{grid}$  array of Jones matrices.
%   T(:, :, 1) =  $\mathbf{T}(\omega_0)$ 
%   T(:, :, 2) =  $\mathbf{T}(\omega_0 + \delta)$ 
%   T(:, :, 3) =  $\mathbf{T}(\omega_0 + 2\delta)$ 

```

```

%      :
%      PMD = 3 × Ngrid array of PMD vectors.
%      PMD(:, 1) =  $\vec{\Omega}(\omega_0)$ 
%      PMD(:, 2) =  $\vec{\Omega}(\omega_0 + \delta)$ 
%      PMD(:, 3) =  $\vec{\Omega}(\omega_0 + 2\delta)$ 
%      :
%—
function [T, PMD] = EqualSpacedPMD(rotAxis, dgd, w0, wgrid)
Ngrid = length(wgrid);
T = zeros(2, 2, Ngrid);
PMD = zeros(3, Ngrid);
for n=1:Ngrid
    % U(:, :, 1) = T
    % U(:, :, 2) = dT/dω
    U = EmulateJonesMatrix(rotAxis, dgd, 1, w0 + wgrid(n));
    T(:, :, n) = U(:, :, 1);
    [notused, PMD(:, n)] = PauliDecomp(2*i*U(:, :, 2)*inv(U(:, :, 1)));
end
return

%—
% CalcInterpParam:
% Evaluates the interpolating Jones matrices
%  $\mathbf{A}_{(n)} = \exp\left[-\frac{i\delta}{6} \vec{\Omega}_{(n)} \cdot \vec{\sigma}\right] \mathbf{T}_{(n)}$  and  $\mathbf{B}_{(n)} = \exp\left[\frac{i\delta}{6} \vec{\Omega}_{(n+1)} \cdot \vec{\sigma}\right] \mathbf{T}_{(n+1)}$ 
% in each frequency interval  $[wgrid_n, wgrid_{n+1}]$ ,  $n = 1, \dots, (N_{grid} - 1)$ .
% Takes as input  $\mathbf{T}(\omega)$  and  $\vec{\Omega}(\omega)$  as calculated by EqualSpacedPMD.
% Input Variables:
% wgrid = Ngrid × 1 array of equally spaced optical frequencies [Grad].
% T = 2 × 2 × Ngrid array of Jones matrices.
% PMD = 3 × Ngrid array of PMD vectors.
% Output Variables:
% InterpParam = Matlab cell array with 4 entries,  $n = 1, \dots, (N_{grid} - 1)$ 
% InterpParam{1}(:, :, n) =  $\mathbf{T}_{(n+1)} \mathbf{B}_{(n)}^{-1}$ 
% InterpParam{2}(:, :, n) =  $\mathbf{B}_{(n)} \mathbf{A}_{(n)}^{-1}$ 
% InterpParam{3}(:, :, n) =  $\mathbf{A}_{(n)} \mathbf{T}_{(n)}^{-1}$ 
% InterpParam{4}(:, :, n) =  $\mathbf{T}_{(n)}$ 
%—

```

```

function [InterpParam] = CalcInterpParam(wgrid, T, PMD);
Ngrid = length(wgrid);
InterpParam{1} = zeros(2,2,Ngrid);
InterpParam{2} = zeros(2,2,Ngrid);
InterpParam{3} = zeros(2,2,Ngrid);
InterpParam{4} = zeros(2,2,Ngrid);
delta = wgrid(2) - wgrid(1); % Assume equal spacing [Grad]
for n = 1:(Ngrid-1)
    A = expm(-i.*delta./6.*VecDotPauli(PMD(:, n))) * T(:, :, n);
    B = expm(i.*delta./6.*VecDotPauli(PMD(:, n+1)))*T(:, :, n+1);
    InterpParam{1}(:, :, n) = T(:, :, n+1) * inv(B);
    InterpParam{2}(:, :, n) = B * inv(A);
    InterpParam{3}(:, :, n) = A * inv(T(:, :, n));
    InterpParam{4}(:, :, n) = T(:, :, n);
end
return

%—
% JonesInterp:
% Interpolates between Jones matrices and PMD vectors
% recorded at a series of frequencies equally spaced by  $\delta$ .
% Use of this function:
% (1) For each emulator realization call EqualSpacedPMD
% to evaluate the Jones matrix and PMD vector at a number
% of equally spaced optical frequencies.
% (2) Call CalcInterpParam once to evaluate interpolation overhead.
% (3) Call JonesInterp to interpolate the Jones matrix at any
% desired number of intermediate frequencies  $\omega$ .
% Input Variables:
% wgrid =  $N_{grid} \times 1$  array of equally spaced optical frequencies [Grad].
% InterpParam = Output of CalcInterpParam.
% w =  $N_{interp} \times 1$  array of intermediate optical frequencies [Grad].
% Output Variables:
% Tinterp =  $2 \times 2 \times N_{interp}$  array of Jones matrices
% interpolated at each  $w_n$ ,  $n = 1, \dots, N_{interp}$ .
%—
function [Tinterp] = JonesInterp(wgrid, InterpParam, w)
Ninterp = length(w);
Tinterp = zeros(2,2,Ninterp);
for n = 1:Ninterp

```

