

Temporal Dynamics of Polarization and Polarization Mode Dispersion and Influence on Optical Fiber Systems

by

George Soliman

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2013

© George Soliman 2013

AUTHOR'S DECLARATION

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

This thesis examines polarization and polarization mode dispersion (PMD) dynamics in optical fibers as well as the evaluation of probability density functions and bit error rates in a realistic wavelength division multiplexed (WDM) optical communication systems.

In the first part of the thesis, experimental studies of the dynamics of polarization in a dispersion compensation module (DCM) are performed in which mechanical shocks are imparted to several different DCMs by dropping a steel ball on the outer casing at different locations and from different heights and the resulting rapid polarization fluctuations are measured. We provide a theoretical model that accounts for the dynamic birefringence generated due to the impact. Next, an experimental technique is proposed to detect the location of temporal polarization activity in WDM systems. It is demonstrated theoretically and in simulations that measurement of both the PMD vector and the Stokes parameters at the WDM frequencies enables the detection of the location of such activity.

Different linear prediction procedures are applied to the differential group delay of an optical fiber link assumed to obey the hinge model. The hinges are modeled as polarization rotators with fixed rotation axes and sinusoidally varying rotation angles. Three prediction methods are investigated and consequently compared: an autoregressive model (AR) with Kalman filter, a pattern imitation method and a Taylor expansion technique. The effect of measurement noise on the prediction horizon is also investigated for each prediction method.

Using a physically reasonable stochastic model for the hinges, we derive analytical expressions for the temporal autocorrelation functions of the state of polarization (SOP) and the PMD vector. The obtained analytical results are compared to simulations.

Finally, we apply the multicanonical method to the probability density function of received symbols and the symbol error ratio (SER) in a dual polarization quadrature phase shift keyed (DP-QPSK) WDM system. We simulate five co propagating channels at a symbol rate of 10.7 GBaud/s and account for PMD and nonlinear effects. We evaluate the performance of the system for two different cases: single mode fibers with full dispersion compensation at the end of the link, effective large area fibers (LEAF) with full dispersion compensation at the end of the link.

Acknowledgements

I would like to thank Professor David Yevick for all his guidance and support during my studies. This thesis would have not been possible without his insightful comments and his broad knowledge of optical communication. I would also like to thank Maurice O' Sullivan for giving me the opportunity of an internship at Nortel (now Ciena) for two months to pursue part of this research and for providing field data.

I would also like to thank my colleagues Michael Reimer and Ahmed Awadalla for easing my coming into this new field of research, for helping me with some of the experimental work and for the insightful discussions we had about various problems. I would also like to thank Michael Reimer for providing part of the code for the WDM system analysis.

Dedication

To my late Uncle Adel who inspired me to pursue a mathematics-related discipline.

To Monica.

Table of Contents

AUTHOR'S DECLARATION.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Dedication.....	v
Table of Contents.....	vi
Glossary.....	ix
List of Figures.....	x
List of Tables.....	xiv
Chapter 1 Introduction.....	1
1.1 Optical Communication.....	1
1.2 Modulation Formats.....	3
1.3 Polarization Mode Dispersion.....	7
1.3.1 Origin of PMD.....	7
1.3.2 Jones Calculus.....	9
1.3.3 Principal States Model.....	12
1.3.4 Stokes Parameter Description.....	15
1.3.5 PMD vector.....	19
1.3.6 Manakov PMD equation.....	21
1.3.7 Statistics of PMD.....	26
1.3.8 Hinge Model.....	29
1.3.9 PMD Compensation.....	30
1.4 Bit Error Ratio.....	31
1.4.1 BER for ASK System.....	31
1.4.2 SER for QPSK System.....	33
1.5 The Multicanonical Method.....	33
1.6 Organization of the Thesis.....	34
Chapter 2 Measurement and Simulation of Polarization Transients in Dispersion Compensation Modules.....	35
2.1 Introduction.....	35

2.2 Experimental Setup and SOP Reproducibility	35
2.3 Birefringence Modeling.....	39
2.4 Measurement and Simulation Results	45
2.4.1 S Parameters and Angular Velocity.....	45
2.4.2 Variation with Fiber Length and Impact Position	50
2.5 Conclusion.....	52
Chapter 3 Locating Regions of Polarization Activity in WDM Systems	54
3.1 Introduction	54
3.2 Theory	55
3.3 Simulation Parameters.....	58
3.4 Results and Experimental Setup.....	59
3.5 Conclusion.....	63
Chapter 4 Differential Group Delay (DGD) Prediction in Optical Fiber Links	64
4.1 Introduction	64
4.2 Model Parameters.....	64
4.3 Prediction Methods.....	67
4.3.1 AR Modeling and Kalman Filters	67
4.3.2 Pattern Imitation [57]	69
4.3.3 Taylor Expansion.....	70
4.3.4 Filter Description.....	70
4.4 Results	71
4.5 Conclusion.....	78
Chapter 5 Temporal Autocorrelation Functions of the PMD and Stokes Vectors in the Anisotropic Hinge Model.....	79
5.1 Introduction	79
5.2 ACF of the Hinge Matrix	79
5.3 ACF for the PMD Vector	82
5.4 Simulations.....	83
5.5 Conclusion.....	85
Chapter 6 Simulation of a WDM Polarization Multiplexed Quadrature Phase Shift Keyed (PM-QPSK) System.....	86

6.1 Introduction.....	86
6.2 System Description and Parameters.....	86
6.3 Initial Field Temporal Profile	87
6.4 SSFM for Numerical Solutions to the Manakov PMD equation	88
6.5 Simulation Results	90
6.6 Conclusion	93
Chapter 7 Conclusions and Future Work.....	94
Bibliography	96

Glossary

Acronym

Definition

ACF	Autocorrelation function
AR	Auto regressive
ASK	Amplitude shift keying
AV	Angular velocity
BER	Bit error ratio
DCM	Dispersion compensation module
DD	Direct detection
DGD	Differential group delay
DSP	Digital signal processing
EDFA	Erbium doped fiber amplifier
FIR	Finite impulse response
FPGA	Field programmable gate array
FTTH	Fiber to the home
GVD	Group velocity dispersion
IM	Intensity modulation
LEAF	Large effective area fiber
NRZ	Non return to zero
OSNR	Optical signal to noise ratio
OTDR	Optical time domain reflectometry
pdf	Probability density function
PM	Polarization multiplexed
PMD	Polarization mode dispersion
PSK	Phase shift keying
PSP	Principal state of polarization
QPSK	Quadrature phase shift keying
RZ	Return to zero
SER	Symbol error ratio
SMF	Single mode fiber
SOP	State of polarization
SSFM	Split step Fourier method
WDM	Wavelength division multiplexing

List of Figures

Figure 1.1. Schematic of an optical communication system. A transmitter launches a modulated optical carrier into an optical fiber. DCMs and repeaters are periodically located along the line. An optical receiver at the end the system converts the optical signal back into an electrical signal.	1
Figure 1.2. A coherent optical receiver [3].	2
Figure 1.3. Modulated optical carrier in the case of ASK and PSK for a specific digital bit pattern [4].	3
Figure 1.4. Comparison of (a) RZ and (b) NRZ signal for the bit pattern shown [4].	4
Figure 1.5. Ideal constellation diagram of QPSK. Each dot represents the symbol that consists of the specific two bit pattern shown on the graph. Each symbol defines a certain carrier phase which can be constructed from a suitable superposition of $m_1(t)$ and $m_2(t)$	6
Figure 1.6. Measured constellation diagram at the receiver after 6400km of transmission [3]......	6
Figure 1.7. QPSK modulator. The binary bit stream is interleaved and encoded by a RZ encoder to generate the I and Q signals which in turn modulate the carrier and a ninety degree phase shifted version of the same carrier respectively.....	7
Figure 1.8. (a) sources of intrinsic birefringence: geometric due to ellipticity of the core or anisotropic stress and (b) extrinsic birefringence due to imperfections occurring due to cabling and installation. .	8
Figure 1.9. Birefringent fiber section oriented at an angle ψ relative to the reference x-y axes.	9
Figure 1.10. A fiber is modeled as a concatenation of birefringent segments. The orientation of each segment is chosen randomly. The arrows indicate the slow and fast axes orientations of each segment.	11
Figure 1.11. Poincare sphere. A few representative polarization states are indicated [10]......	16
Figure 1.12. The concatenation of two fiber segments and the resulting PMD vector.	21
Figure 1.13. Schematic of the PMD nulling technique. The PC and the birefringent element are dynamically controlled to cancel the output PMD vector at the carrier frequency [15]......	30
Figure 2.1. In our experiments, a steel ball is secured and later released onto a DCM by an electromagnet generating fast polarization changes. The acoustics signal from the impact is detected by a microphone and employed to trigger the scope, which then samples the output from a fast polarimeter.	37

Figure 2.2. Different impact positions on the DCM as follows: points 1-4 are located 1.5 cm, 3 cm, 4.5 cm and 6 cm from the inner core. Points 6 and 8 and 10 and 12 are at the same radial distances as points 2 and 4 but displaced at 45° and 90° angles.	37
Figure 2.3. Ten successive measurements of the s_1 component. The inset displays the region from 0.7 ms to 0.74 ms showing the individual measurement curves.	38
Figure 2.4. The mean squared error of the polarization trace versus drop height for different DCMs for impact location 1. The error is largest for a height of 9cm.	39
Figure 2.5. The geometry of the DCM. Our model assumes that the steel ball impacts a certain area of the DCM resulting in the oscillation of the fiber coils beneath it. Here $R_{in}=5.25$ cm and $R_{out}=12.4$ cm. $H = 4$ cm.	42
Figure 2.6. The AV for DCM10 with impact location 1, and release height 3 cm.	46
Figure 2.7. The AV for DCM20 with impact location 1, and release height 3 cm.	46
Figure 2.8. The AV for DCM40 with impact location 1, and release height 3 cm.	47
Figure 2.9. The AV for DCM100 with impact location 1, and release height 3 cm.	48
Figure 2.10. The simulated AV of fiber B for an active length of 600m, $z_{i1}=100$ m.	49
Figure 2.11. The simulated AV for fiber B, for active length equal to the total fiber length.	49
Figure 2.12. The S parameters for fiber C with an active length of 20 m and $z_{i1}=200$ m. The fluctuations are clearly very small compared to the experimental measurements (figure 2.3).	50
Figure 2.13. The dependence of the arc length on the angular velocity for different lengths of fiber B.	51
Figure 2.14. The dependence of the experimental arc length on the angular velocity for DCM10 with a ball release height of 9 cm.	52
Figure 2.15. The experimental arc length for impact location 1, a release height of 9 cm, and different DCM structures.	52
Figure 3.1. A rapidly varying localized polarization variation modeled as a hinge time varying rotation matrix $H(t)$. The fiber before the hinge is section 1, with PMD vector and rotation matrix $\vec{\tau}_1$ and R_1 while section 2 follows the hinge with PMD vector and rotation matrix $\vec{\tau}_2$ and R_2	55
Figure 3.2. Absolute difference between X from sequences 1 and 2 for the two hinge models used and for 80 and 1000 realizations.	62

Figure 3.3. Proposed experiment. A tunable laser launches light into a concatenation of two fibers connected through a polarization controller acting as a hinge. A power splitter at the output of fiber 2 splits light into PMD and Stokes vector measurement apparatus for simultaneous measurements. The frequency of the laser is then varied and the process is repeated and ensemble averages are formed. The computer is used to control the time variation of the polarization controller thus implementing the hinge in the link. 63

Figure 4.1. Simulated optical link. Five fiber spans S1 to S5 are connected via a series of time varying hinges H1 to H4 modeled as polarization rotators. Each fiber span consists of 50 birefringent sections each having randomly oriented rotation axes. 65

Figure 4.2. Correlation function of simulated DGD time series. The correlation index is defined as the distance between the maximum of the correlation function and the first minimum. 66

Figure 4.3. DGD time series with mean removed. The first 7200 data points are used to estimate the model and the following 7000 points of the series is employed for validation. 67

Figure 4.4. The metric versus prediction horizon for the different methods employed. The legend indicates model numbers as described in section 4.4. (1) an AR(5) model applied to τ (dotted line), (2) an AR(8) model applied to τ^2 (dashed line), (3) a combination of three separate AR(9), AR(9) and AR(8) models applied to τ_x, τ_y and τ_z respectively (circle markers), (4) a seventh-order Taylor expansion for τ_x, τ_y and τ_z (square markers), and (5) the pattern imitation procedure applied to τ (triangle markers). 72

Figure 4.5. Simulated and predicted DGD time series for model (3) for $P = 130$ and $\Phi = 86\%$ 73

Figure 4.6. Metric plotted as a function of the prediction horizon after filtering. 74

Figure 4.7. Simulated DGD and predicted DGD time series for noisy data for AR(10) model applied to the DGD. 75

Figure 4.8. Predicted and measured DGD. 76

Figure 4.9. Predicted and measured DGD for fiber 2. 77

Figure 5.1. Optical fiber link with hinges. 80

Figure 5.2. Averaged autocovariance of $\psi(t)$ for 5000 realizations. 81

Figure 5.3. Normalized ACF for the PMD vector and the SOP as a function of α in days. 84

Figure 5.4. Simulated and analytic ACF's of the PMD vector for three different sets of invariant hinge rotation axes. 85

Figure 6.1. A schematic diagram of the simulated system.	86
Figure 6.2. Real part of A_{ix} versus time in ns.	88
Figure 6.3. Log_{10} of the joint pdf of the xI, xQ components for a $xI = xQ = 1$ symbol launched into a single polarization in the center channel, for four 3×10^5 multicanonical samples per iteration and the Monte Carlo procedure. The smallest pdf value shown is 10^{-11}	91
Figure 6.4. The variation in the SER with OSNR for a SMF for two step sizes.	92
Figure 6.5. As in Figure 6.4 but for a LEAF.	93

List of Tables

Table 3-1. Exact and computed values of X for 80 realizations and two different sequences of random numbers for model 1	60
Table 3-2. Exact and computed values of X for 80 realizations and two different sequences of random numbers for model 2	60
Table 3-3. Exact and computed values of X for 1000 realizations and two different sequences of random numbers for model 1	61
Table 3-4. Exact and computed values of X for 1000 realizations and two different sequences of random numbers for model 2	61
Table 4-1. ψ compared for the three methods applied to simulated and measured data.	77

Chapter 1

Introduction

1.1 Optical Communication

Optical fiber communication has revolutionized global telecommunications since the early 1970's. In particular, the large bandwidth due to small losses and dispersion in optical fibers led to the adoption of fiber systems initially for transoceanic communication and in long haul systems for e.g. internet traffic, but subsequently for transmission over increasingly smaller distances. Currently, fiber to the home (FTTH) is replacing conventional copper cables in carrying internet, telephone, and cable TV services to the homes.

In a digital communication system, a digital electrical signal modulates a high frequency carrier that in turn is transmitted over a suitable medium. For example, electrical systems employ microwave carriers which oscillate at frequencies ranging typically from 1 to 10 GHz and then are transmitted in free space, over coaxial cables or over closed waveguides. Since only a few percent of the carrier can be used as modulation bandwidth, such systems cannot operate at bit rates larger than 0.2 Gb/s. An optical carrier on the other hand possesses a frequency of 100 THz and a few percent of the carrier frequency corresponds to a bit rate of 1Tb/s. To harness the full bandwidth of optical communication however requires a suitable transmission medium which was first provided with the advent of low loss optical fibers. Lightwave systems usually relied on intensity modulation (IM) of the optical wave in an on/off fashion and direct detection (DD). DD means that the detector only responds to the optical power of the signal; phase or polarization information is lost in the detection process.

A typical optical communication system consists of an optical transmitter which in turn consists of a laser diode and a modulator, the optical fiber with repeaters and dispersion compensator modules (DCMs) along the line, and finally an optical receiver as shown in

Figure 1.1. The repeater's function was to regenerate the signal completely in the electronic domain to compensate for losses and dispersion roughly every 20km. Each repeater consisted of a complete receiver-transmitter system, where the optical signal was converted back into the electrical domain in the form of a binary stream that was then used to modulate a laser to regenerate the optical signal that is launched again in the optical fiber. The process of regeneration was expensive and different schemes to avoid its use or increase the distance between the repeaters were being investigated.

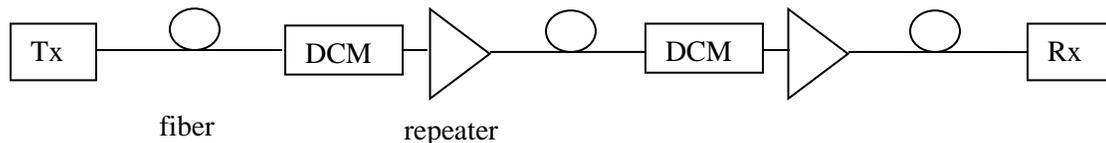


Figure 1.1. Schematic of an optical communication system. A transmitter launches a modulated optical carrier into an optical fiber. DCMs and repeaters are periodically located along the line. An optical receiver at the end the system converts the optical signal back into an electrical signal.