```

% Bracket wgrid(klo) ≤ wn ≤ wgrid(khi)
klo = max(find(wgrid < w(n)));
if (isempty(klo)); klo = 1; end;
khi = klo + 1;
if (khi > length(wgrid))
    Tinterp(:, :, n) = InterpParam{4}(:, :, end);
else
    % Interpolate
    t = (w(n) - wgrid(klo))/(wgrid(khi) - wgrid(klo)); % 0 ≤ t ≤ 1
    b1 = 1 - (1 - t)^3;
    b2 = t^2 * (3 - 2*t);
    b3 = t^3;
    Tinterp(:, :, n) = InterpParam{1}(:, :, klo)^b3 * ...
        InterpParam{2}(:, :, klo)^b2 * ...
        InterpParam{3}(:, :, klo)^b1 * ...
        InterpParam{4}(:, :, klo);
end
end
return

```

C.5 Jones matrix Padé approximant

```

%—
% JonesPade:
% Rearranges the 2 × 2 × N array
% of Taylor series coefficients as calculated
% by EmulateJonesMatrix into the corresponding
% series coefficients for the [Q,M]-Padé approximant:
% 
$$\sum_{n=0}^N \mathbf{T}_n \Delta\omega^n = \left[ \sum_{q=0}^Q \mathbf{F}_q \Delta\omega^q \right] \left[ \mathbf{I} + \sum_{m=1}^M \mathbf{G}_m \Delta\omega^m \right]^{-1}$$

% Note that this implementation requires Q + M ≤ N.
% Input Variables:
% T = 2 × 2 × N array of Taylor coefficients
% as returned by the function EmulateJonesMatrix
% Q, M = Positive integers satisfying Q + M ≤ N
% Output Variables:
% F = 2 × 2 × (Q + 1) array of Padé series coefficients
% G = 2 × 2 × M array of Padé series coefficients
%—

```

```

function [F, G] = JonesPade(T, Q, M)
[r, c, N] = size(T);

% Calculate required block matrices
TG = zeros(2*M, 2); % M x 1 block matrix used to calculate G
C = zeros(2*M, 2*M); % M x M block matrix
for k = 1:M
    TG( (2*k-1):2*k, :) = T(:, :, 1 + Q+k);
    for n = 1:min(M, Q+k)
        C( (2*k-1):2*k, (2*n-1):2*n ) = T(:, :, 1 + Q+k-n);
    end
end
TF = zeros(2*Q, 2); % Q x 1 block matrix used to calculate F
D = zeros(2*Q, 2*M); % Q x M block matrix
for k = 1:Q
    TF( (2*k-1):2*k, :) = T(:, :, 1 + k);
    for n = 1:min(M, k)
        D( (2*k-1):2*k, (2*n-1):2*n ) = T(:, :, 1 + k-n);
    end
end
% Calculate F and G coefficients in block matrix form
G = -inv(C) * TG;
F1 = TF + D * G;

% Rearrange block matrices into arrays of 2 x 2 matrices
F1 = reshape(transpose(F1), 2, 2, Q);
G = reshape(transpose(G), 2, 2, M);
F = zeros(2, 2, Q+1);
F(:, :, 1) = T(:, :, 1);
for k = 1:Q
    F(:, :, 1 + k) = transpose(F1(:, :, k));
end
for k = 1:M
    G(:, :, k) = transpose(G(:, :, k));
end
return

%—
% JonesPadeEval:
% Evaluates the Jones matrix Pade approximant

```

```

%  $\mathbf{T}_{pade} \equiv \left[ \sum_{q=0}^Q \mathbf{F}_q \Delta\omega^q \right] \left[ \mathbf{I} + \sum_{m=1}^M \mathbf{G}_m \Delta\omega^m \right]^{-1}$ 
% at specified values of  $\Delta\omega$ .
% Input Variables:
% Takes as input  $\mathbf{F}$  and  $\mathbf{G}$  as returned by JonesPade.
%  $F = 2 \times 2 \times (Q+1)$  array of Pade series coefficients
%  $G = 2 \times 2 \times M$  array of Pade series coefficients
%  $deltaw = \Delta\omega = N_{\Delta\omega} \times 1$  array of frequency deviations
% relative to the optical carrier [Grad].
% Output Variables:
%  $Tpade = 2 \times 2 \times N_{\Delta\omega}$  array of Jones matrices
% calculated at each  $\Delta\omega_i$ ,  $i = 1, \dots, N_{\Delta\omega}$ .
%—
function [Tpade] = JonesPadeEval(F, G, deltaw)
[r,c,Q] = size(F);
[r,c,M] = size(G);
Ndelta = length(deltaw);
Tpade = zeros(2,2,Ndelta);
for wix = 1:Ndelta
    SUM1 = F(:, :, 1);
    for q = 1:(Q-1)
        SUM1 = SUM1 + F(:, :, 1 + q) .* deltaw(wix).^q;
    end
    SUM2 = eye(2,2);
    for m = 1:M
        SUM2 = SUM2 + G(:, :, m) .* deltaw(wix).^m;
    end
    Tpade(:, :, wix) = SUM1 * inv(SUM2);
end
return

```

C.6 Optimum biased multicanonical normalization

```

%—
% NormalizationFactors:
% Given  $N$  estimates of a PDF within overlapping
% biasing regions, this routine iteratively solves
% for the normalization constants minimizing
% the mean square error (MSE) of the combined

```

```

% PDF estimate.
% Let nbins = number of histogram bins
%     nbias = number of biasing regions
%
% Input variables:
%   x = (nbins × 1) array
%       Contains the midpoint of each histogram bin.
%   h = (nbins × nbias) matrix
%       The n:th column of h corresponds to the histogram
%       as calculated for the n:th biasing region.
%       The first column of h is the Monte Carlo PDF estimate.
%
% Output variables:
%   combinedPDF = (nbins × 1) array
%       Contains the optimum MSE estimate of the
%       PDF after combining all nbias biased estimates.
%   A = (nbias × 1) array
%       Array of the converged normalization factors
%—
function [combinedPDF, A] = NormalizationFactors(x, h)
maxIter = 5e4; % Avoid infinite loop
iterTol = 1e-3; % Stop when error reaches this value

% Initialization
dx = x(2) - x(1); % Assume uniform bin width
[nbins, nbias] = size(h);
p = zeros(size(h));
I = zeros(size(h));
N = zeros(nbias, 1);
A = zeros(nbias, 1);
for n = 1:nbias
    p(:,n) = h(:,n) ./ sum(h(:,n)) ./ dx; % PDF estimate
    I(:,n) = h(:,n) > 0; % One if histogram bin > 0
    k{n} = find(I(:,n)); % Indices for histogram bin > 0
    N(n) = sum(h(:,n)); % Number of histogram samples
end

% Initialize the normalization factors
% by minimizing the vertical offset between
% adjacent PDF estimates

```

```

A(1) = 1;
for n = 2:nbias
    A(n) = A(n-1) .* sum(p(k{n}, n-1) .* p(k{n}, n)) ./ ...
        sum(p(k{n}, n) .* p(k{n}, n));
end

% Iterative solution for normalization factors.
Aprev = A;
for iterIx = 1:maxIter
    % Combined estimate of the PDF
    combinedPDF = sum(p*diag(N), 2) ./ sum(I*diag(N./A), 2);

    % Integrate combined PDF over n:th biasing region
    for n = 1:nbias; A(n) = sum( combinedPDF(k{n}) ); end;
    A = A ./ sum(combinedPDF); % Ensures PDF sums to one

    % Check for convergence
    if (max(abs((A-Aprev)./Aprev)*100) < iterTol); break; end;
    Aprev = A;
end
return

```

C.7 General polarization functionality

```

%—
% binomial:
% Input:
% n, k = Positive integers satisfying 0 ≤ k ≤ n
% Output:
% Binomial coefficient  $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{k} \cdot \frac{n-1}{k-1} \dots \frac{n-k+1}{1}$ 
%—
function [y] = binomial(n,k)
if (k <= 0 | k > n) y = 1;
else y = prod( ((n-k+1):n) ./ (1:k) );
end
return

%—
% JonesRotation:

```

```

% Input:
%   rotAngle =  $\psi$  = Scalar rotation angle.
%   rotAxis =  $\hat{r}$  =  $3 \times 1$  Stokes space unit vector.
% Output:
%    $U = \mathbf{U} = \exp[-i\psi(\hat{r} \cdot \vec{\sigma})/2] = \cos\left(\frac{\psi}{2}\right)\sigma_0 - i\sin\left(\frac{\psi}{2}\right)(\hat{r} \cdot \vec{\sigma})$ 
%—
function [U] = JonesRotation(rotAngle, rotAxis)
rotAxis = rotAxis ./ sqrt(transpose(rotAxis)*rotAxis); % Ensure  $|\hat{r}| = 1$ 
U = cos(0.5*rotAngle)*eye(2,2) ...
    - i*sin(0.5*rotAngle)*VecDotPauli(rotAxis);
return

%—
% JonesToMueller:
% Input:
%   T =  $\mathbf{T}$  = Arbitrary  $2 \times 2$  complex Jones matrix.
% Output:
%   M =  $\mathbf{M} \equiv \mathbf{A}(\mathbf{T} \otimes \mathbf{T}^*)\mathbf{A}^\dagger = 4 \times 4$  real-valued Mueller matrix.
%—
function [M] = JonesToMueller(T)
A = 1/sqrt(2) * [1 0 0 1; 1 0 0 -1; 0 1 1 0; 0 i -i 0];
M = A * kron(T, conj(T)) * ctranspose(A);
return

%—
% JonesToStokes:
% Input:
%   jv =  $|s\rangle = 2 \times 1$  complex Jones vector.
% Output:
%   sv =  $\vec{s} \equiv \langle s | \vec{\sigma} | s \rangle = 3 \times 1$  real Stokes vector
%—
function [sv] = JonesToStokes(jv)
sv = zeros(3,1);
jvt = ctranspose(jv);
sv(1) = jvt * [1, 0; 0, -1] * jv;
sv(2) = jvt * [0, 1; 1, 0] * jv;
sv(3) = jvt * [0, -i; i, 0] * jv;
return

```

```

%—
% MatrixCross:
% Input:
%   svec =  $\vec{s} = 3 \times 1$  Stokes vector.
% Output:
%   scross =  $(\vec{s} \times) = 3 \times 3$  Stokes matrix
%—
function [scross] = MatrixCross(svec)
scross = [0, -svec(3), svec(2); ...
          svec(3), 0, -svec(1); ...
          -svec(2), svec(1), 0];
return