Coherent optical communications was researched extensively in the 1980's as a potential solution to the repeater distance problem [1]. The receiver architecture consisted of a local oscillator, a directional coupler, and a balanced photodiode detector as shown in Figure 1.2 as opposed to a single photodiode in DD. The received optical signal is mixed with a local oscillator signal in the directional coupler. Both outputs of the directional coupler are then detected by a photodiode and the difference between the resulting currents is fed to the electronic circuits. The signal is passed through a polarization controller (PC) before mixing [2].

By controlling the local oscillator power, coherent detection achieved a signal to noise ratio very close to the quantum limit of detection resulting in a higher sensitivity compared to the standard receiver by values ranging from 20 dB to 30 dB [1] thus increasing repeater distance. Further, information on the carrier frequency, phase and polarization can be retained [3], allowing polarization multiplexing and the employment of higher level modulation formats. The disadvantage of this approach was the complexity involved in tracking the carrier frequency, phase and polarization which required difficult to implement optical phase locked loops to allow accurate detection.

Interest in coherent communications faded due to the discovery of Erbium doped fiber amplifiers (EDFA's) allowing amplification of the signal in the optical instead of the electrical domain. Also, the wide bandwidth of 20nm offered by the EDFA along with reliable tunable lasers made wavelength division multiplexing (WDM) transmission possible where several IM/DD channels at different wavelengths are multiplexed together and propagated over the same fiber. For example, the standard Dense WDM (DWDM) consists of 80 channels separated by 50 GHz in the C band which extends from a wavelength of 1.53 μm to 1.57 μm . Consequently, WDM/EDFA systems were widely deployed during the 1990's replacing the need for repeaters.

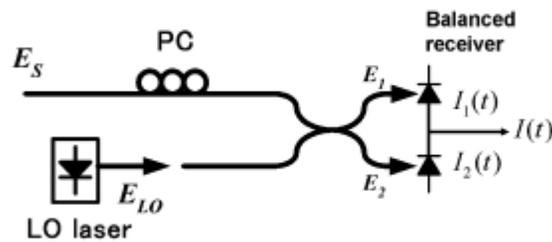


Figure 1.2. A coherent optical receiver [3].

With the ever increasing bandwidth demand in the first decade of the twenty first century, WDM systems' capacity required a boost which was provided by the technology of coherent communications. The ability to detect the frequency or phase of the carrier enables higher level modulation formats which rely on modulating the carrier phase with any number of values and hence increases the transmission bit rate employing the same bandwidth offered by the channel. The problems of carrier frequency, phase and polarization tracking still remained but were solved in the

electronic domain using high speed digital signal processing (DSP) made possible through advances in extremely fast sampling analog to digital converters [3]. Nowadays, DSP is also used to compensate for other propagation impairments such as chromatic dispersion, nonlinearities and polarization mode dispersion (PMD).

A common figure of merit for communication systems is the BL product ; the bit rate, B , the distance, L , product [4]. The distance is defined as the maximum length over which the optical signal can be transmitted without need of a repeater or an amplifier. Over four generations of optical communication systems, BL has increased from 10 Gb/s – km in 1975 to 10^6 Gb/s – km in the early 2000's.

1.2 Modulation Formats

Two types of basic modulation formats are commonly used for digital transmission. Amplitude shift keying (ASK) which involves modulating the intensity (also called IM as explained above) of the optical carrier with a digital one represented by a burst of optical power and a digital zero represented as no optical power. In phase shift keying (PSK), digital ones and zeros are represented with a carrier phase of zero and π respectively. A schematic of the modulation formats is shown in

Figure 1.3.

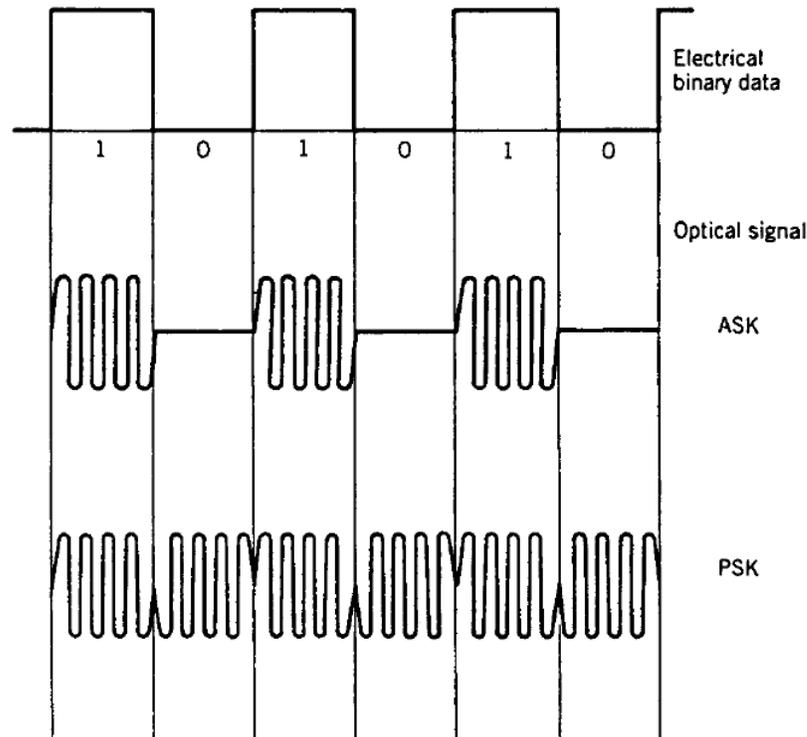


Figure 1.3. Modulated optical carrier in the case of ASK and PSK for a specific digital bit pattern [4].

As can be observed from Figure 1.3, the envelope of the PSK signal is constant while that of ASK is not. However, for a continuous stream of ones, the ASK envelope is also constant. The constant envelope can generate a problem at the receiver which functions by extracting a sine wave operating at the bit rate from the data, named the clock, in order to sample the data and make a decision. Transitions are required in the data stream to extract the clock. We distinguish between two types of transmission, nonreturn to zero (NRZ) and return to zero (RZ) formats.

In the NRZ format, the optical pulse representing the one bit occupies the entire bit slot. On the other hand, in RZ format the optical pulse representing the one bit is chosen to be shorter than the bit slot and its amplitude returns to zero before the bit duration is over [4] as shown in Figure 1.4. In PSK, it is necessary to use RZ format to enable clock extraction.

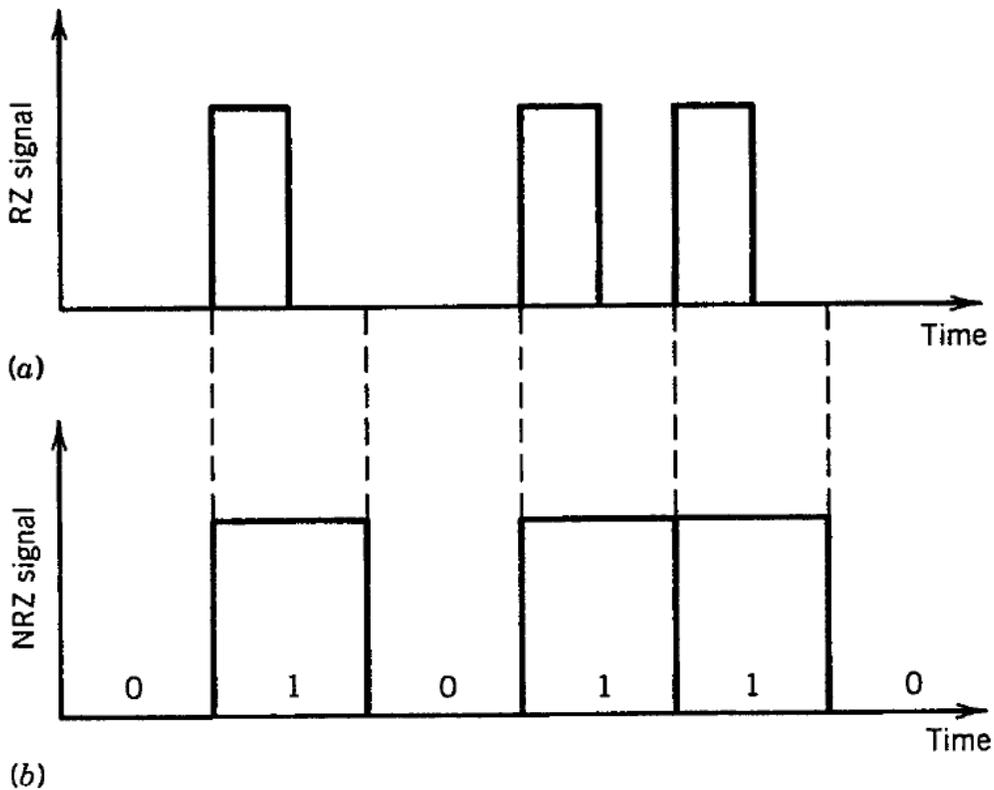


Figure 1.4. Comparison of (a) RZ and (b) NRZ signal for the bit pattern shown [4].

Modulation formats discussed so far transmit one optical symbol per bit. A currently widely employed higher level modulation format is quadrature phase shift keying (QPSK) in which each symbol is represented by two digital bits. Each symbol is in turn represented by a carrier with a phase of $\theta_s = \pm\pi/4, \pm3\pi/4$ depending on the bit combination. We can write for the optical field of the QPSK signal

$$v(t) = a(t)\sin(\omega_s t + \theta_s) \quad (1.1)$$

where $a(t)$ is the pulse shape and ω_s is the carrier angular frequency . Dropping the pulse shape as it is common for all symbols and expanding the sine term we can express (1.1) as

$$\begin{aligned} v(t) &= \frac{1}{\sqrt{2}}\sin(\omega_s t) + \frac{1}{\sqrt{2}}\cos(\omega_s t), \theta_s = \pi/4 \\ v(t) &= -\frac{1}{\sqrt{2}}\sin(\omega_s t) + \frac{1}{\sqrt{2}}\cos(\omega_s t), \theta_s = 3\pi/4 \\ v(t) &= \frac{1}{\sqrt{2}}\sin(\omega_s t) - \frac{1}{\sqrt{2}}\cos(\omega_s t), \theta_s = -\pi/4 \\ v(t) &= -\frac{1}{\sqrt{2}}\sin(\omega_s t) - \frac{1}{\sqrt{2}}\cos(\omega_s t), \theta_s = -3\pi/4 \end{aligned} \quad (1.2)$$

Eq. (1.2) can be represented in a graphical form called the constellation diagram. Writing $m_1(t) = \sin(\omega_s t)/\sqrt{2}$ and $m_2(t) = \cos(\omega_s t)/\sqrt{2}$ and plotting them as x and y axes respectively we obtain the constellation diagram displayed in Figure 1.5. Such a constellation diagram can be obtained in practice after the receiver by sampling the resulting digital sequence and plotting the samples on the $m_1 - m_2$ axes. If the transmission is perfect, all samples will align and we get four dots corresponding to the four different symbols. However, transmission impairments such as dispersion and nonlinearity cause the sampled sequence to be spread out over an area in a scatter plot. A more realistic constellation diagram is shown in Figure 1.6 [3].

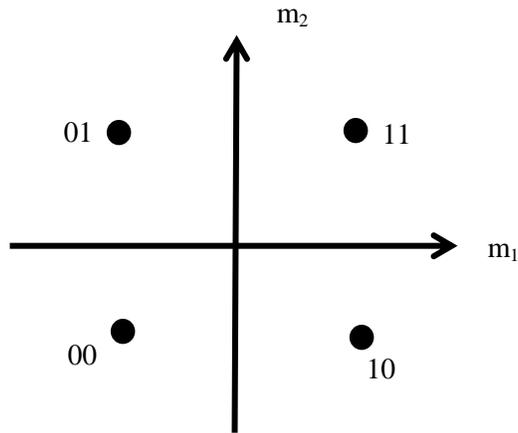


Figure 1.5. Ideal constellation diagram of QPSK. Each dot represents the symbol that consists of the specific two bit pattern shown on the graph. Each symbol defines a certain carrier phase which can be constructed from a suitable superposition of $m_1(t)$ and $m_2(t)$.

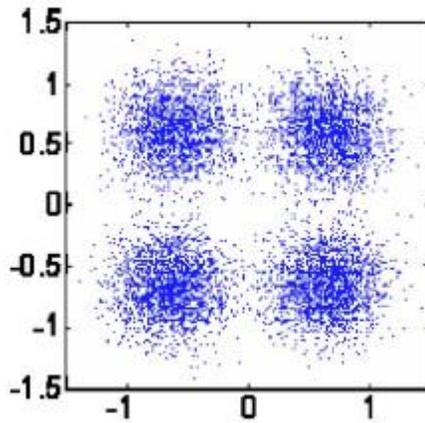


Figure 1.6. Measured constellation diagram at the receiver after 6400km of transmission [3].

Eq. (1.2) also suggests a method for modulating the carrier in QPSK format. The bit stream is interleaved resulting in two sequences called I and Q. A one bit is represented by amplitude of 1 and a zero bit by amplitude of -1. Each stream then modulates $\sin(\omega_s t)$ and $\cos(\omega_s t)$ independently where $\cos(\omega_s t)$ is obtained from $\sin(\omega_s t)$ through a $\pi/2$ phase shift. Finally both modulated signals are multiplexed to form the transmit signal. A schematic diagram of a QPSK modulator is shown in Figure 1.7.

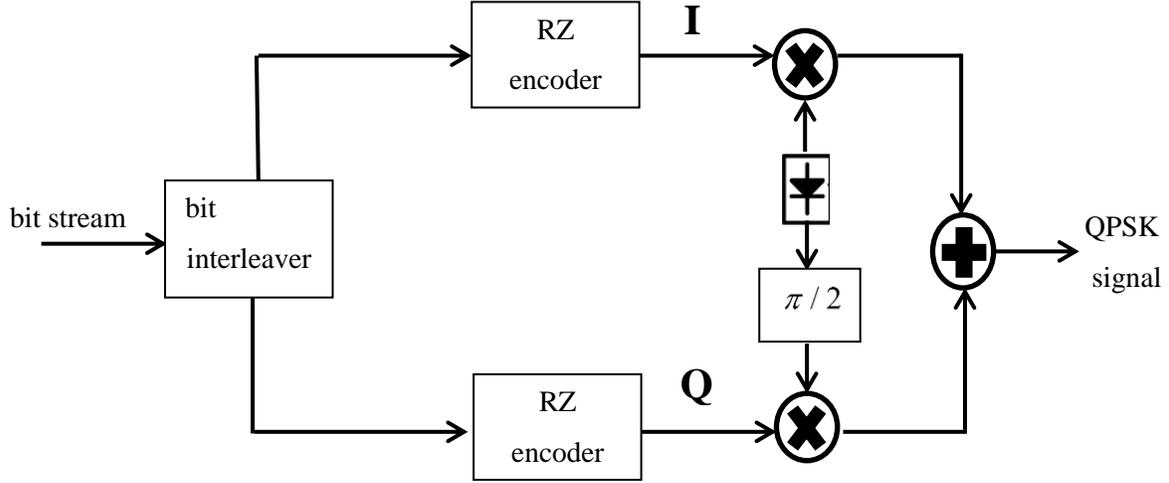


Figure 1.7. QPSK modulator. The binary bit stream is interleaved and encoded by a RZ encoder to generate the I and Q signals which in turn modulate the carrier and a ninety degree phase shifted version of the same carrier respectively.

State of the art optical communication systems use polarization multiplexed (PM) - QPSK format to further increase the effective bit rate by a factor of 2 over just simple QPSK. Each polarization carries a different data stream and each symbol carries two bits effectively quadrupling the bit rate compared to an IM/DD system thus achieving high spectral efficiency [5]. Further, PM-QPSK coding has been shown to yield excellent optical noise tolerance with modest implementation complexity [6] and has thus been widely employed in 40Gb/s transmission [7].

Optical fiber impairments restrict further increase in data rates. One of the most important factors that limit communication at bit rates beyond 10 Gb/s per WDM channel is PMD. This impairment in particular is not constant and drifts randomly with time and hence in general requires a statistical treatment. In the rest of this chapter, we explain the origin of this impairment and describe the theory and consequences.

1.3 Polarization Mode Dispersion

1.3.1 Origin of PMD

PMD has its origin in fiber birefringence. A weakly guiding single mode fiber supports two orthogonally polarized modes, termed LP₀₁ (linearly polarized) [8]. The electric fields \vec{E} of these modes can be expressed as

$$\vec{E} = \hat{x}E_x F(x, y)e^{j(\omega t - \beta_x z)} \quad (1.3)$$

$$\vec{E} = \hat{y}E_y F(x, y)e^{j(\omega t - \beta_y z)} \quad (1.4)$$

where E_x, E_y are the complex field amplitudes, $F(x, y)$ is the transverse field distribution, ω is the angular frequency, β_x, β_y are the propagation constants and z is the axial distance along the fiber. When light of arbitrary polarization is launched into the fiber, the power is divided between both modes. In the ideal fiber, $\beta_x = \beta_y$ because of the perfect circular symmetry of the fiber cross section and hence the two modes are degenerate thus they propagate with the same velocity and no distortion occurs. In a real fiber, however, circular symmetry is broken either during manufacture due to the ellipticity of the core or cladding regions and anisotropic internal stress; or due to fiber cabling that induces bends and twists also introducing stress. Consequently, the modes are no longer degenerate hence light of arbitrary polarization launched at the input will arrive distorted at the output in what is called PMD. Writing $\beta_{x,y} = k_0 n_{effx,y}$ with $k_0 = 2\pi/\lambda_0$, λ_0 being the free space wavelength and n_{eff} being the effective refractive index of the mode, we define the birefringence of the fiber as $\Delta n_{eff} = n_{effx} - n_{effy}$. Figure 1.8 displays the different mechanisms that lead to birefringence. They are grouped under intrinsic and extrinsic birefringence. Intrinsic birefringence occurs due to the imperfections in the manufacturing process while extrinsic birefringence occurs during cabling.