%—
% MinkowskiPolarDecomp:
% Input:
%   M =  $\mathbf{M} =$  Arbitrary  $4 \times 4$  Mueller matrix
% Output:
%   H =  $\mathbf{H} = 4 \times 4$  symmetric Mueller matrix. Can be complex.
%   U =  $\mathbf{U} = 4 \times 4$  orthogonal Mueller matrix. Can be complex.
% Algorithm:
%   Decomposes  $\mathbf{M}$  according to  $\mathbf{M} \equiv \mathbf{H}\mathbf{U}$ 
%   such that  $\mathbf{H} = \mathbf{g}\mathbf{H}^T\mathbf{g}$  and  $\mathbf{U}^T\mathbf{g}\mathbf{U} = \mathbf{g}$ 
%   for  $\mathbf{g} = \text{diag}(-1, 1, 1, 1)$  the Minkowski metric.
%—
function [H,U] = MinkowskiPolarDecomp(M)
g = diag([-1, 1, 1, 1]);
H = (M*g*transpose(M)*g)^0.5;
U = inv(H)*M;
return

%—
% PauliDecomp:
% Input:
%   T =  $\mathbf{T} =$  Arbitrary  $2 \times 2$  complex Jones matrix
% Output
%   t0 =  $t_0 = \text{Tr}(\mathbf{T})/2$ 
%   tvec =  $\vec{t} \equiv (t_1, t_2, t_3)$  where  $t_i = \text{Tr}(\sigma_i \mathbf{T})/2$ . Complex  $3 \times 1$  vector
% Algorithm:
%   Decomposes  $\mathbf{T}$  into  $t_0$  and  $\vec{t} = (t_1, t_2, t_3)$ 

```

```

%    satisfying  $\mathbf{T} \equiv t_0 \boldsymbol{\sigma}_0 + \vec{t} \cdot \vec{\boldsymbol{\sigma}} = t_0 \boldsymbol{\sigma}_0 + t_1 \boldsymbol{\sigma}_1 + t_2 \boldsymbol{\sigma}_2 + t_3 \boldsymbol{\sigma}_3$ .
%    This decomposition exists for all  $2 \times 2$  complex matrices.
%—
function [t0, tvec] = PauliDecomp(T)
tvec = zeros(3,1);
t0    = 0.5 * (T(1,1) + T(2,2));
tvec(1) = 0.5 * (T(1,1) - T(2,2));
tvec(2) = 0.5 * (T(1,2) + T(2,1));
tvec(3) = 0.5 * i * (T(1,2) - T(2,1));
return

%—
% PolarDecomp:
% Input:
%  $T = \mathbf{T} = N \times N$  complex-valued matrix.
% Output:
%  $H = \mathbf{H} =$  Hermitian  $N \times N$  complex-valued matrix.
%  $U = \mathbf{U} =$  Unitary  $N \times N$  complex-valued matrix.
% Algorithm:
% Decomposes  $\mathbf{T}$  according to  $\mathbf{T} \equiv \mathbf{H}\mathbf{U}$ 
% such that  $\mathbf{H} = \mathbf{H}^\dagger$  and  $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}_N$ 
% for  $\mathbf{I}_N$  the  $N \times N$  identity matrix.
%—
function [H,U] = PolarDecomp(T)
H = (T*ctranspose(T))^0.5;
U = inv(H) * T;
return

%—
% RandomUnitVector:
% Input:
%  $N =$  Number of random unit vectors to calculate
% Output:
%  $\text{unitvecs} = 3 \times N$  real-valued matrix. Each column of "unitvecs"
% is a unit-vector uniformly distributed on the Poincare sphere.
% Note:
% This routine generates the rotation axes required
% in the functions EmulateJonesMatrix and EmulatePMDVector.
%—
function [unitvecs] = RandomUnitVector(N)

```

```

z = 2 .* rand(1, N) - 1; % Uniform z ∈ [-1,1]
t = 2 .* pi .* rand(1, N); % Uniform angle t ∈ [0,2π]
r = sqrt(1 - z.*z);
unitvecs = [r.*cos(t); r.*sin(t); z];
return

%—
% StokesRotation:
% Input:
%   rotAngle = ψ = Scalar rotation angle. Real valued.
%   rotAxis =  $\hat{r}$  = 3 × 1 Stokes space unit vector. Real valued.
% Output:
%   R =  $\mathbf{R}$  = exp[ψ $\hat{n}$ ×] = cos ψ $\mathbf{I}$  + (1 - cos ψ) $\hat{n}\hat{n}^T$  + sin ψ( $\hat{n}$ ×)
%—
function [R] = StokesRotation(rotAngle, rotAxis)
c = cos(rotAngle);
R = c * eye(3,3) + ...
    (1-c) * rotAxis * transpose(rotAxis) + ...
    sin(rotAngle) * MatrixCross(rotAxis);
return

%—
% StokesToJones:
% Input:
%   sv =  $\vec{s}$  = 3 × 1 real Stokes UNIT vector
% Output:
%   jv = |s⟩ = 2 × 1 complex Jones vector.
% Algorithm:
%   Solves the defining eigenvalue equation ( $\vec{s} \cdot \vec{\sigma}$ )|s⟩ = |s⟩
%   for the Jones vector |s⟩. As implemented this
%   function assumes | $\vec{s}$ | = 1.
%—
function [jv] = StokesToJones(sv)
sv = sv ./ sqrt(dot(sv,sv)); % Ensure | $\vec{s}$ | = 1
if (sv(1) == 1) % Handles the case where  $\vec{s} = (1,0,0)$ 
    jv = [1; 0];
else
    jv = [1; (1-sv(1))./(sv(2)-i*sv(3))];
    jv = jv ./ sqrt(ctranspose(jv)*jv);
end

```

```

return

%—
% UnitaryDecomp:
% Input:
%  $U = \mathbf{U} = 2 \times 2$  UNITARY Jones matrix
% Output:
%  $rotAngle = \psi =$  real-valued scalar rotation angle
%  $rotAxis = \hat{r} =$  real-valued  $3 \times 1$  unit vector.
% Algorithm:
% Decomposes the unitary matrix  $\mathbf{U}$  into a real-valued
% rotation angle,  $\psi$ , and rotation axis,  $\hat{r}$ ,
% such that  $\mathbf{U} \equiv \exp[-i\psi(\hat{r} \cdot \vec{\sigma})/2]$ .
% As implemented, this function assumes  $\mathbf{U}\mathbf{U}^\dagger \equiv \sigma_0$ 
%—
function [rotAngle, rotAxis] = UnitaryDecomp(U)
[rotAngle, rotAxis] = PauliDecomp(U);
rotAngle = -real(2.0*acos(rotAngle));
rotAxis = real(-i .* rotAxis ./ sqrt(dot(rotAxis, rotAxis)));
return

%—
% VecDotPauli:
% Input:
%  $svec = \vec{s} = 3 \times 1$  Stokes space vector.
% Output:
%  $U = \mathbf{U} = \vec{s} \cdot \vec{\sigma}$ 
%—
function [U] = VecDotPauli(svec)
U = [svec(1), svec(2) - i*svec(3); ...
     svec(2) + i*svec(3), -svec(1)];
return

```

Bibliography

- [1] NIST/SEMATECH e-Handbook of Statistical Methods. Online, 2006. <http://www.itl.nist.gov/div898/handbook/>.
- [2] A. Acharya. *Free energy differences: Representations, estimators and sampling strategies*. PhD thesis, University of Edinburgh, 2004.
- [3] G. Agrawal. *Fiber-Optic communications systems*. John Wiley & Sons, 605 Third Avenue, New York, second edition, 1997.
- [4] A. Aiello and J.P. Woerdman. Linear algebra for Mueller calculus, 2004. arXiv:math-ph/0412061.
- [5] N. Alves, B. Berg, and R. Villanova. Ising-model Monte Carlo simulations: Density of states and mass gap. *Physical Review B*, 41(1):383–394, 1990.
- [6] C. Antonelli, A. Mecozzi, K. Cornick, M. Brodsky, and M. Boroditsky. Pmd-induced penalty statistics in fiber links. *Photon. Technol. Lett.*, 17(5):1013–1015, 2005.
- [7] S. Balasubramanian. Time-development operator method in quantum mechanics. *American Journal of Physics*, 69(4):508–511, 2001.
- [8] R. Barakat. Exponential versions of the Jones and Mueller-Jones polarization matrices. *J. Opt. Soc. Amer. A*, 13(1):158–163, 1996.
- [9] W. Baylis, editor. *Clifford (geometric) algebras, with applications in physics, mathematics, and engineering*. Birkhauser, 1996.

- [10] W. E. Baylis. Special relativity with 2×2 matrices. *American Journal of Physics*, 48(11):918–925, 1980.
- [11] W. E. Baylis, J. Bonenfant, J. Derbyshire, and J. Huschilt. Light polarization: A geometric-algebra approach. *American Journal of Physics*, 61(6):534–545, 1993.
- [12] C. Bennett. Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics*, 22:245–268, 1976.
- [13] B. Berg. Multicanonical Recursions. *Journal of Statistical Physics*, 82:323–342, 1996.
- [14] B. Berg. Algorithmic aspects of multicanonical simulations. *Nuclear Physics B*, 63A-C:982–984, 1998.
- [15] B. Berg. Introduction to Multicanonical Monte Carlo simulations. *Fields Institute Communications*, 26(1):1–24, 2000.
- [16] B. Berg and T. Neuhaus. Multicanonical ensemble: A new approach to simulate first-order phase transitions. *Physical Review Letters*, 68(1):9–12, 1992.
- [17] N. Bergano, C. Poole, and R. Wagner. Investigation of polarization dispersion in long lengths of single-mode fiber using multilongitudinal mode lasers. *J. Lightwave Technol.*, LT-5(11):1618–1622, November 1987.
- [18] M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, The Edinburgh Building, Cambridge, CB2 2RU, UK, seventh edition, 1999.
- [19] M. Boroditsky, K. Cornick, C. Antonelli, M. Brodsky, and S. Dods. Comparison of system penalties from first- and multiorde polarization mode dispersion. *Photon. Technol. Lett.*, 17(8):1650–1652, 2005.
- [20] H. Chen, R. Jopson, and H. Kogelnik. On the bandwidth of higher-order polarization-mode dispersion: the taylor series expansion. *Optics Express*, 11(11):1270–1282, 2003.

- [21] L. Chen, O. Chen, S. Hadjifaradji, and X. Bao. Polarization-mode dispersion measurement in a system with polarization dependent loss or gain. *Photon. Technol. Lett.*, 16(1):206–208, 2004.
- [22] R. Chipman and R. Kinnera. High-order polarization mode dispersion emulator. *Optical Engineering*, 41(5):932–937, 2002.
- [23] E. Collett. *Polarized light in fiber optics*, chapter 14, pages 379–431. The PolaWave Group, Lincroft, New Jersey, 2003.
- [24] E. Dam, M. Koch, and M. Lillholm. Quaternions, interpolation and animation. Technical Report DIKU-TR-98/5, University of Copenhagen, 1998. (available online).
- [25] G. Dattoli, J. Gallardo, and A. Torre. An algebraic view to the operatorial ordering and its applications to optics. *Rivista Del Nuovo Cimento*, 11(11):1–79, 1988.
- [26] A. Djupsjobacka. On differential group-delay statistics for polarization-mode dispersion emulators. *J. Lightwave Technol.*, 19(2):285–290, 2001.