The perturbations illustrated in Figure 1.8 vary along the fiber length randomly. The reference x-y axes employed in (1.3) and (1.4) will not coincide with the true polarization directions of the fundamental modes and the birefringence might also change along the axis of the fiber. Hence, for a treatment of PMD in lightwave systems, the fiber has to be represented as a concatenation of birefringent sections each having different orientation and birefringence. In the next section we introduce Jones calculus [9] which is the mathematical tool that is used to describe propagation in birefringent systems.

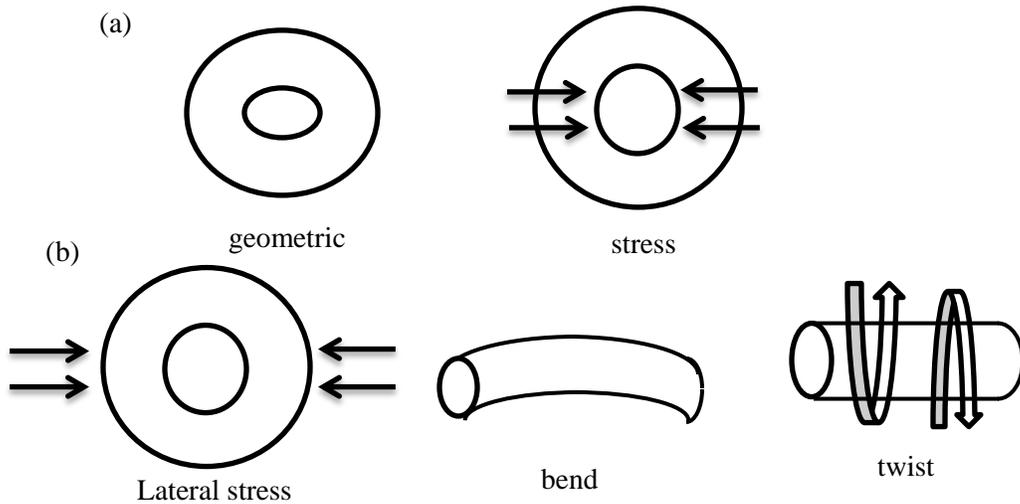


Figure 1.8. (a) sources of intrinsic birefringence: geometric due to ellipticity of the core or anisotropic stress and (b) extrinsic birefringence due to imperfections occurring due to cabling and installation.

1.3.2 Jones Calculus

Dropping the transverse field distribution, the time dependence, setting $z = 0$ in Eq's. (1.3) and (1.4) and grouping both electric field components into a column vector, we describe the polarization of the field by the Jones vector [10]

$$\mathbf{E} = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} A_x e^{j\delta_x} \\ A_y e^{j\delta_y} \end{pmatrix} \quad (1.5)$$

where A_x and A_y are the x and y amplitudes of the fields, δ_x and δ_y are the phases. It is customary to normalize the Jones vector such that $\mathbf{E}^\dagger \mathbf{E} = 1$, where \mathbf{E}^\dagger is the Hermitian conjugate of \mathbf{E} . This is equivalent to setting the intensity of the light beam to unity. To describe the propagation of the Jones vector in a fiber, we consider propagation through a birefringent fiber section with LP modes oriented at an angle ψ to the reference x-y axes as shown in Figure 1.9. The section axes are called slow “s” and fast “f” depending on the values of the propagation constants of the modes. The polarization of these modes can be described by the following orthogonal Jones vectors

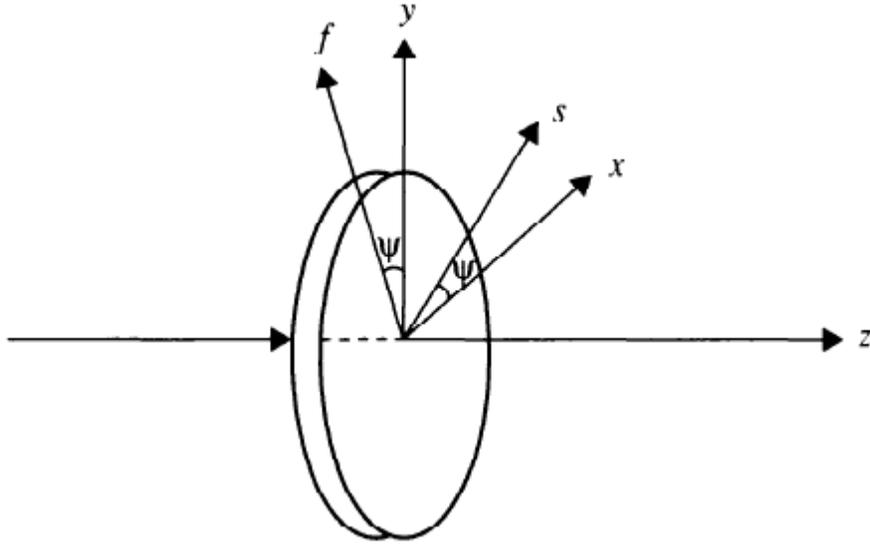


Figure 1.9. Birefringent fiber section oriented at an angle ψ relative to the reference x-y axes.

$$\begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix} \text{ and } \begin{pmatrix} -\sin \psi \\ \cos \psi \end{pmatrix} \quad (1.6)$$

Decomposing the incident light \mathbf{E}_{in} into a linear combination of the $s-f$ modes is achieved through the coordinate transformation

$$\begin{pmatrix} \mathbf{E}_s \\ \mathbf{E}_f \end{pmatrix} = \xi(\psi) \mathbf{E}_{in} = \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix} \mathbf{E}_{in} \quad (1.7)$$

The propagation through the fiber section is achieved through the expression

$$\begin{pmatrix} \mathbf{E}'_s \\ \mathbf{E}'_f \end{pmatrix} = e^{-\frac{\alpha d}{2}} \begin{pmatrix} e^{-j\beta_s d} & 0 \\ 0 & e^{-j\beta_f d} \end{pmatrix} \begin{pmatrix} \mathbf{E}_s \\ \mathbf{E}_f \end{pmatrix} \quad (1.8)$$

where β_s, β_f are the propagation constants of the slow and fast modes respectively, α is the power attenuation coefficient and d is the thickness of the section. Pulling out a common factor $e^{-\frac{\alpha d}{2}} e^{-j\beta_{av} d}$

where $\beta_{av} = \frac{1}{2}(\beta_s + \beta_f)$ we can rewrite the propagation matrix as

$$W = e^{-\frac{\alpha d}{2}} e^{-j\beta_{av} d} \begin{pmatrix} e^{-j\Gamma/2} & 0 \\ 0 & e^{j\Gamma/2} \end{pmatrix} \quad (1.9)$$

where $\Gamma = (\beta_s - \beta_f) d = \frac{2\pi}{\lambda_0} (\Delta n_{eff}) d$, Δn_{eff} being the birefringence of the section. The common

phase and loss terms are usually dropped in pure polarization calculations because they do not affect the state of polarization (SOP). To transform back the field to the reference x-y axes we need another coordinate transformation given by the matrix $\xi(-\psi)$. The whole section can be represented by a Jones matrix, which is the product of all three matrices as $U = \xi(-\psi) W \xi(\psi)$. The relationship between the output and input electric field vectors can consequently be expressed as

$$\mathbf{E}_{out} = U \mathbf{E}_{in} \quad (1.10)$$

For a concatenation of fiber sections, the fiber can be described by the product of the individual Jones matrices of the different sections. Figure 1.10 illustrates how a fiber can be modeled as a concatenation of different sections with random axes orientations.

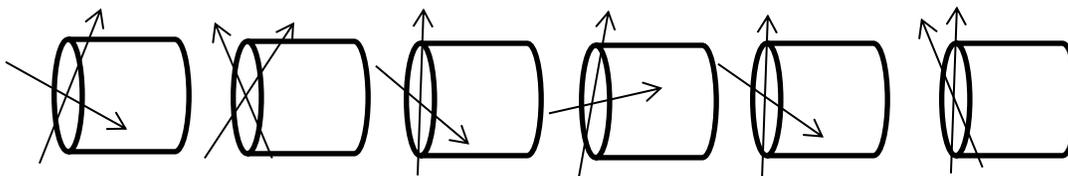


Figure 1.10. A fiber is modeled as a concatenation of birefringent segments. The orientation of each segment is chosen randomly. The arrows indicate the slow and fast axes orientations of each segment.

Jones matrices satisfy a number of important properties. Time reversal symmetry and reciprocity principles lead to the unitary property of Jones matrices, hence $U^\dagger U = I$ where I is the identity matrix [10]. If a Jones matrix of a system is represented as

$$U = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad (1.11)$$

The unitary property of the matrix dictates that $D = A^*$, $C = -B^*$ where A^* is the complex conjugate of A . Hence any Jones matrix can be written as

$$U = \begin{bmatrix} A & B \\ -B^* & A^* \end{bmatrix} \quad (1.12)$$

It's important to note that these conditions hold for optical fibers but not for magneto-optical materials because they do not satisfy the reciprocity principle. In the absence of polarization dependent loss which also holds for optical fibers, the intensity of the electric field vector does not change during propagation apart from a common loss factor for both polarizations that can be accounted for by multiplying the Jones matrix of the system by $e^{-\alpha z/2}$. Therefore, $AA^* + BB^* = 1$ and hence $\det(U) = 1$.

A very useful phenomenological model for the description of PMD in optical fibers depending on the global parameters of the fiber's Jones matrix was first developed in 1986 [11]. We discuss this principal states model on which all PMD modeling and measurements are based in the next section.

1.3.3 Principal States Model

The principal states of an optical fiber are output polarization states that do not change their polarization to first order in frequency. A fiber can be described (as discussed in the previous section) by a frequency dependent matrix given by

$$T(\omega) = e^{\gamma(\omega)}U(\omega) \quad (1.13)$$

where ω is the angular frequency, $\gamma(\omega) = \alpha(\omega)/2 - j\beta(\omega)$ is in general complex as it includes the common phase and loss terms. $U(\omega)$ is the unitary Jones matrix given by

$$U(\omega) = \begin{bmatrix} u_1(\omega) & u_2(\omega) \\ -u_2^*(\omega) & u_1^*(\omega) \end{bmatrix} \quad (1.14)$$

where $|u_1|^2 + |u_2|^2 = 1$. The relation between the output and input polarizations is

$$\mathbf{E}_{out} = T(\omega)\mathbf{E}_{in} \quad (1.15)$$

We can write the complex electric fields at the input and output as follows

$$\mathbf{E}_{out/in} = \varepsilon_{out/in} e^{j\varphi_{out/in}} \boldsymbol{\varepsilon}_{out/in} \quad (1.16)$$

ε is the field amplitude, φ is the phase and $\boldsymbol{\varepsilon}$ is a unit vector describing the direction of polarization. Differentiating Eq. (1.15) and assuming that the input polarization is independent of frequency we have

$$\frac{d\mathbf{E}_{out}}{d\omega} = \frac{dT}{d\omega} \mathbf{E}_{in} = e^{\beta} [\gamma' U + U'] \mathbf{E}_{in} \quad (1.17)$$

where the prime denotes differentiation with respect to frequency. Differentiating Eq. (1.16), we obtain

$$\frac{d\mathbf{E}_{out}}{d\omega} = \left[\frac{\dot{\boldsymbol{\varepsilon}}_{out}}{\boldsymbol{\varepsilon}} + j\dot{\varphi}_{out} \right] \mathbf{E}_{out} + \boldsymbol{\varepsilon}_{out} e^{j\varphi_{out}} \frac{d\boldsymbol{\varepsilon}_{out}}{d\omega} \quad (1.18)$$

Combining Eq. (1.17) and Eq. (1.18) and making use of Eq. (1.15) we obtain

$$\boldsymbol{\varepsilon}_{out} e^{j\varphi_{out}} \frac{d\boldsymbol{\varepsilon}_{out}}{d\omega} = e^\gamma [U' - jkU] \mathbf{E}_{in} \quad (1.19)$$

where

$$k = \dot{\varphi}_{out} + j \left[\gamma' - \frac{\dot{\boldsymbol{\varepsilon}}_{out}}{\boldsymbol{\varepsilon}_{out}} \right] \quad (1.20)$$

We are looking for polarization states that do not change to first order with frequency. Hence we set $d\boldsymbol{\varepsilon}_{out}/d\omega = 0$ to obtain the following eigenvalue equation

$$[U' - jkU] \boldsymbol{\varepsilon}_{in} = 0 \quad (1.21)$$

Eq. (1.21) implies that for such states to exist the value of k has to take on specific values that can be evaluated by setting the determinant of the matrix $U' - jkU$ to zero. Employing $|u_1|^2 + |u_2|^2 = 1$ and its frequency derivative we find the allowed values of k to be

$$k_{\pm} = \pm \sqrt{|u_1'|^2 + |u_2'|^2} \quad (1.22)$$

The corresponding eigenvectors which represent the principal states at the input are given by

$$\boldsymbol{\varepsilon}_{in\pm} = e^{j\rho} \begin{bmatrix} \frac{[u_2' - jk_{\pm}u_2']}{D_{\pm}} \\ -\frac{[u_1' - jk_{\pm}u_1']}{D_{\pm}} \end{bmatrix} \quad (1.23)$$

where ρ is an arbitrary phase and $D_{\pm} = \sqrt{2k_{\pm} \left(k_{\pm} - \text{Im} \left[u_1^* u_1' + u_2^* u_2' \right] \right)}$ with Im indicating imaginary part. It can be easily shown that the two principal states form an orthonormal pair

$$\boldsymbol{\varepsilon}_{in+} \cdot \boldsymbol{\varepsilon}_{in-}^* = 0 \quad \text{and} \quad \boldsymbol{\varepsilon}_{in\pm} \cdot \boldsymbol{\varepsilon}_{in\pm}^* = 1 \quad (1.24)$$

These principal states represent two orthogonal input polarization states whose corresponding output polarization states are independent of frequency to first order. A pulse aligned to the polarization of one of the principal states will remain polarized to first order in frequency at the fiber output. This implies that a pulse with a small enough frequency spectrum will not broaden as a result of propagation in the fiber to first order. The advantage of the principal states model is that it provides a simple picture for pulse broadening inside an optical fiber. A pulse of arbitrary polarization can be decomposed into the principal states at the input of the fiber. The two waves arrive at the output of the fiber each polarized along the principal state direction at the output. The resulting output field is determined by the relative magnitude and phase of each of these components. The time of flight of the principal states is in general different. Consequently, a pulse travelling along the fiber is distorted at the output. To determine the propagation delay of the two pulses we consider Eq. (1.20) and Eq. (1.22). For both of these equations to be satisfied at the same time we must have

$$\text{Re}(\gamma') = \frac{\mathcal{E}'_{out+}}{\mathcal{E}_{out+}} = \frac{\mathcal{E}'_{out-}}{\mathcal{E}_{out-}} \quad (1.25)$$

and

$$\varphi'_{out\pm} = \text{Im}(\gamma') \pm \sqrt{|u_1'|^2 + |u_2'|^2} \quad (1.26)$$

The frequency derivative of the output phase of the principal state in Eq. (1.26) is the propagation delay. As can be seen, one state experiences more delay and hence is named the slow state and the

other is named the fast state. The common time delay is designated by τ_o and the differential time delay, $\tau = 2\sqrt{|u_1|^2 + |u_2|^2}$, is termed the differential group delay (DGD). Hence Eq. (1.26) states that the time delay for the slow and fast states is $\tau_o \pm \tau/2$. It is worth noting that even for waves launched in the principal states, pulse broadening can still occur depending on the frequency dependence of the time delays and the second order variation of the state of polarization.

Another description for polarization of light is provided by the Stokes parameters. We define these parameters in the next section.

1.3.4 Stokes Parameter Description

The Stokes parameters provide another description of polarization phenomenon as opposed to Jones calculus. From this point onwards in this chapter, we describe the electric field in Jones space using Dirac notation as a ket expressible as a column vector $|s\rangle = (s_x, s_y)$ such that $s_x s_x^* + s_y s_y^* = 1$. The real 3D Stokes vector, $\hat{s} = (s_1, s_2, s_3)$ parameters are defined by

$$\begin{aligned} s_1 &= s_x s_x^* - s_y s_y^* \\ s_2 &= s_x s_y^* + s_y s_x^* \\ s_3 &= j(s_x s_y^* - s_y s_x^*) \end{aligned} \quad (1.27)$$

The Stokes vector also satisfies $s_1^2 + s_2^2 + s_3^2 = 1$. The most general SOP can be described by the following Jones vector

$$\begin{pmatrix} \cos \psi \\ \sin \psi e^{j\delta} \end{pmatrix} \quad (1.28)$$

where δ is the phase shift between the x and y components. Substituting into Eq. (1.27), we find the Stokes parameters to be

$$\begin{aligned} s_1 &= \cos 2\psi \\ s_2 &= \sin 2\psi \cos \delta \\ s_3 &= \sin 2\psi \sin \delta \end{aligned} \quad (1.29)$$

The form of Eq. (1.29) suggests that the Stokes vector can be represented on a unit sphere called the Poincare sphere. Each Stokes parameter is plotted as a Cartesian axis where each point on the surface of the sphere represents a unique polarization state. The Poincare sphere is displayed in Figure 1.11.