- [27] C. Doran and A. Lasenby. *Geometric algebra for physicists*. Cambridge University Press, 2003.
- [28] L. Dorst. First order error propagation of the procrustes method for 3d attitude estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):221–229, 2005.
- [29] W. Eickhoff, Y. Yen, and R. Ulrich. Wavelength dependence of birefringence in single-mode fiber. *Applied Optics*, 20(19):3428–3435, October 1981.
- [30] G. Einarsson. *Principles of Lightwave Communications*. John Wiley & Sons, 605 Third Avenue, New York, first edition, 1996.
- [31] A. Eyal, W. K. Marshall, M. Tur, and A. Yariv. Representation of second-order polarization mode dispersion. *Electronics Letters*, 41(5):1658–1659, 1999.

- [32] A. Eyal and M. Tur. A modified Poincare sphere technique for the determination of polarization-mode dispersion in the presence of differential gain/loss, ThR1. In *OFC Technical Digest*, page 340. OFC/NFOEC, 1998.
- [33] A. Ferrenberg, D. Landau, and R. Swendsen. Statistical errors in histogram reweighting. *Physical Review E*, 51(2):5092–5100, 1995.
- [34] A. Ferrenberg and R. Swendsen. Optimized Monte Carlo data analysis. *Physical Review Letters*, 63(12):1195–1198, 1989.
- [35] G. Foschini, R. Jopson, L. Nelson, and H. Kogelnik. The statistics of pmd-induced chromatic fiber dispersion. *J. Lightwave Technol.*, 17(9):1560–1565, 1999.
- [36] G. Foschini, L. Nelson, R. Jopson, and H. Kogelnik. Statistics of second-order pmd depolarization. *J. Lightwave Technol.*, 19(12):1882–1886, 2001.
- [37] G. Foschini and C. Poole. Statistical theory of polarization dispersion in single mode fibers. *J. Lightwave Technol.*, 9(11):1439–1456, November 1991.
- [38] N. Frigo. A generalized geometrical representation of coupled mode theory. *IEEE Journal of Quantum Electronics*, QE-22(11):2131–2140, 1986.
- [39] N. Gisin and J. Von der Weid and J. Pellaux. Polarization mode dispersion of short and long single-mode fibers. *J. Lightwave Technol.*, 9(7):821–827, July 1991.
- [40] N. Gisin and B. Huttner. Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers. *Opt. Commun.*, 142(1–3):119–125, October 1997.
- [41] N. Gisin and J. Pellaux. Polarization mode dispersion: time versus frequency domains. *Opt. Commun.*, 89(2-4):316–323, May 1992.
- [42] M. Glasner, D. Yevick, and T. Fysik. Generalized propagation techniques for longitudinally varying refractive index distributions. *Mathematical Computer Modeling*, 16(10):177–182, 1992.

- [43] M. Glasner, D. Yevick, and B. Hermansson. Generalized propagation formulas of arbitrarily high order. *Journal of Chemical Physics*, 95(11):8266–8272, 1991.
- [44] M. Glasner, D. Yevick, and B. Hermansson. High-order generalized propagation techniques. *J. Opt. Soc. Amer. B*, 8(2):413–415, 1991.
- [45] J. Gordon and H. Kogelnik. PMD fundamentals: Polarization mode dispersion in optical fibers. *Proceedings of the National Academy of Sciences*, 97(9):4541–4550, 2000.
- [46] D. Gottlieb. Skew symmetric bundle maps on space-time. *Contemporary Mathematics*, 220(11):117–141, 1998.
- [47] B. Heffner. Automated measurement of polarization mode dispersion using Jones matrix eigenanalysis. *Photon. Technol. Lett.*, 4(9):1066–1068, 1992.
- [48] F. Heismann. Accurate Jones matrix expansion for all orders of polarization mode dispersion. *Optics Letters*, 28(21):2013–2015, 2003.
- [49] D. Hestenes and G. Sobczyk. *Clifford algebra to geometric calculus*. D. Reidel Publishing Company, 3300 AA Dordrecht, Holland, 1984.
- [50] B. Huttner, C. Geiser, and N. Gisin. Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses. *J. Sel. Top. Quant. Electron.*, 6(2):317–329, 2000.
- [51] B. Huttner and N. Gisin. Anomalous pulse spreading in birefringent optical fibers with polarization-dependent losses. *Optics Letters*, 22(8):504–506, 1997.
- [52] R. Jopson, L. Nelson, and H. Kogelnik. Measurement of second-order polarization-mode dispersion vectors in optical fibers. *Photon. Technol. Lett.*, 11(9):1553–1555, 1999.
- [53] I. Kaminow. Polarization in optical fibers. *IEEE Journal of Quantum Electronics*, QE-17(1):15–22, 1981.

- [54] M. Karlsson. Probability density functions of the differential group delay in optical fiber communication Systems. *J. Lightwave Technol.*, 19(3):324–331, 2001.
- [55] M. Karlsson and M. Petersson. Quaternion approach to PMD and PDL phenomena in optical fiber systems. *J. Lightwave Technol.*, 22(4):1137–1146, 2004.
- [56] Myoung-Jun Kim, Myung-Soo Kim, and Sung Yong Shin. A general construction scheme for unit quaternion curves with simple high order derivatives. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 369–376, New York, NY, USA, 1995. ACM Press.
- [57] U. Kintzel. *Polar decompositions and Procrustes problems in finite dimensional indefinite scalar product spaces*. PhD thesis, Technical University of Berlin, 2005.
- [58] U. Kintzel. Procrustes problems in finite dimensional indefinite scalar product spaces. *Linear Algebra and its Applications*, 402:1–28, 2005.
- [59] H. Kogelnik, L. E. Nelson, and J. P. Gordon. Emulation and inversion of polarization-mode dispersion. *J. Lightwave Technol.*, 21(2):482–495, 2003.
- [60] T. Kudou, M. Iguchi, M. Masuda, and T. Ozeki. Theoretical basis of polarization mode dispersion equalization up to the second order. *J. Lightwave Technol.*, 18(4):614–617, 2000.
- [61] S. Kumar, D. Bouzids, R. Swendsen, P. Kollman, and J. Rosebberg. The weighted histogram analysis method for free-energy calculations on biomolecules. I. the method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992.
- [62] Y. Li and A. Yariv. Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses. *J. Opt. Soc. Amer. B*, 17(11):1821–1827, 2000.
- [63] J. Love, R. Sammut, and A. Snyder. Birefringence in elliptically deformed optical fibers. *Electron. Lett.*, 15(20):615–616, 1979.

- [64] P. Lu, L. Chen, and X. Bao. Statistical distributino of polarization-dependent loss in the presence of polarization-mode dispersion in single-mode fibers. *Photon. Technol. Lett.*, 13(5):451–453, May 2001.
- [65] S. Lu and R. Chipman. Mueller matrices and the degree of polarization. *Optics Communications*, 146:11–14, 1998.
- [66] T. Lu. *The application of boundary element and multicanonical methods in optical communications*. PhD thesis, University of Waterloo, December 2005.
- [67] T. Lu, W. Huang, D. Yevick, M. O’Sullivan, and M. Reimer. Multicanonical comparison of polarization-mode dispersion compensator performance. *J. Opt. Soc. Amer. A*, 22(12):2804–2809, 2005.
- [68] T. Lu and D. Yevick. Biased Multicanonical Sampling. *Photon. Technol. Lett.*, 17(7):1420–1422, 2005.
- [69] T. Lu, D. Yevick, B. Hamilton, D. Dumas, and M. Reimer. An Experimental Realization of Biased Multicanonical Sampling. *Photon. Technol. Lett.*, 17(12):1583–2585, 2005.
- [70] T. Lu, D. Yevick, L. Yan, B. Zhang, and A. E. Willner. An Experimental Approach to Multicanonical Sampling. *Photon. Technol. Lett.*, 16(8):1978–1980, 2004.
- [71] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications in Pure and Applied Mathematics*, VII:649–673, 1954.
- [72] A. Mecozzi and M. Shtaif. The statistics of polarization-dependent loss in optical communication systems. *Photon. Technol. Lett.*, 14(3):313–315, March 2002.
- [73] N. Ohta and K. Kanatani. Optimal estimation of three-dimensional rotation and reliability evaluation. *IEICE Transactions on Information and Systems*, E81D(11):1247–1252, 1998.

- [74] A. Orlandini and L. Vincetti. A simple and useful model for Jones matrix to evaluate higher order polarization-mode dispersion effects. *Photon. Technol. Lett.*, 13(11):1176–1178, 2001.
- [75] J. Oteo and J. Ros. From time ordered products to Magnus expansion. *Journal of Mathematical Physics*, 41(5):3268–3277, 2000.
- [76] P. Pechukas and J. Light. On the exponential form of time-displacement operators in quantum mechanics. *J. Chem. Phys.*, 44(10):3897–3912, May 1966.
- [77] P. B. Phua, J. M. Fini, and H. A. Haus. Real-time first- and second-order PMD characterization using averaged state-of-polarization of filtered signal and polarization scrambling. *J. Lightwave Technol.*, 21(4):982–989, 2003.
- [78] C. Poole. Statistical treatment of polarization dispersion in single-mode fiber. *Optics Letters*, 13(8):687–689, August 1988.
- [79] C. Poole, N. Bergano, R. Wagner, and H. Schulte. Polarization dispersion and principal states in a 147km undersea lightwave cable. *J. Lightwave Technol.*, 6(7):1185–1190, July 1988.
- [80] C. Poole, J. Winters, and J. Nagel. Dynamical equation for polarization dispersion. *Optics Letters*, 16(6):372–374, March 1991.
- [81] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes: The art of scientific computing*. Cambridge Univ. Press, Cambridge, U.K., second edition, 1986.
- [82] S. Rashleigh and R. Ulrich. High birefringence in tension-coiled single-mode fibers. *Optics Letters*, 5(8):354–356, 1980.
- [83] M. Reimer, A. Awadalla, D. Yevick, and T. Lu. Alignment methods for biased multicanonical sampling. *IEEE Transactions on Communications*, 2006. (submitted).

- [84] M. Reimer and D. Yevick. A Clifford algebra analysis of polarization-mode dispersion and polarization-dependent loss. *Photon. Technol. Lett.*, 18:734–736, 2006.
- [85] M. Reimer and D. Yevick. Application of the magnus expansion to polarization-mode dispersion and polarization-dependent loss. In *OWI40, OFC Technical Digest. OFC/NFOEC*, 2006.