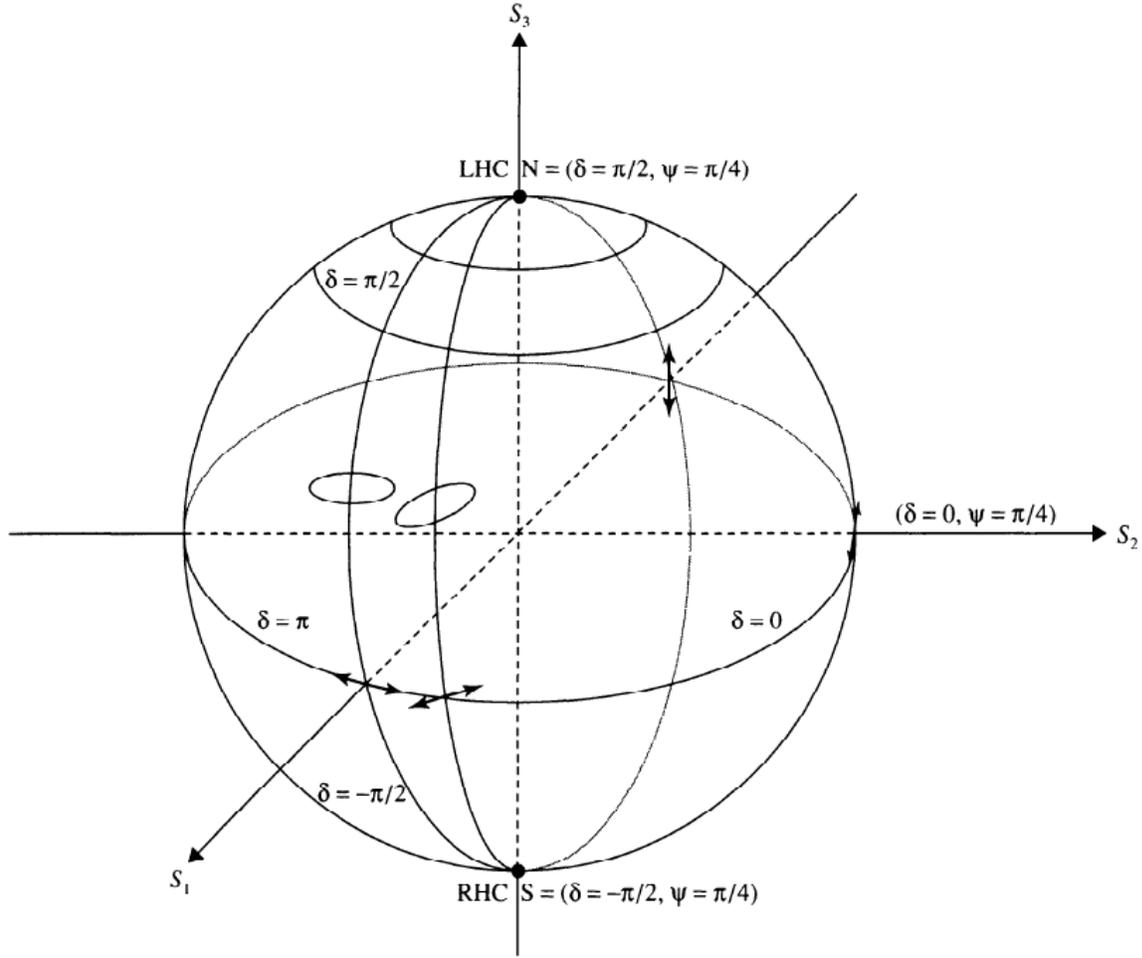


Figure 1.11. Poincare sphere. A few representative polarization states are indicated [10].

Linear polarization states occur for $\delta = 0, \pm\pi$, hence are represented as points on the equator in the $s_1 - s_2$ plane. The north and south poles of the sphere represent left circularly polarized and right circularly polarized states respectively. Any other point on the surface corresponds to elliptical polarization.

If the output polarization state in Jones space is $|t\rangle$, the relation between the output and input polarization states in a birefringent system is $|t\rangle = U|s\rangle$ where $|s\rangle$ is the input state of polarization

and U is the 2x2 Jones matrix of the system. A similar relationship exists between the input and output states of polarization in Stokes space \hat{s} and \hat{t} respectively. Namely,

$$\hat{t} = R\hat{s} \quad (1.30)$$

where R is a 3x3 rotation matrix that is also unitary. There is a one to one correspondence between the Jones matrix and the rotation matrix of any system. Given the Jones vector $|s\rangle$, we can express the corresponding Stokes vector conveniently using the Pauli spin matrices [10]

$$\hat{s} = \begin{pmatrix} \langle s|\sigma_1|s\rangle \\ \langle s|\sigma_2|s\rangle \\ \langle s|\sigma_3|s\rangle \end{pmatrix} \quad (1.31)$$

where σ_1, σ_2 and σ_3 are the 2x2 Pauli matrices $\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}$. Eq.

(1.31) is usually written compactly as $\hat{s} = \langle s|\vec{\sigma}|s\rangle$, where $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the Pauli spin vector in Stokes space. The connection between U and R can be obtained as follows. From Eqs.(1.30) and (1.31) we have $\hat{v} = \langle v|\vec{\sigma}|v\rangle = R\langle s|\vec{\sigma}|s\rangle = \langle s|R\vec{\sigma}|s\rangle = \langle s|U^\dagger\vec{\sigma}U|s\rangle$. Hence

$$R\vec{\sigma} = U^\dagger\vec{\sigma}U \quad (1.32)$$

The projection operator is given by

$$|s\rangle\langle s| = \frac{1}{2}(I + \hat{s}\cdot\vec{\sigma}) \quad (1.33)$$

where I is the identity matrix. The inner product of two arbitrary polarization states $|p\rangle$ and $|q\rangle$ can be expressed in Stokes space by setting $s = p$ in Eq. (1.33) and multiplying with $|q\rangle$ from the right and $\langle q|$ from the left to obtain

$$\langle q|p\rangle\langle p|q\rangle = \frac{1}{2}(I + \langle q|\hat{p}\cdot\vec{\sigma}|q\rangle) = \frac{1}{2}(I + \hat{p}\cdot\langle q|\vec{\sigma}|q\rangle) = \frac{1}{2}(I + \hat{p}\cdot\hat{q}) \quad (1.34)$$

From Eq. (1.34), if two states are orthogonal in Jones space, then it follows that $\hat{p} \cdot \hat{q} = -1$.

Let the eigenvectors of the Jones matrix be $|r\rangle$ and $|r_{-}\rangle$. By definition, any SOP proportional to any of these eigenvectors will not change state upon propagation through the system. Since these states form a complete orthogonal set of Jones vectors, we can express U as $\rho_1|r\rangle\langle r| + \rho_2|r_{-}\rangle\langle r_{-}|$ where ρ_1 and ρ_2 are its eigenvalues. Using the unitary and unity determinant properties of the Jones matrix, we find that both eigenvalues have unit amplitude and their product must be unity. Hence we can write U in the following form

$$U = e^{-j\phi}|r\rangle\langle r| + e^{j\phi}|r_{-}\rangle\langle r_{-}| \quad (1.35)$$

Using the projection operator along with the orthogonality of $|r\rangle$ and $|r_{-}\rangle$, we obtain the rotational form of the Jones matrix

$$U = \cos(\phi)I - j \sin(\phi)\hat{r} \cdot \vec{\sigma} = e^{-j\phi\hat{r} \cdot \vec{\sigma}} \quad (1.36)$$

Using Eq. (1.32), R can be expressed as

$$R = \cos(2\phi)I + (1 - \cos(2\phi))\hat{r}\hat{r} + \sin(2\phi)\hat{r} \times \quad (1.37)$$

where $\hat{r}\hat{r}$ is the 3D dyadic operator and $\hat{r} \times$ is the cross product operator defined in matrix notation as

$$\hat{r}\hat{r} = \begin{pmatrix} r_1r_1 & r_1r_2 & r_1r_3 \\ r_2r_1 & r_2r_2 & r_2r_3 \\ r_3r_1 & r_3r_2 & r_3r_3 \end{pmatrix}; \hat{r} \times = \begin{pmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{pmatrix} \quad (1.38)$$

R can also be expressed in analogy to Eq. (1.37) as [9]

$$R = e^{2\phi\hat{r} \times} \quad (1.39)$$

The operation represented by Eq. (1.30) can be viewed as a rotation on the Poincare sphere of the initial state of polarization about the axis \hat{r} by a rotation angle 2ϕ [10]. If the input polarization is aligned with the rotation axis, the output and the input SOP coincide and this can be deduced by applying Eq. (1.37).

The propagation operation $\xi(-\psi)W\xi(\psi)$ through a fiber section can be given an interpretation in Stokes space by writing each matrix in terms of the Pauli spin vector and comparing the results to the rotational form of Eq. (1.36). The first operation is a rotation about the s_3 axis with an angle ψ followed by a rotation about the axis s_1 by an angle $\Gamma/2$ followed finally by another rotation about s_3 with angle ψ . Expressing the propagation operation in rotational form (Eq.(1.36)) we find that

$$\xi(-\psi)W\xi(\psi) = \cos\left(\frac{\Gamma}{2}\right)I - j \sin\left(\frac{\Gamma}{2}\right)(\cos 2\psi\sigma_1 + \sin 2\psi\sigma_2) \quad (1.40)$$

Hence the corresponding rotation matrix in Stokes space has the parameters

$$2\phi = \Gamma; \hat{r} = (\cos 2\psi, \sin 2\psi, 0) \quad (1.41)$$

In the next section, we describe the PMD vector in Stokes space and state some important results on the evolution of the polarization with distance along the fiber and with frequency. We also state the concatenation rule for PMD vectors.

1.3.5 PMD vector

The PMD vector, $\vec{\tau}$ is formed from the multiplication of the DGD with the Stokes vector of the slow principal state. The transformation of the input PMD vector to the output by a birefringent network with a rotation matrix R is given by [9]

$$\vec{\tau}_t = R\vec{\tau}_s \quad (1.42)$$

where $\vec{\tau}_s$ is the input PMD vector and $\vec{\tau}_t$ is the output PMD vector. Eq. (1.42) is the same relationship that links the input and output polarizations. The output polarization evolves with frequency as [9]

$$\frac{\partial \hat{s}}{\partial \omega} = \vec{\tau} \times \hat{s} \quad (1.43)$$

Eq. (1.43) can be given a geometrical interpretation. The output SOP, precesses on the Poincare sphere around the PMD vector when the frequency is varied. To obtain the variation of the SOP with distance, we note that an infinitesimal fiber segment of length dz can be described by a rotation matrix having an angle $2\phi = \frac{2\pi}{\lambda} \Delta n dz$ and a rotation axis \hat{r} given by Eq. (1.41). Using Eq. (1.37), R can be written as $I + 2\phi \hat{r} \times$. The polarization at the output of the section can be obtained from

$$\hat{s} + d\hat{s} = R\hat{s} = \hat{s} + \frac{2\pi}{\lambda} \Delta n dz \hat{r} \times \hat{s} \quad (1.44)$$

Defining the local birefringence vector as $\vec{\beta} = \frac{2\pi}{\lambda} \Delta n \hat{r}$, we obtain the evolution equation for the polarization along the fiber as

$$\frac{\partial \hat{s}}{\partial z} = \vec{\beta} \times \hat{s} \quad (1.45)$$

This equation, similar to (1.43), describes a rotation of the SOP on the Poincare sphere around the birefringence vector as light propagates along the fiber. Using Eqs. (1.43) and (1.45) we obtain the evolution equation for the PMD vector [12]

$$\frac{\partial \vec{\tau}}{\partial z} = \frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\tau} \quad (1.46)$$

Finally, we state the concatenation rule of $\vec{\tau}$ for two concatenated sections. Given two fiber sections with PMD vectors $\vec{\tau}_1$ and $\vec{\tau}_2$ and rotation matrices R_1 and R_2 as shown in Figure 1.12. The PMD vector of the concatenation of the two segments is

$$\vec{\tau} = \vec{\tau}_2 + R_2 \vec{\tau}_1 \quad (1.47)$$

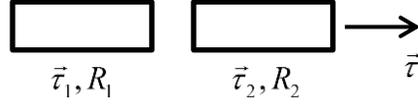


Figure 1.12. The concatenation of two fiber segments and the resulting PMD vector.

1.3.6 Manakov PMD equation

In order to assess the performance of an optical communication system we need to take into account all propagation effects that occur during transmission over the optical fiber. For this purpose we develop in this section the evolution equation for the propagation of the electric field along the fiber that takes into account group velocity dispersion (GVD), PMD and the nonlinear response of the dielectric constant. The presence of EDFA's introduces ASE (amplified spontaneous emission) which is a source of noise that accumulates along the fiber length since repeaters are absent. In order to maintain signal integrity at the receiver, a large input power is usually required to overcome ASE, typically in excess of 1mW. With such high input power, the nonlinear response of the dielectric constant of the fiber cannot be neglected. However, because the nonlinear contribution is much smaller than the linear contribution, we treat the former as a perturbation [13], [14].

In our first analysis, we will assume that the nonlinear contribution to the dielectric constant vanishes. We start by considering the Jones vector of the electric field in the frequency domain given by

$$|A(z, \omega)\rangle = \begin{pmatrix} A_x(z, \omega) \\ A_y(z, \omega) \end{pmatrix} \quad (1.48)$$

Consider the propagation of this Jones vector through a very small section of fiber of length Δz neglecting nonlinear effects. While we have so far considered LP modes, we are going to treat the more general case by assuming that the fiber section supports two orthogonal elliptically polarized modes. These can be expressed in normalized form as

$$|m_1\rangle = \begin{pmatrix} \cos \psi \\ \sin \psi e^{j\delta} \end{pmatrix} \quad \text{and} \quad |m_2\rangle = \begin{pmatrix} -\sin \psi e^{-j\delta} \\ \cos \psi \end{pmatrix} \quad (1.49)$$

In the case of LP modes, $\delta = 0$. Both ψ and δ vary with z and represent the varying direction of birefringence along the fiber. We write Eq. (1.48) in terms of $|m_1\rangle$ and $|m_2\rangle$ as

$$A_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + A_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} = A_x |s_1\rangle + A_y |s_2\rangle = \chi_1 |m_1\rangle + \chi_2 |m_2\rangle \quad (1.50)$$

where the coefficients χ_1 and χ_2 can be determined from the orthogonality of $|m_1\rangle$ and $|m_2\rangle$. Multiplying Eq. (1.50) by $\langle m_1|$ and $\langle m_2|$ we obtain

$$\begin{aligned} \chi_1 &= \langle m_1 | s_1 \rangle A_x + \langle m_1 | s_2 \rangle A_y \\ \chi_2 &= \langle m_2 | s_1 \rangle A_x + \langle m_2 | s_2 \rangle A_y \end{aligned} \quad (1.51)$$

Rewriting this result in matrix form we have

$$\begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \xi(\psi, \delta) \begin{pmatrix} A_x \\ A_y \end{pmatrix} \quad (1.52)$$

$$\xi(\psi, \delta) = \begin{pmatrix} \cos \psi & \sin \psi e^{-j\delta} \\ -\sin \psi e^{j\delta} & \cos \psi \end{pmatrix} \quad (1.53)$$

which is a more general form for the transformation matrix of Eq. (1.7). Similar to Eq. (1.9), the propagation step can be represented as

$$e^{-\frac{\alpha}{2}\Delta z} e^{-j\beta_{av}\Delta z} \begin{pmatrix} e^{-j\Gamma/2} & 0 \\ 0 & e^{j\Gamma/2} \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \quad (1.54)$$

with $\Gamma = \Delta\beta\Delta z$. It is important to note that β_{av} and $\Delta\beta$ are functions of frequency. The attenuation coefficient has a very weak dependence on frequency, thus it can be considered constant.