- [86] M. Reimer and D. Yevick. Least-squares analysis of the Mueller matrix. *Optics Letters*, 31:2399–2401, 2006.
- [87] M. Reimer and D. Yevick. Least squares procedure for measuring mueller matrices. In *OWI41, OFC Technical Digest. OFC/NFOEC*, 2006.
- [88] M. Reimer and D. Yevick. Mueller matrix description of polarization mode dispersion and polarization dependent loss. *J. Opt. Soc. Amer. A*, 23(6):1503–1508, 2006.
- [89] M. Renardy. Singular value decomposition in Minkowski space. *Linear Algebra and its Applications*, 236:53–58, 1996.
- [90] W. Salzman. Convergence of magnus and magnus-like expansions in the schrodinger representation. *Journal of Chemical Physics*, 85(8):4605–4613, 1986.
- [91] W. Salzman. On the convergence of exponential perturbation theories in the schrodinger representation. *Chemical Physics Letters*, 124(6):531–533, 1986.
- [92] W. Salzman. New criterion for convergence of exponential perturbation theory in the schrodinger representation. *Physical Review A*, 36(10):5074–5076, 1987.
- [93] R. Sammut. Birefringence in slightly elliptical optical fibers. *Electron. Lett.*, 16(19):728–729, 1980.
- [94] D. Sandel, V. Mirvoda, S. Bhandare, F. Wust, and R. Noe. Some enabling techniques for polarization mode dispersion compensation. *J. Lightwave Technol.*, 21(5):1198–1210, 2003.

- [95] S. Savory and F. Payne. Pulse propagation in fibers with polarization-mode dispersion. *J. Lightwave Technol.*, 19(3):350–357, 2001.
- [96] D. Scott. *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, New York, 1992.
- [97] Ken Shoemake. Animating rotation with quaternion curves. In *SIGGRAPH '85: Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, pages 245–254, New York, NY, USA, 1985. ACM Press.
- [98] M. Shtaif and M. Boroditsky. The effect of the frequency dependence of pmd on the performance of optical communications systems. *Photon. Tehcnol. Lett.*, 15(10):1369–1371, 2003.
- [99] A. Simon and R. Ulrich. Evolution of polarization along a single-mode fiber. *Applied Physics Letters*, 31(8):517–520, 1977.
- [100] G. Smith and A. Bruce. A study of the multi-canonical Monte Carlo method. *Journal of Physics A: Mathematical and General*, 28:6623–6643, 1995.
- [101] M. Souaille and B. Roux. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135:40–57, 2001.
- [102] G. Torrie and J. Valleau. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters*, 28(4):578–581, 1974.
- [103] G. Torrie and J. Valleau. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23:187–199, 1977.
- [104] D. Tweed. Estimating rigid motions via the conformal model of Euclidean space. In *Proceedings of the 17th International Conference on Pattern Recognition*, volume 2, pages 171–174. IEEE Computer Society, 2004.

- [105] R. Ulrich, S. Rashleigh, and W. Eikhoff. Bending-induced birefringence in single-mode fiber. *Optics Letters*, 5(6):273–275, 1980.
- [106] R. Ulrich and A. Simon. Polarization optics of twisted single-mode fibers. *Applied Optics*, 18(13):2241–2251, 1979.
- [107] E. Veach. *Robust monte carlo methods for light transport simulation*. PhD thesis, Stanford University, 1997.
- [108] P. Virnau and M. Muller. Calculation of free energy through successive umbrella sampling. *Journal of Chemical Physics*, 120(23):10925–10930, 2004.
- [109] T. Wanner, B. Marks, C. Menyuk, and J. Zweck. Polarization decorrelation in optical fibers with randomly varying elliptical birefringence. *Opt. Lett.*, 28(19):1799–1801, October 2003.
- [110] E. W. Weisstein. Matrix direct product. From MathWorld-A Wolfram Web Resource. <http://mathworld.wolfram.com/MatrixDirectProduct.html>.
- [111] R. Wilcox. Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics*, 8(4):962–982, 1967.
- [112] D. Yevick. Multicanonical Communication System Modeling-Application to PMD Statistics. *Photon. Technol. Lett.*, 14(11):1512–1514, 2002.
- [113] D. Yevick. The accuracy of Multicanonical system models. *Photon. Technol. Lett.*, 15(2):224–226, 2003.
- [114] D. Yevick. *A first course in computational physics and object-oriented programming with C++*. Cambridge University Press, 2005.
- [115] D. Yevick, M. Chanachowicz, M. Reimer, M. O’Sullivan, W. Huang, and T. Lu. Chebyshev and Taylor approximations of polarization mode dispersion for improved compensation bandwidth. *J. Opt. Soc. Amer. A*, 22(8):1662–1667, 2005.

- [116] D. Yevick and B. Hermansson. Convergence properties of wide-angle techniques. *Photon. Technol. Lett.*, 16(12):1457–1459, 1994.
- [117] D. Yevick and T. Lu. Improved Multicanonical Algorithms. *J. Opt. Soc. Amer. A*, 23(11):2912–2918, November 2006.
- [118] D. Yevick, T. Lu, W. Huang, and W. Bardyszewski. Operator expansions for PMD analysis and compensation. *J. Opt. Soc. Amer. A*, 23(2):455–460, 2006.