After the propagation step, we transform the electric field back to the x-y reference frame. For this purpose we follow the procedure leading to Eq. (1.51) to convert from the $m_1 - m_2$ reference frame to the $x - y$ reference frame. The required matrix is $\xi(-\psi, \delta)$. Therefore we can write an expression for $|A(z + \Delta z, \omega)\rangle$ as follows

$$|A(z + \Delta z, \omega)\rangle = \xi(-\psi, \delta) e^{-j\beta_{av}(\omega)\Delta z} \begin{pmatrix} e^{-j\Gamma/2} & 0 \\ 0 & e^{j\Gamma/2} \end{pmatrix} \xi(\psi, \delta) |A(z, \omega)\rangle \quad (1.55)$$

Expanding both sides as a function of z and keeping only first order terms in Δz we obtain the required evolution equation

$$\frac{\partial |A(z, \omega)\rangle}{\partial z} = -j \left(\beta_{av}(\omega) + \frac{\Delta\beta(\omega)}{2} M \right) |A(z, \omega)\rangle \quad (1.56)$$

where M is a random matrix given by

$$M = \xi(-\psi, \delta) \sigma_1 \xi(\psi, \delta) \quad (1.57)$$

where σ_1 is the Pauli matrix defined in section 1.3.4. If we write the matrix ξ in terms of the Pauli matrices, we can show that M can be expressed as

$$M = \sigma_1 \cos 2\psi + \sigma_2 \sin 2\psi \cos \delta + \sigma_3 \sin 2\psi \sin \delta \quad (1.58)$$

or in more compact form as

$$M = (\cos 2\psi, \sin 2\psi \cos \delta, \sin 2\psi \sin \delta) \cdot \vec{\sigma} \quad (1.59)$$

where $\vec{\sigma}$ is the Pauli spin vector defined also in section 1.3.4. It's important to note that the vector $\hat{s} = (\cos 2\psi, \sin 2\psi \cos \delta, \sin 2\psi \sin \delta)$ is the Stokes vector of the eigen polarizations for the short section of length Δz . Since both ψ and δ vary randomly along the fiber, \hat{s} can be considered a random vector on the Poincare sphere and M a unitary random rotation matrix.

We are interested in the evolution of the field with time. To transform Eq. (1.56) into the time domain would require knowledge of $\beta_{av}(\omega)$ and $\Delta\beta(\omega)$. Such information is not available in analytical form. To be able to proceed further, we assume that the spectral width of the electric field pulses is much smaller than the carrier frequency which is a valid approximation for pulse widths as small as 0.1 ps or, equivalently, up to a symbol rate of 10Tb/s which is a condition met by all state of the art communication systems. We therefore expand $\beta_{av}(\omega)$ and $\Delta\beta(\omega)$ in a Taylor series around the carrier frequency ω_0 as

$$\beta_{av}(\omega) \approx \beta_0 + \beta_1(\omega - \omega_0) + \frac{\beta_2}{2}(\omega - \omega_0)^2 \quad (1.60)$$

$$\Delta\beta \approx \Delta\beta_0 + \Delta\beta_1(\omega - \omega_0) \quad (1.61)$$

And keep terms up to second order in $\beta_{av}(\omega)$ and first order in $\Delta\beta(\omega)$ [4]. Substituting Eq. (1.60) and (1.61) into (1.56), noting the association $\frac{\partial}{\partial t} \leftrightarrow j(\omega - \omega_0)$, and taking the inverse Fourier transform we arrive at

$$\begin{aligned} \frac{\partial}{\partial z} |A(z, t)\rangle = & -j\beta_0 |A(z, t)\rangle - j\Delta\beta_0 M |A(z, t)\rangle - \beta_1 \frac{\partial}{\partial t} |A(z, t)\rangle + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} |A(z, t)\rangle \\ & - \frac{\Delta\beta_1}{2} M \frac{\partial}{\partial t} |A(z, t)\rangle - \frac{\alpha}{2} |A(z, t)\rangle \end{aligned} \quad (1.62)$$

Up to this point we have neglected the nonlinear term. To include nonlinearity, it can be shown that we need to add a term of the form [4], [14]

$$-j\gamma \left(\langle A|A\rangle |A\rangle - \frac{1}{3} \langle A|M_3|A\rangle M_3|A\rangle \right) \quad (1.63)$$

to the right hand side of Eq. (1.62). The constant γ is given by $2\pi n_2 / \lambda_0 A_{eff}$ where n_2 is a constant and A_{eff} is the effective area of the fiber mode. M_3 is a random rotation matrix given by $M_3 = \xi(-\psi, \delta) \sigma_3 \xi(\psi, \delta)$. The first two terms on the right hand side of Eq. (1.62) can be eliminated by the transformation $|A\rangle = M'|A'\rangle$ where the matrix M' satisfies

$$\frac{\partial M'}{\partial z} = -j(\beta_0 I + \Delta\beta_0 M)M' \quad (1.64)$$

Substituting in (1.62), using (1.63), (1.64) and dropping the prime from the Jones vector we obtain

$$\begin{aligned} \frac{\partial}{\partial z}|A\rangle = & -\beta_1 \frac{\partial}{\partial t}|A\rangle + \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2}|A\rangle - \frac{\Delta\beta_1}{2} M'^T M M' \frac{\partial}{\partial t}|A\rangle - \frac{\alpha}{2}|A\rangle + \\ & -j\gamma \left(\langle A|A\rangle|A\rangle - \frac{1}{3} \langle A|M'^T M_3 M'|A\rangle M'^T M_3 M'|A\rangle \right) \end{aligned} \quad (1.65)$$

where a capital T superscript indicates the transpose of a matrix. The first term on the right hand side of Eq. (1.65) can be eliminated by substituting $t' = t - \beta_1 z$. This term corresponds to a constant delay as the signal propagates down the fiber and the substitution enables us to work in a reference frame moving with the signal. Nonlinear changes occur over lengths of a few kms while birefringence changes occur over lengths of tens of meters [14]. This permits us to employ an approximation for solving (1.65) by averaging the nonlinear term over the birefringence fluctuations. The result is [4]

$$\text{mean}(\langle A|M'^T M_3 M'|A\rangle M'^T M_3 M'|A\rangle) = \frac{1}{3} \langle A|A\rangle|A\rangle \quad (1.66)$$

Substituting back into Eq. (1.65) and dropping the prime from the new time variable we arrive at the central result of this section, the Manakov PMD equation

$$\frac{\partial}{\partial z}|A\rangle = \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2}|A\rangle - \frac{1}{2} \vec{b}(z) \cdot \vec{\sigma} \frac{\partial}{\partial t}|A\rangle - \frac{\alpha}{2}|A\rangle - j\gamma \frac{8}{9} \langle A|A\rangle|A\rangle \quad (1.67)$$

where we have written the rotation matrix $M'^T M M' = \hat{r} \cdot \vec{\sigma}$ as in Eq. (1.59) where \hat{r} is a random vector on the Poincare sphere and

$$\vec{b}(z) = \Delta\beta_1 \hat{r} \quad (1.68)$$

Since $\Delta\beta(\omega) = \frac{\omega}{c} \Delta n_{eff}(\omega)$, where c is the speed of light, and the frequency dependence of Δn_{eff} is negligible, $\Delta\beta_1 \approx \frac{\Delta n_{eff}}{c}$ and hence

$$\vec{\beta} = \omega \vec{b}(z) \quad (1.69)$$

is the local birefringence vector defined in section 1.3.5. The β_2 term represents GVD. In order to solve this equation, we need to treat the random dependence of PMD that enters through the $\vec{b}(z)$ term. In the next section we discuss the statistics of PMD and statistical models that best describe fiber systems.

1.3.7 Statistics of PMD

Lightwave system impairments such as those due to chromatic dispersion, attenuation or nonlinearities will have some uncertainty caused by fabrication tolerances, aging, imprecise wavelength control and other factors. The probability density functions of these impairments vanish beyond a certain value [15] and hence allow system design to tolerate a worst case impairment. However, probability density functions of PMD have asymptotic tails that extend to unacceptable levels of impairment. Instead, when PMD is relevant, a system cannot be designed according to the worst PMD impairment, but rather with a specified probability of outage.

Quantities of interest in statistical analysis of PMD include the PMD vector (also known as the first order PMD), the DGD, the second order PMD (the frequency derivative of the PMD vector) and the autocorrelation functions. An important assumption that has been verified experimentally in statistical analysis of PMD is that a statistical average calculated over a range of optical frequencies is equivalent to an average calculated over time if hinge behavior (to be explained in the next section) is not present. This assumption is necessary because PMD usually changes slowly over time and it is not often practical to make long time measurements of a system to determine its statistical properties.

PMD statistics has been studied thoroughly by several researchers [16], [17]. We here follow a simplified model presented in [4] that leads to the same results. The components of the vector \vec{b} of Eq. (1.68) are assumed to be statistically independent, stationary stochastic processes with first two moments given by

$$\langle b_i(z) \rangle = 0, \quad \langle b_i(z) b_j(z) \rangle = \frac{1}{3} \delta_{ij} (\Delta b)^2 e^{(-|z-z'|/L_c)} \quad (1.70)$$

where i, j take on the values x, y and z , δ_{ij} is the Kronecker delta. The correlation length L_c , is a measure of decorrelation of the birefringence components and $(\Delta b)^2$ is a measure of the standard deviation of the birefringence. With these assumptions, Eq. (1.46) becomes

$$\frac{\partial \vec{\tau}}{\partial z} = \vec{b} + \omega \vec{b} \times \vec{\tau} \quad (1.71)$$

The second term of Eq. (1.71) represents a rotation of the PMD vector and hence does not affect its amplitude. Neglecting this term, the components of $\vec{\tau}$ can be obtained by direct integration

$$\tau_i = \int_0^z b_i(z') dz' \approx \sum_{n=1}^N b_i(z_n) \Delta z \quad (1.72)$$

Where we have approximated the integral by a summation which is equivalent to physically dividing the fiber into N segments each with length Δz . If we assume that the birefringence fluctuations in the various segments are independent, then τ_i is the sum of a large number of independent random variables with the same statistical distribution. It follows from the central limit theorem that the three components of $\vec{\tau}$ satisfy Gaussian distributions with zero mean given by

$$p(\tau_i) = \frac{1}{\sqrt{2\pi \langle \tau_i^2 \rangle}} e^{-\tau_i^2 / (2 \langle \tau_i^2 \rangle)} \quad (1.73)$$

The variance can be obtained using Eq. (1.72) as

$$\langle \tau_i \tau_j \rangle = \delta_{ij} \frac{2}{3} (\Delta b)^2 L_c^2 \left[e^{-z/L_c} + z/L_c - 1 \right] \quad (1.74)$$

The DGD is given by $\tau = \sqrt{\tau_x^2 + \tau_y^2 + \tau_z^2}$ and hence assumes a Maxwellian distribution. In terms of the mean DGD, the distribution has the form

$$p(\tau) = \frac{8}{\pi^2 \langle \tau \rangle} \left(\frac{2\tau}{\langle \tau \rangle} \right)^2 e^{-(2\tau/\langle \tau \rangle)^2/\pi} \quad (1.75)$$

which is defined for $\tau \geq 0$ and is zero otherwise. In terms of the mean DGD, we can also rewrite Eq. (1.73) as

$$p(\tau_i) = \frac{2}{\pi \langle \tau \rangle} e^{-(2\tau/\langle \tau \rangle)^2/\pi} \quad (1.76)$$

We note that the components of $\vec{\tau}$ are independent random variables. We also note that the variance of the DGD is given by $\langle \tau^2 \rangle = \frac{3\pi}{8} \langle \tau \rangle^2$ and the variance for each of the components of $\vec{\tau}$ is given by $\langle \tau_i^2 \rangle = \frac{\pi}{8} \langle \tau^2 \rangle$ and therefore $\langle \tau_i^2 \rangle = \langle \tau^2 \rangle / 3$. Observing Eq. (1.74) we see that for $z \ll L_c$, $\langle \tau^2 \rangle$ will be proportional to z^2 and for $z \gg L_c$,

$$\langle \tau^2 \rangle = \frac{2}{3} (\Delta b)^2 L_c z = D_{PMD}^2 z \quad (1.77)$$

proportional to z where

$$D_{PMD} = \Delta b \sqrt{\frac{2}{3} L_c} \quad (1.78)$$

is the PMD coefficient of the fiber in units of ps / \sqrt{km} . As it's usually preferable to minimize the DGD, the correlation length should be as small as possible such that the DGD scales as \sqrt{L} , L

being the fiber length, fast as possible. Compared to other types of dispersion that scale as L , the randomness of the orientation of the fiber sections reduces the impact of PMD causing it to scale as \sqrt{L} .

To simplify the numerical solution of the Manakov PMD Eq. (1.67), the correlation function of Eq.

(1.70) is simplified to the form $\langle b_i(z)b_j(z') \rangle = \frac{2}{3} \delta_{ij} (\Delta b)^2 L_c \delta(z-z')$, obtained by assuming a

very small L_c and hence $e^{(-|z-z'|/L_c)} \rightarrow 2L_c \delta(z-z')$. Upon comparison to Eq. (1.78) we have

$\langle b_i(z)b_j(z') \rangle = D_{PMD}^2 \delta_{ij} \delta(z-z')$. Therefore, the random vector $\vec{b}(z)$ is treated as a white noise

Gaussian process with autocorrelation function $\langle \vec{b}(z) \cdot \vec{b}(z') \rangle = D_{PMD}^2 \delta(z-z')$.

1.3.8 Hinge Model

The statistics described so far assumes that all the fiber sections are influenced by random variations with time. Consequently, the instantaneous PMD of the fiber will occasionally sample the tails of the Maxwellian distribution leading to system outages. However, measurements on field fibers indicated that this “fast mixing assumption” is not realized in practice and that the PMD varies with time in a slower and more restricted manner [18]. In reality, fibers are usually buried underground for large distances and parts of them emerge at different places along the network. These particular places might be bridges where the fiber is exposed to the environment or at amplifier and DCM huts where these devices are also subject to environmental perturbations. Mounting experimental evidence has indicated that the buried fiber does not vary with time and the areas of the network that are exposed to the environment act as the main contributor to the PMD temporal variation. These exposed areas have negligible PMD themselves but act as polarization rotators and hence can vary the direction of the PMD vector. In this “hinge” model, the link is viewed consisting of long spans of stationary sections of fiber for which the PMD vectors are stable connected by hinges which essentially rotate those larger sections. The model does not preclude the possibility of the stationary sections varying with time; however, for shorter time scales (weeks or months) the hinge model will be a more accurate description.

To illustrate the difference between the hinge model and the fast mixing assumption, consider a link in the form of two frozen fiber sections and a hinge connecting them. Using the concatenation rule, we can write the PMD vector as

$$\vec{\tau} = \vec{\tau}_2 + R_2 H(t) \vec{\tau}_1 \quad (1.79)$$

where $H(t)$ is the time dependent rotation matrix of the hinge and $\vec{\tau}_1, \vec{\tau}_2$ are the PMD vectors of the two sections and R_2 is the rotation matrix of section 2. In the fast mixing assumption, $\vec{\tau}_1, \vec{\tau}_2$ can take

on any value with time and hence there is no upper limit to the value of the DGD. However, in the hinge model, since both PMD vectors are fixed, an upper limit on the DGD exists given by $|\vec{\tau}_2| + |\vec{\tau}_1|$ when the rotator $H(t)$ aligns the two vectors and a lower limit given by $||\vec{\tau}_2| - |\vec{\tau}_1||$. Consequently, the statistical distribution of the DGD diverges from Maxwellian form.

1.3.9 PMD Compensation

The purpose of PMD compensation is to reduce the PMD impairment caused by the combination of the fiber and the compensating element. Compensation in the past has been achieved using optical elements but currently optical compensation is being replaced by electronic compensation. In this section, we describe briefly one optical compensation technique and then outline how electronic compensation is achieved.

Optical compensation techniques involve an adaptive element and a control algorithm. In Figure 1.13, we display the setup for the PMD vector nulling technique [15].

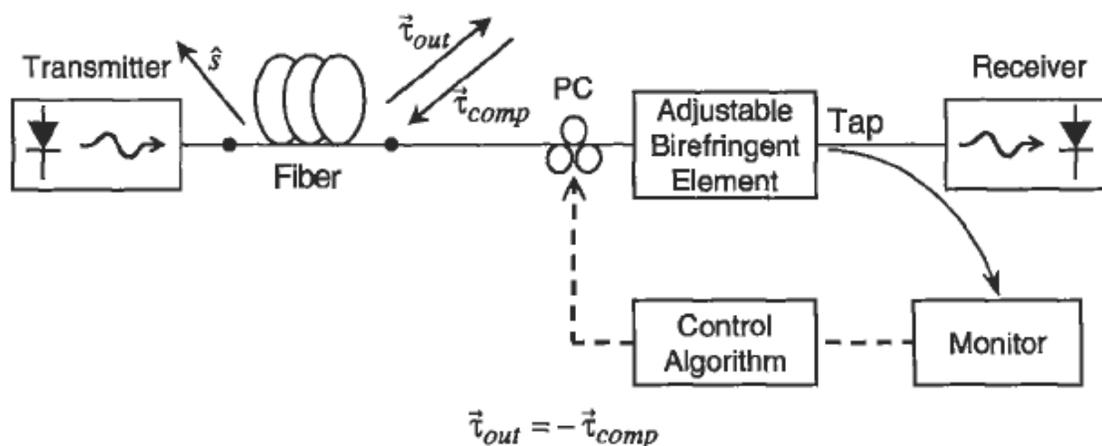


Figure 1.13. Schematic of the PMD nulling technique. The PC and the birefringent element are dynamically controlled to cancel the output PMD vector at the carrier frequency [15].

The adjustable birefringent element compensates the magnitude of the output PMD vector while the PC compensates for the direction such that $\vec{\tau}_{out}(\omega_0) = -\vec{\tau}_{comp}(\omega_0)$, ω_0 being the carrier frequency. $\vec{\tau}_{comp}$ is adjusted dynamically using the control algorithm.

Electronic compensation has the advantage that it can be done in software and thus gets rid of the bulky components required in optical compensation as well as being cost effective. An adaptive FIR

(finite impulse response) digital filter is used to invert the Jones matrix of the fiber to recover the transmitted polarization. The optimal tap weights can be determined using various algorithms [3]. However one commonly used algorithm that exploits the properties of the data in a PM-QPSK system is the constant modulus algorithm, where deviations of the amplitude of the received electric field from unity is used to dynamically vary the tap weights.

1.4 Bit Error Ratio

The performance of a communication system is governed by the bit error ratio (BER) which is the average number of bits received in error after the detection process. The BER is a function of the signal to noise ratio (SNR) at the receiver; the higher the SNR, the smaller the BER. In this section, we demonstrate the calculation of the BER for a communication system with ASK modulation where the only impairment is the noise at the receiver as it's analytically tractable. Then we will show how we can numerically calculate the symbol error ratio (SER) in a QPSK system.

1.4.1 BER for ASK System

Let A be the received amplitude at the receiver when transmitting a logical 1. If the received signal is corrupted by an additive white Gaussian noise process, $w(t)$, such that $\langle w \rangle = 0$ and $\langle w(t)w(t') \rangle = \mu^2 \delta(t-t')$, where μ^2 is the noise power spectral density. The random process representing the received signal at the receiver is therefore

$$A_r = A + w(t) \quad (1.80)$$

This is also a Gaussian process with a mean A and a variance given by μ^2 . Therefore the conditional probability density function (pdf) of the received signal given that a logical 1 was transmitted is given by

$$f_1 = f_{A_r}(A_r/1) = \frac{1}{\mu\sqrt{2\pi}} \exp\left(-\frac{(A_r - A)^2}{2\mu^2}\right) \quad (1.81)$$

In case a logical 0 was transmitted, the received signal is $A_r = w(t)$. Consequently the pdf given that a logical 0 is transmitted is given by

$$f_0 = f_{A_r}(A_r/0) = \frac{1}{\mu\sqrt{2\pi}} \exp\left(-\frac{A_r^2}{2\mu^2}\right) \quad (1.82)$$

The receiver compares the received signal with a certain threshold A_{th} and if $A_r > A_{th}$ it decides that a 1 was transmitted otherwise if $A_r < A_{th}$, it decides for a 0. Hence there are two types of error to consider. The error that the receiver detects 1 given that 0 is transmitted and the error that 0 is detected given that 1 is transmitted. The probabilities of these two events are given by

$$p_{1/0} = \int_{A_{th}}^{\infty} f_0 dA_r = \frac{1}{\mu\sqrt{2\pi}} \int_{A_{th}}^{\infty} \exp\left(-\frac{A_r^2}{2\mu^2}\right) dA_r \quad (1.83)$$

and

$$p_{0/1} = \int_{-\infty}^{A_{th}} f_1 dA_r = \frac{1}{\mu\sqrt{2\pi}} \int_{-\infty}^{A_{th}} \exp\left(-\frac{(A_r - A)^2}{2\mu^2}\right) dA_r \quad (1.84)$$

respectively. If the probabilities of transmitting 1 and 0 are p_1 and p_0 respectively, then the BER is given by

$$BER = p_1 p_{0/1} + p_0 p_{1/0} \quad (1.85)$$

Since the data stream is random, $p_1 = p_0 = 1/2$. If we substitute Eq. (1.83) and (1.84) into (1.85) and differentiate with respect to A_{th} , we obtain an optimal value $A_{th} = A/2$. Substituting back into Eq. (1.85) we obtain an expression for the BER

$$BER = \frac{1}{2} \operatorname{erfc}\left(\frac{A}{2\sqrt{2}\mu}\right) \quad (1.86)$$

where $\operatorname{erfc}(x)$ is the complimentary error function. The ratio A/μ is proportional to \sqrt{SNR} as the signal power is proportional to A^2 and the noise power to μ^2 .

1.4.2 SER for QPSK System

The calculation of the *SER* in QPSK, requires the evaluation of four conditional pdfs. If the received symbol lies outside its corresponding quadrant, this results in an error. To illustrate, let the conditional pdf's of each symbol S be denoted by $p_k(S/S_k)$ where S_k is symbol k , $k = 1$ to 4 . The *SER* is defined as

$$SER = \sum_k q_k \iint_{\text{area of error}} p_k(S/S_k) \quad (1.87)$$

where q_k is the probability of occurrence of symbol k and the area of error is the area outside the corresponding quadrant of symbol k . However, from the symmetry of the problem, the *SER* can be obtained from the conditional pdf of any symbol since all symbols experience the same impairments and receiver noise. Furthermore, all symbols occur equally likely. Consequently, $q_k = 1/4$ and the values of $\iint_{\text{area of error}} p_k(S/S_k)$ are identical. Hence, the *SER* is given by $\iint_{\text{area of error}} p_k(S/S_k)$ where k can be any symbol.

1.5 The Multicanonical Method

As seen in the previous section, the performance of a communication system, quantified by its *SER*, is computed from the pdf of the received symbols. As this requires simulated pdf values of the order of 10^{-9} , standard Monte Carlo methods must be supplemented by variance reduction techniques. The multicanonical method [19], unlike, for example, importance sampling is both easily programmed and does not require any prior knowledge regarding the structure of the pdf [19], [20]. Consequently, stochastic functions relevant to optical communication system performance are often calculated with multicanonical techniques [21], [22], [23], [24].

In the multicanonical method the parameters governing the behavior of a system are initially assigned random values and the relevant output (control) variables are calculated. Subsequently, these parameters are changed slightly and the output variables reevaluated. From the change in the output variables the new set of parameters replaces the previous quantities with a probability governed by a selection rule that favors transitions to less probable states. While the pdf of the output states (whose dimension equals the number of output variables) is not known in advance, an initial estimate can be obtained through a Monte-Carlo simulation. Subsequent iterations of the multicanonical method lead to increasingly accurate estimates of the tails of the pdf [25], [26].

To illustrate, consider a variable X that is a function of several random variables expressed as a vector $\vec{v} = (v_1, v_2, v_3 \dots)$ and we wish to calculate the pdf $p(X)$. We start by generating a histogram for X over all the possible expected values of X . For example, if we want to calculate the pdf of the

DGD of a 15 section fiber, each section with a DGD of 1ps, the minimum and maximum values of X are going to be 0 and 15ps respectively. Therefore a histogram $H(X)$ is constructed for values of $0 \leq X \leq 15ps$ each entry containing one and we set $p_0(X)$ to be the normalized version of $H(X)$. We employ a Metropolis Hastings algorithm [19] to obtain samples using the biased pdf given by $1/p_0(X)$ and update the histogram $H(X)$ accordingly. The set of obtained samples form a new histogram $H_0(X)$. A new pdf, $p_1(X)$, is obtained by normalizing $H_0(X)p_0(X)$. The procedure is repeated using a new histogram $H(X)$ with all unity entries but with a biased pdf given by $1/p_1(X)$. Again, a Metropolis Hastings algorithm is used to update $H(X)$ based on the samples obtained using the algorithm to obtain a new histogram $H_1(X)$. This iteration generates a new pdf given by normalizing $H_1(X)p_1(X)$. The procedure is again repeated using the new biased pdf until the desired reach in $p(X)$ is obtained.

1.6 Organization of the Thesis

The structure of the thesis is as follows: Chapter 1 introduced the background knowledge of PMD. In Chapter 2, we present measurements of very rapid time variations of the SOP at the output of DCM's and a physical model to describe these changes. In Chapter 3, we propose a technique to determine the position of temporal polarization activity i.e. a signal fast hinge in a WDM link and test it using simulations. In Chapter 4, we study the ability and compare between various linear predictions models applied to the DGD for a link within a deterministic hinge model. In Chapter 5, we derive analytical expressions for the SOP and PMD vector temporal autocorrelation functions using a reasonable stochastic model for the hinges. Finally in Chapter 6, we employ the multicanonical method to evaluate the SER as a function of the SNR in a WDM system taking into account fiber nonlinearities and PMD.

Chapter 2

Measurement and Simulation of Polarization Transients in Dispersion Compensation Modules

2.1 Introduction

Coherent optical communication systems require state of polarization (SOP) tracking to enable accurate demultiplexing of polarization multiplexed data [3]. High speed polarization changes in excess of the receiver's tracking capacity can adversely affect the performance of the PMD compensator and therefore yield a system outage. Electronic polarization tracking at least partly avoids such failures; for example, an FPGA (field programmable gate array) based system has been developed that provides endless tracking up to speeds of 9 krad/s [27].

Rapid polarization changes have long been known to result from mechanical impacts on patch cords and fibers spools when these components fall certain distance or are struck with a metallic tool such as a screwdrivers [28]. Since devices such as DCMs are placed in huts where they can experience similar mechanical impacts, the accurate characterization and modeling of such effects is necessary to predict even qualitatively the frequency and duration of outages. Based on these observations, a procedure that generates high frequency transients in the SOP and DGD (differential group delay) of a DCM was described in [29]. However, the procedure of [29], which was later transferred to industry and described in more detail in [30], employs a carefully calibrated system to release a steel ball on the DCM casing. This method was unexpectedly found to produce nearly identical PMD traces when the experiment was repeated, which enabled the determination of time-dependent quantities such as the DGD in an extremely convenient fashion. Additionally, we observed angular velocities (AV's) for the Stokes parameters over the Poincare sphere as large as 550 krad/s, which are far higher values than those that could be compensated by the system described in [27].

In this chapter we examine, both theoretically and experimentally, rapid polarization transients generated by mechanical impacts on DCMs. Experimentally we can measure AV's of the Stokes vector over the Poincare sphere exceeding 100krad/s for all measured DCMs. Theoretically, we demonstrate with a simple model for the excitation that the patterns of the AV observed in experiments can be reproduced through simulation and that the amplitude of the AV increases with the volume of the fiber affected by the impact. Our model is sufficiently simple to be employed in system simulations.

2.2 Experimental Setup and SOP Reproducibility

To examine the time evolution of the rotation matrix describing the output state in DCM's, we employed the experimental setup of Figure 2.1. Light from a Photonetics Tunics-PRI tunable laser is first passed through an inline polarization controller that selects the input polarization which is then inserted into a Nortel DCM10, DCM20, DCM40 or DCM100 dispersion compensation module, all of

which possess identical casings. The output SOP was measured with a Lightwaves2020 inline polarimeter and sampled at a 2.5×10^6 Hz data rate with a Tektronix digitizing oscilloscope. The polarimeter was calibrated against an HP 8509B polarization analyzer. Successive measurements of the SOP Stokes vector on the Poincaré sphere, $\hat{s} = [s_1 \ s_2 \ s_3]^T$, coincided to within a mean angular error of 0.07° and standard deviation of 0.050° with the HP 8509B. The magnitude of the error further remained $< 0.22^\circ$ over 200 independent measurements.

The experiment is performed as follows. A random input polarization is launched into the DCM with the input polarization controller. A 2.6 g, 1.3 cm diameter steel ball is dropped on the DCM casing and the resulting output polarization transients recorded. The experiment is repeated for several impact locations at heights of 3.0, 6.0 and 9.0 cm. The positions of the different impact locations are shown in Figure 2.2. We have found that if this experiment is repeated several times under identical conditions, provided that the difference in measurement times does not exceed the time scale associated with the slow thermal drift of the system (typically 10min-1hr), the output polarization follows nearly identical paths on the Poincare sphere. In Figure 2.3, we display the results of 10 measurements of the time variation of the s_1 component of the light output from the DCM40, for a ball drop height of 9cm and drop position number 1. Apart from a small time shift between consecutive measurements, the plots of s_1 are almost identical. This time shift in the curves results from fluctuations in the trigger to the oscilloscope which was provided by monitoring the audio signal generated from a microphone resting on the DCM. Defining a measure of reproducibility

by the mean squared error $\frac{1}{T} \int_0^T \sqrt{(s_1^1 - s_1^{10})^2 + (s_2^1 - s_2^{10})^2 + (s_3^1 - s_3^{10})^2} dt$ in which \hat{s}^1 and \hat{s}^{10} are the

first and tenth measured polarizations and $T = 25.0$ ms is the record length, yields the error versus ball drop height shown in Figure 2.4 at position 1 for the different DCMs under consideration. This error is far smaller in magnitude than the amplitude of the polarization fluctuations. Releasing the ball from a height exceeding 12cm or delaying successive measurements by a long time interval (for our setup approximately 24 hours) over which thermal drift is substantial, induces a substantial change in the internal state of the DCM changes so that the SOP traces a different path on the Poincare sphere for identical measurement conditions.

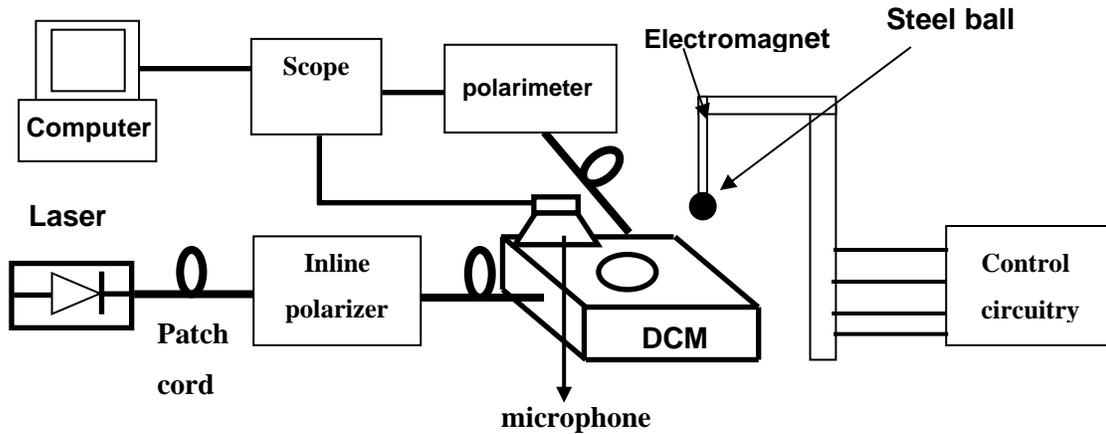


Figure 2.1. In our experiments, a steel ball is secured and later released onto a DCM by an electromagnet generating fast polarization changes. The acoustics signal from the impact is detected by a microphone and employed to trigger the scope, which then samples the output from a fast polarimeter.

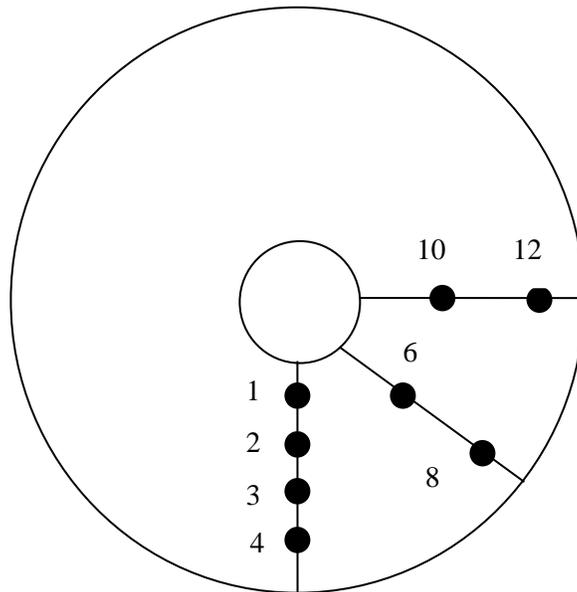


Figure 2.2. Different impact positions on the DCM as follows: points 1-4 are located 1.5 cm, 3 cm, 4.5 cm and 6 cm from the inner core. Points 6 and 8 and 10 and 12 are at the same radial distances as points 2 and 4 but displaced at 45° and 90° angles.

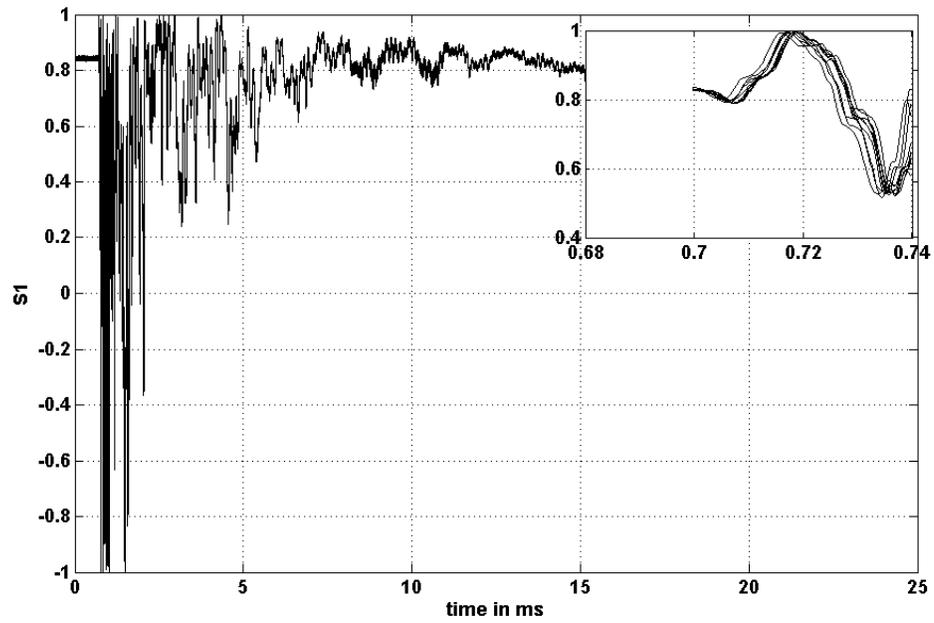


Figure 2.3. Ten successive measurements of the s_1 component. The inset displays the region from 0.7 ms to 0.74 ms showing the individual measurement curves.

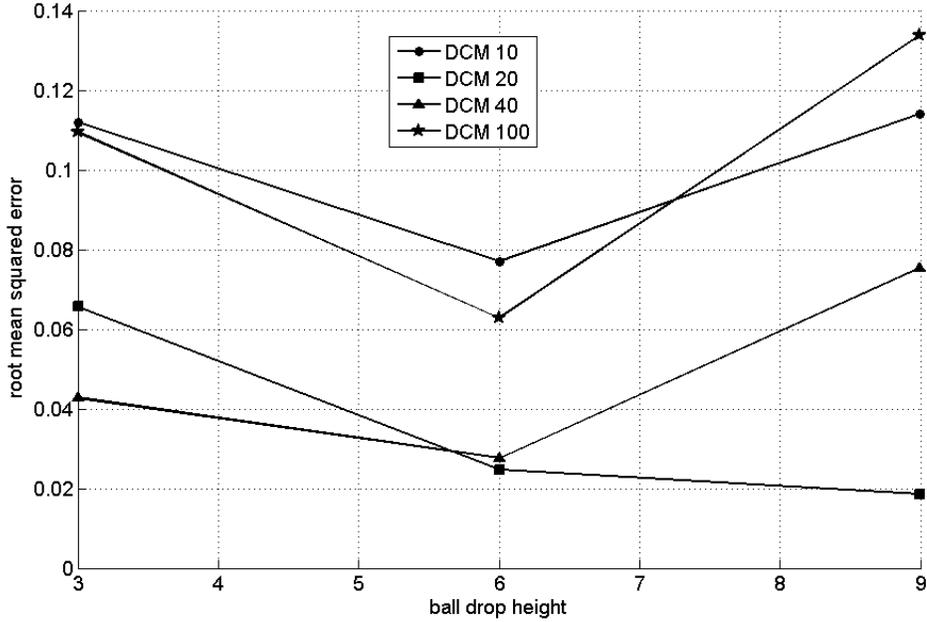


Figure 2.4. The mean squared error of the polarization trace versus drop height for different DCMs for impact location 1. The error is largest for a height of 9cm.

2.3 Birefringence Modeling

The time dependence of the output Stokes parameters can be determined theoretically if the birefringence of the fiber is known as a function of space and time. As a result, our measurement results for the time dependent behavior of the output polarization can be interpreted, at least qualitatively, with a simple theoretical model for the birefringence.

The underlying equations for a model of polarization evolution are derived from the relationship between the variation in the SOP, \hat{s} , with distance, z along the optical fiber in terms of the local birefringence vector $\vec{\beta}(z, t)$ as in Eq. (1.45), but now with time included as

$$\frac{\partial \hat{s}(z, t)}{\partial z} = \vec{\beta}(z, t) \times \hat{s}(z, t) \quad (2.1)$$

Consequently, the output Stokes vector, $\hat{s}_{out} = \hat{s}(L, t)$, where L is the DCM fiber length, can be estimated given a birefringence model and input polarization. The relation between \hat{s}_{out} and the input

Stokes vector \hat{s}_{in} is described by a rotation matrix R such that $\hat{s}_{out} = R \hat{s}_{in}$, which is the quantity that is measured experimentally. In terms of this rotation matrix, the AV vector $\vec{\omega}$ is determined from

$$\vec{\omega} \times = \frac{dR}{dt} R^T \quad (2.2)$$

as demonstrated in [31]. The amplitude of the AV vector, $|\vec{\omega}|$, determines the rate at which the output polarization varies and is independent of the input polarization. Therefore, this quantity is a property of the DCM under test and is directly related to the dynamics of the problem. The AV is related to the birefringence vector in an analogous manner to the relationship between the PMD vector and the birefringence vector, namely

$$\frac{\partial \vec{\omega}(z, t)}{\partial z} = \frac{\partial \vec{\beta}}{\partial t} + \vec{\omega}(z, t) \times \vec{\beta} \quad (2.3)$$

Here $\vec{\omega}(z, t)$ is the angular velocity at point z and time t where we simulate or measure $\vec{\omega}(L, t)$. Since the term $\vec{\omega} \times \vec{\beta}$ represents a rotation, the principle source of the AV is the time derivative of the birefringence.

Our fiber model separates the birefringence into two components: the steady-state birefringence before the mechanical impact and a time-dependent part generated by the acoustic wave which travels along the fiber away from the position of the mechanical impact. The unperturbed birefringence of the DCM is simulated by a random modulus model according to the Langevin equation [32], [33]

$$\frac{d\beta_i}{dz} = -\frac{\beta_i}{L_c} + g(z), i = x, y \quad (2.4)$$

in which $g(z)dz$ is a Wiener process and L_c is the birefringence correlation length. The standard deviation of the Wiener process is $\sigma = \sqrt{b^2/L_c}$, where the beat length is defined by $L_b = 2\pi/b$.

To describe the time dependent component of the birefringence generated by the impact, we observe that torsional, longitudinal and flexural or bending waves can be present in a cylindrical rod [34], [35]. Dispersion curves for the sound velocity in a cylindrical rod have been derived for the three wave types as a function of the dimensionless parameter fa/c_t , where f is the acoustic frequency, a is the radius of the rod and c_t is the bulk transverse velocity [34]. For a fiber radius of

125 μm , which includes the plastic acrylate jacket and a sound velocity of $c_t = 3.37 \times 10^3 \text{ m/s}$ [36], we obtain $fa/c_t = f(3.7 \times 10^{-8})$. We estimate f as $v/2H$ where v is the speed of sound in the fiber with jacket and H is the height of the DCM, yielding $fa/c_t = 1.7 \times 10^{-3}$. Hence, for the range of acoustic frequencies resulting in our experiments from the mechanical impacts on the DCM casing, $fa/c_t \ll 1$. In this case, only the fundamental bending, longitudinal and torsional modes can propagate as all higher orders are cutoff [34]. However, a torsional wave requires that the fiber be twisted by the impact, which is precluded by the DCM casing. Further, bending waves possess a significantly smaller propagation velocity compared to a longitudinal wave from an analysis of the dispersion curves. However, we will demonstrate later that AV magnitudes of the order of the experimentally observed values imply that a large fiber length is excited immediately after impact, which requires a large propagation velocity. That is, consider the DCM geometry in Figure 2.5 with the casing dimensions (height, inner and outer radii) indicated in the figure. When the ball impacts the DCM at a certain radius r , the part of the fiber coils immediately beneath the area of impact begin to oscillate. Initially only the top few coils are excited while subsequently the excitation propagates vertically downwards vibrating lower fiber coils. Simultaneously, the wave also travels horizontally along the fiber length. We estimate the time taken by the wave to travel vertically to the bottom most wrap to be less than $4 \text{ cm}/3645 \text{ m/s} = 0.01 \text{ ms}$, where 3645 m/s is the speed of sound in the fiber taking the jacket into consideration [36]. Within this interval the measured AV's are relatively small. After impact, the wave also travels around the circumference of the fiber coils, traversing half of each coil for $r = 10 \text{ cm}$ in 0.086 ms for a longitudinal wave which travels at the speed of sound. Indeed, the experimentally measured polarization activity is significant over time scales of this order. However, these figures are qualitative in nature and only give an estimate of the order of magnitude of the physical effects. Examining the dispersion curves in [34], a bending wave requires roughly 0.5 ms to propagate half-way around the circumference of the fiber spool for $fa/c_t = 10^{-2}$ which is an order of magnitude less than the estimated value. Thus the actual time is going to be even greater than that. Most of the significant oscillations we observed in experiment occur within the first 1 ms , hence a bending wave cannot significantly contribute to the experimentally observed AV.

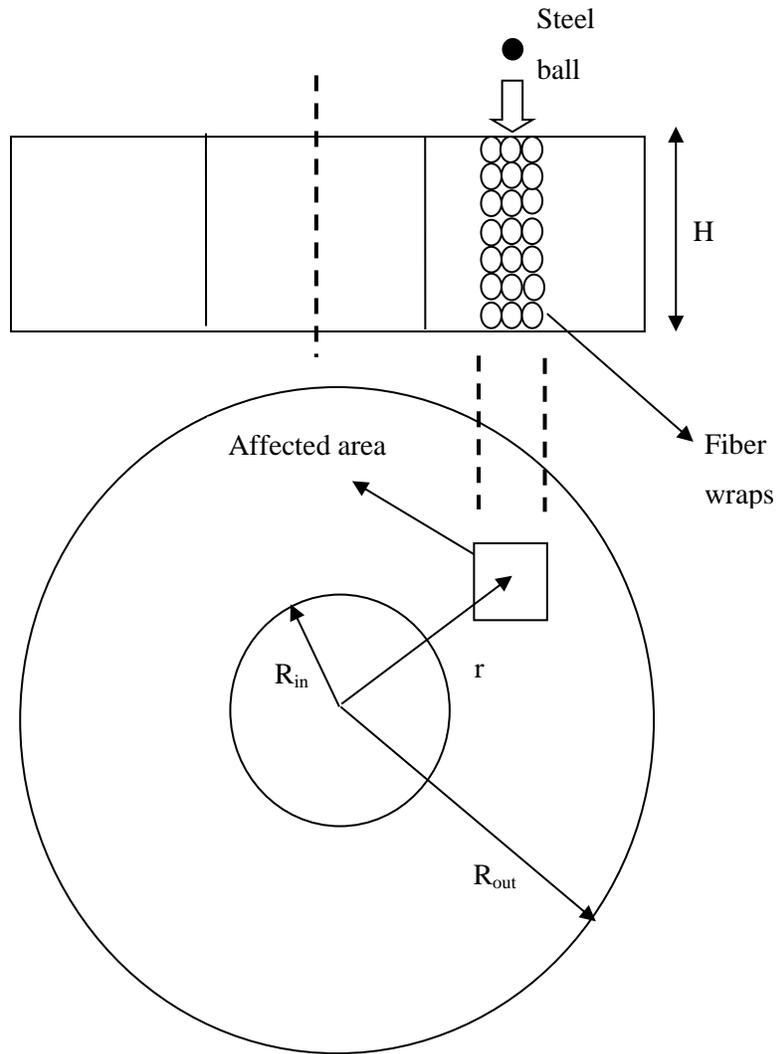


Figure 2.5. The geometry of the DCM. Our model assumes that the steel ball impacts a certain area of the DCM resulting in the oscillation of the fiber coils beneath it. Here $R_{in}=5.25$ cm and $R_{out}=12.4$ cm. $H = 4$ cm.

Starting with the above general discussion, we now model the time dependent birefringence generated by a longitudinal wave. A longitudinal wave traveling along a bent fiber induces a linear birefringence. At a wavelength of $1.55 \mu\text{m}$, this is given by [37]

$$\delta\beta(z,t) = -5.11 \times 10^5 \left(\frac{a_{cl}}{r} \right)^2 - 1.8 \times 10^6 \frac{a_{cl}}{r} \frac{\partial u_z(z,t)}{\partial z} \quad (2.5)$$

Here a_{cl} and r are the cladding radius and radius of curvature and $\partial u_z / \partial z$ is the axial tensile strain. The first term of Eq. (2.5) is the bending-induced birefringence from [38]. The time-dependent part of the birefringence is accordingly

$$\delta\beta(z,t) = -1.8 \times 10^6 \frac{a_{cl}}{r} \frac{\partial u_z(z,t)}{\partial z} \quad (2.6)$$

We model the displacement u_z from the equilibrium position by

$$\begin{aligned} u_z &= q(t)F(z) = q(t) \sin\left(\frac{2\pi z}{\lambda}\right) W(z) \\ &= q(t) \sin\left(\frac{2\pi z}{\lambda}\right) \frac{(\tanh(z - z_{i1}) - \tanh(z - z_{i2}))}{2} \end{aligned} \quad (2.7)$$

in which λ is the excitation wavelength, $q(t)$ is a function describing the time dependence, z_{i1} and z_{i2} define the z coordinates of the fiber over which the wave is excited, hence the active length of the fiber would be $L_{act} = z_{i2} - z_{i1}$. The hyperbolic tangent function effectively represents a window function that defines the extent of the excited length and therefore yields a smooth transition of the birefringence between the vibrating and stationary portions of the fiber. Since we cannot examine the internal structure of the DCM, we neglect the spreading of the wave in our model and instead consider the excited length as the average length that is excited over the duration of the oscillations. However, by varying this length parameter from a small value to $L_{act} = L$, which corresponds to equal excitation of the entire fiber length, we can examine the variation of the characteristic time response with the physical extent of the excitation. Large values of the length parameter correspond to case of coupling of vibrational energy between a large area of the DCM casing and the fiber while a small value of the parameter instead indicates that the excitation is damped rapidly as it passes through the fiber coils or does not acquire energy from a large fraction of the DCM casing. The axial strain satisfies the damped wave equation [35]. Hence the function $q(t) \sin\left(\frac{2\pi z}{\lambda}\right)$ is taken as a solution to the damped wave equation [35]

$$\frac{\partial^2 f}{\partial z^2} = \frac{1}{v_g^2} \frac{\partial^2 f}{\partial t^2} + b \frac{\partial f}{\partial t} \quad (2.8)$$

in which v_g is the wave velocity and b the damping coefficient. Substituting $f(t) = q(t) \sin\left(\frac{2\pi z}{\lambda}\right)$ in (2.8) yields

$$\ddot{q} + 2\xi\omega_o\dot{q} + \omega_o^2 q = 0 \quad (2.9)$$

With $\omega_o = 2\pi v_g / \lambda$, $\xi = b v_g \lambda / 4\pi = \lambda / \lambda_o$ and $\lambda_o = 4\pi / b v_g$.

Eq. (2.9) admits underdamped, overdamped and critically damped solutions. With the initial condition $g(0) = 0$, we have, $q(t) = A(e^{-\alpha_1 t} - e^{-\alpha_2 t})$, $\alpha_1 = \omega_o(\xi + \sqrt{\xi^2 - 1})$ and $\alpha_2 = \omega_o(\xi - \sqrt{\xi^2 - 1})$ for an overdamped system ($\xi > 1$) and A is the vibration amplitude. We will employ these formulas since our simulations demonstrated that underdamped systems exhibit a periodicity in the AV that is not present in our experimental results.

To estimate λ_o , we note that the attenuation coefficient of longitudinal sinusoidal waves in a fiber was demonstrated experimentally to possess the nearly frequency-independent value of 0.8m^{-1} for small fiber losses [36]. Eq. (2.8) can then be employed to estimate b by Fourier transforming Eq.(2.8) and solving for f in the limit that the loss is small compared to the propagation constant. This yields $b v_g = 1.6\text{m}^{-1}$ and $\lambda_o = 7.9\text{m}$.

It remains to model the static fiber bending radius. We approximate this as a linear function of z since the DCM length extends over a few km and the radii are on the order of cm. We denote the radius of curvature at the beginning of winding by R_1 and that at the end of winding by R_2 . For a fiber of length L we have

$$r(z) = R_1 + \frac{(R_2 - R_1)}{L} z \quad (2.10)$$

Denoting the time independent local birefringence vector by $(\beta_x(z), \beta_y(z), 0)$ and the birefringence vector due to bending and wave excitation by $(\beta_p, 0, 0)$, the total birefringence vector is now

$$\vec{\beta} = \begin{pmatrix} \beta_x(z) - \beta_p \\ \beta_y(z) \\ 0 \end{pmatrix} \quad (2.11)$$

The components $\beta_x(z)$ and $\beta_y(z)$ are determined through Eq. (2.4) while β_p is given according to Eqs. (2.5) and (2.7) as

$$\beta_p = 5.11 \times 10^5 \left(\frac{a_{cl}}{r} \right)^2 + 1.8 \times 10^6 A \left(e^{-\alpha_1 t} - e^{-\alpha_2 t} \right) \frac{a_{cl}}{r} \frac{\partial F}{\partial z} \quad (2.12)$$

in which F is defined as in Eq. (2.7).

2.4 Measurement and Simulation Results

2.4.1 S Parameters and Angular Velocity

We now compare measured and simulated results for the AV and the S parameters. Here we first release a ball at position 1 and height 3cm and determine the rotation matrix R by repeating the experiment for 10 different input polarizations. A least squares technique based on singular value decomposition (SVD) was employed to obtain the elements of the rotation matrix. We display our resulting calculated values of the AV for our DCM10 to DCM100 components in Figure 2.6- Figure 2.9. Evidently the DCM40 exhibits the largest AV magnitudes of all the DCMs. Further, the AV increases with release height for all DCMs as expected, reaching a maximum value of 550 krad/s for the DCM40, impact position 1 and release height of 9 cm.

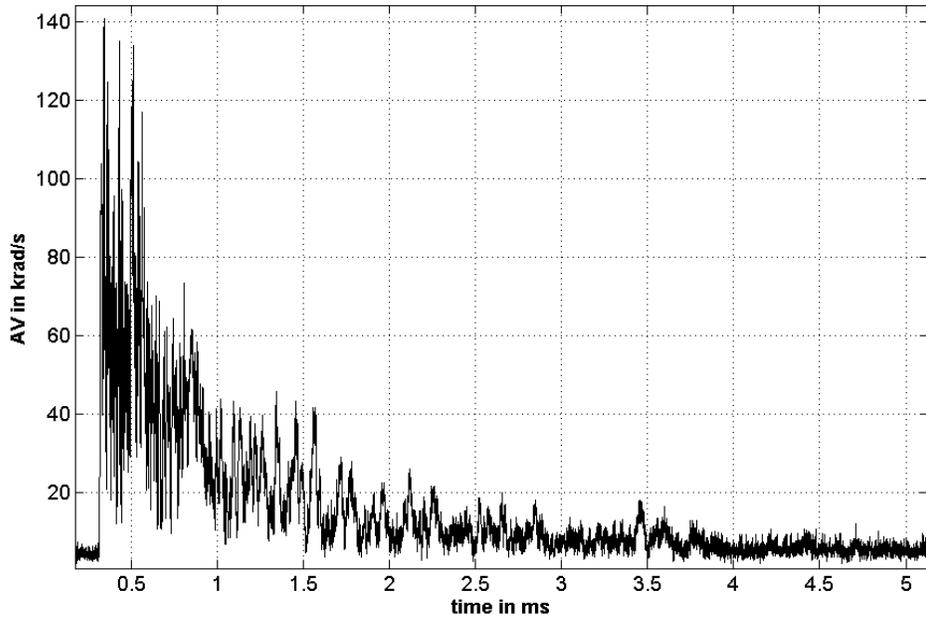


Figure 2.6. The AV for DCM10 with impact location 1, and release height 3 cm.

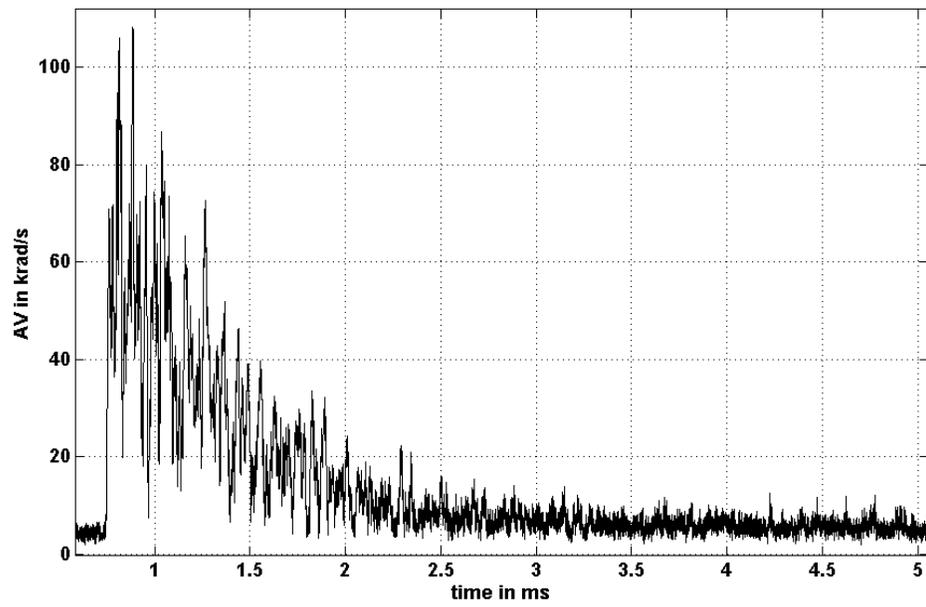


Figure 2.7. The AV for DCM20 with impact location 1, and release height 3 cm.

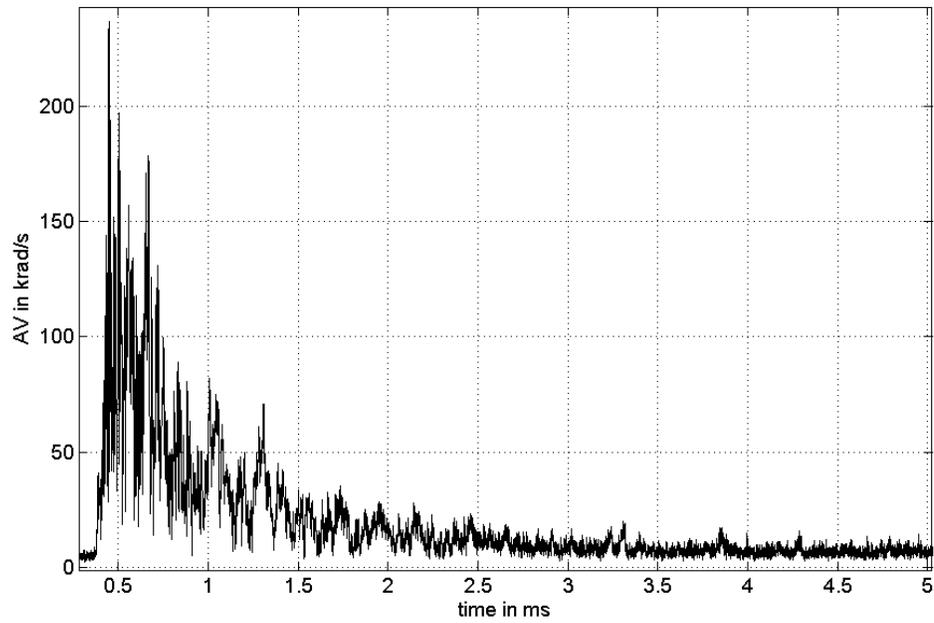


Figure 2.8. The AV for DCM40 with impact location 1, and release height 3 cm.

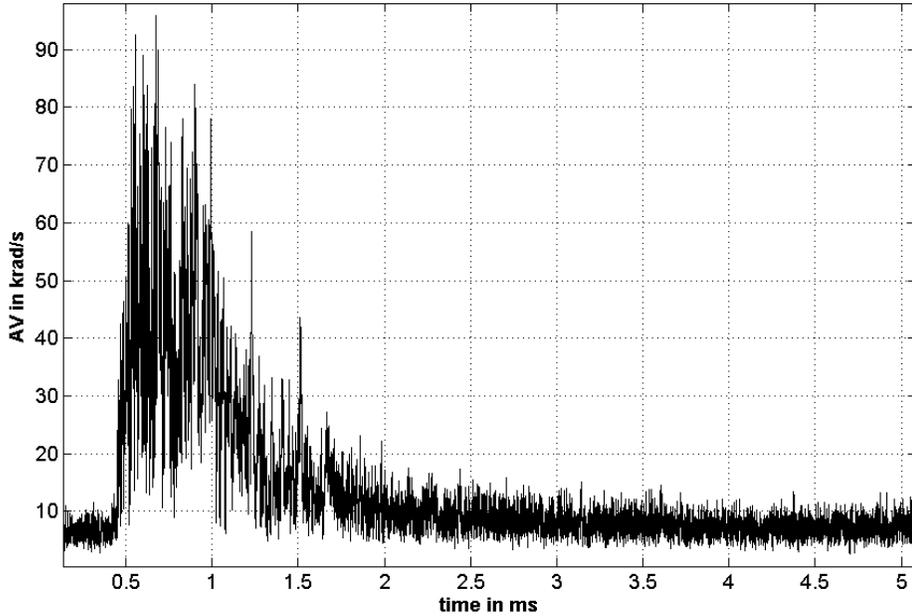


Figure 2.9. The AV for DCM100 with impact location 1, and release height 3 cm.

We next present for comparison simulated AV curves. The parameters employed are a wave velocity $v_g = 10^3 \text{ m/s}$, $\lambda_o = 10.0 \text{ m}$, $\lambda = 50.0 \text{ m}$ (and hence $\xi = 5$), $a_{cl} = 125 \mu\text{m}$, $R_1 = 5.5 \text{ cm}$, $R_2 = 9.0 \text{ cm}$, $A = 2.5 \text{ cm}$ and a fiber length $L = 1 \text{ km}$. Since we are concerned with the region of large AV, we only carry out simulations for the first 5 ms where the time derivative of the birefringence attains its largest values. The S parameters, which respond to the time scale of the birefringence rather than its time derivative, display oscillations after impact that decay much more slowly than those of the AV. Indeed, a comparison Figure 2.6 - Figure 2.9 with Figure 2.3, demonstrates that while AV decays after approximately 2ms, the S parameters continue to oscillate for a substantially longer time. We consider three different fibers, fibers A, B and C with different values of the steady-state birefringence and therefore different solutions of Eq. (2.4) but with identical beat and correlation lengths, $L_b = 15 \text{ m}$ and $L_c = 20 \text{ m}$, respectively, which are typical values for communication fibers [39]. Since noise effects are absent in the simulations, three orthogonal input polarizations are employed after which R and hence the AV is determined. The SVD technique is again employed to obtain the rotation matrix. In Figure 2.10, we display the simulated AV for fiber B, for $L_{act} = 600 \text{ m}$ and $z_{i1} = 100 \text{ m}$. Evidently, our simple model for the excitation yields an AV that displays the salient features of the experimental results. The model appears well suited to system simulations since experimental behavior can be reproduced by varying only a small number of parameters. In Figure 2.11, we simulate the AV for $L_{act} = 1.0 \text{ km}$, which corresponds to the entire fiber length. The magnitude of the AV corresponds to the largest value expected from a DCM with the parameters of the simulation.

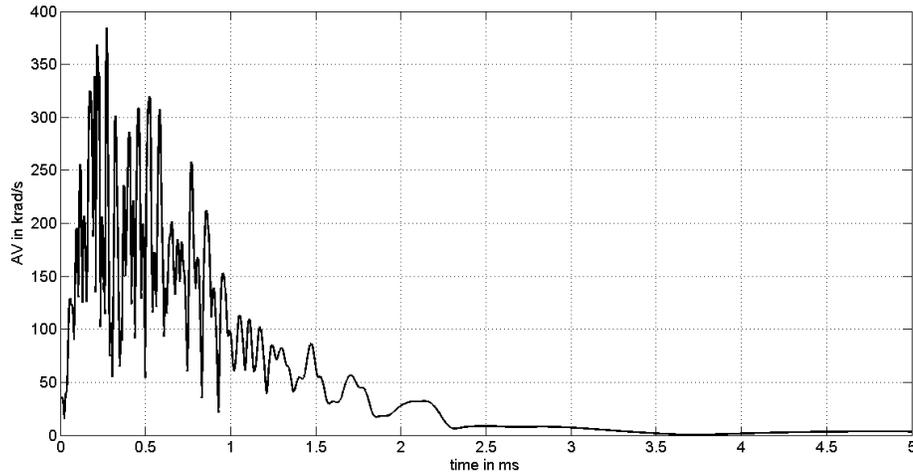


Figure 2.10. The simulated AV of fiber B for an active length of 600m, $z_{i1}=100\text{m}$.

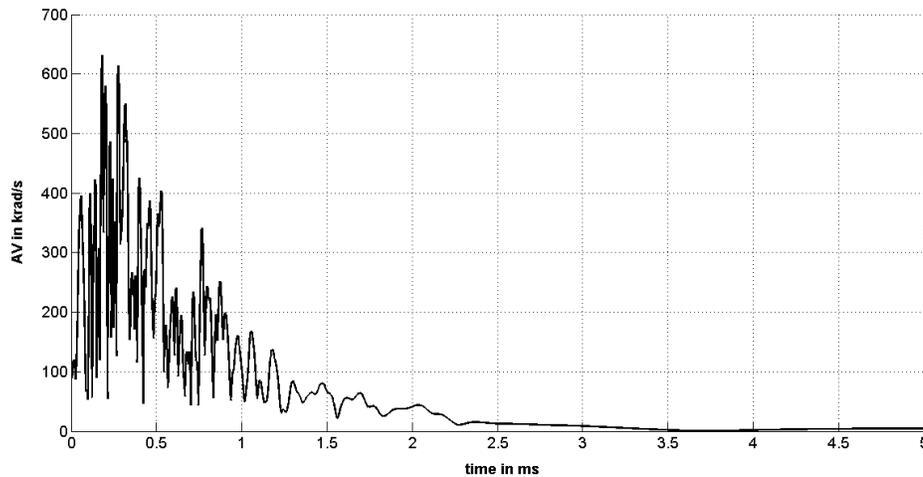


Figure 2.11. The simulated AV for fiber B, for active length equal to the total fiber length.

We finally simulate in Figure 2.12 the S parameters of fiber C for a 20 m active length. The range of the S parameters obviously does not agree with that observed in the experiments. As well, the maximum AV is calculated to be 18 krad/s. However, the simulated S parameters ranges and AV results were in reasonable agreement with experiment for active region lengths in excess of 100 m, yielding maximum AV's in the 100 krad/s range for an active length of 100 m. Hence, we conclude that the length of the fiber that must be excited at the beginning of oscillation must be relatively large which cannot be accomplished simply by the propagation of bending waves.

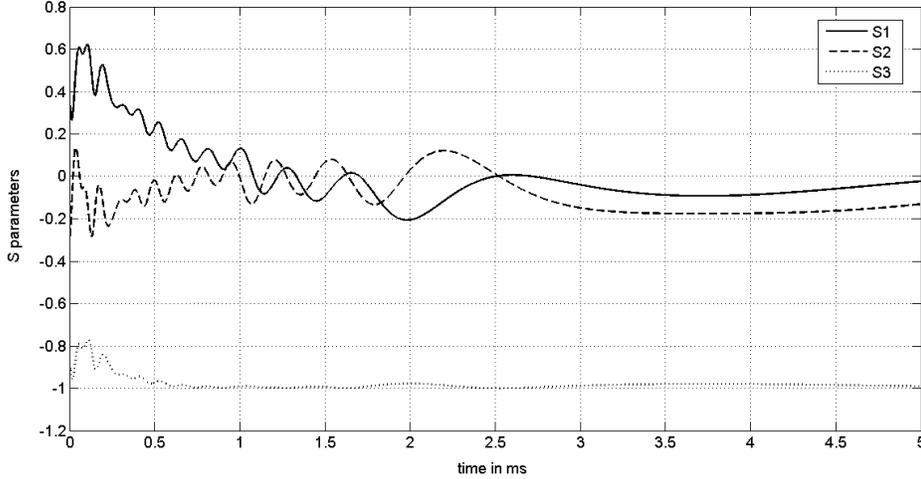


Figure 2.12. The S parameters for fiber C with an active length of 20 m and $z_{i1}=200$ m. The fluctuations are clearly very small compared to the experimental measurements (figure 2.3).

2.4.2 Variation with Fiber Length and Impact Position

We finally investigate both theoretically and experimentally the influence of the active fiber length and impact position on the AV. Our comparison is based on a parameter we term the arc length defined for a given value of the angular velocity, ω_o , as

$$a_l(\omega_o) = \int_0^T |\vec{\omega}| \theta(|\vec{\omega}| - \omega_o) dt \quad (2.13)$$

where θ is the Heaviside step function. Thus a_l corresponds to length scanned by the output polarization over the Poincare sphere for angular velocities greater than a certain value. While the fraction of the arc length covered at low angular velocities increases with T , the time interval $[0, T]$ can be restricted to the physically interesting region of rapid oscillations and large angular velocities shortly after the initial impact. In our simulations, we employ $T=5$ ms while in our measurements, $T=12$ ms. In Figure 2.13, we plot a_l as a function of ω_o for fiber B. The simulated curves are for $L_{act}=600$ m with $z_{i1}=100, 200$ and 300 m and $L_{act}=100$ m, with $z_{i1}=200, 500$ and 800 m. Evidently, a_l depends on both the amount of the active fiber length that is excited and the excitation position with the maximum AV monotonically increasing as a function of this length. Further, the maximum angular velocity is observed to vary with excitation position (z_{i1}). The results for the fibers A and C

are qualitatively similar. Our experimental results for a_l are displayed in Figure 2.14 for two different ball drop positions, 6 and 8 for the same DCM. The arc length is clearly greater for the inner point 6 compared to the outer point 8. Similar results also have been observed for points 10 and 12, independent of the DCM type and the release heights. This leads us to conclude that a greater length of fiber is excited if the ball is dropped near the inner core of the DCM, which may be the result of reflections of the vibrational motion from the solid inner core of the spool as well as the transmission of the acoustic wave through the annular metal region before entering the fiber. A detailed mechanical analysis of the system, which we cannot undertake at present since the DCM cannot be opened nondestructively, would presumably show the precise evolution of the acoustic wave and enhance the accuracy of our predictions in this regard, but such a calculation is outside the scope of our work.

By varying the active length of fiber B from 100m to its full length in our simulation, we found that the maximum AV increased from ~ 100 krad/s to 620krad/s. At the same time, when the release position was altered experimentally for a release height of 9cm, the maximum AV values were observed to vary between 153krad/s and 550krad/s for the DCM40. Thus our model provides a possible explanation of this variation as resulting from the change in the excited fiber length with release position. We have finally examined the arc length of all the DCMs, yielding results such as those of Figure 2.15. The DCM40 displays the largest arc length for all positions and heights employed although only a single position is considered in the figure. The difference between the different DCM behavior indicates the existence of different active lengths and might possibly yield information about the performance fabrication tolerances of a DCM device.

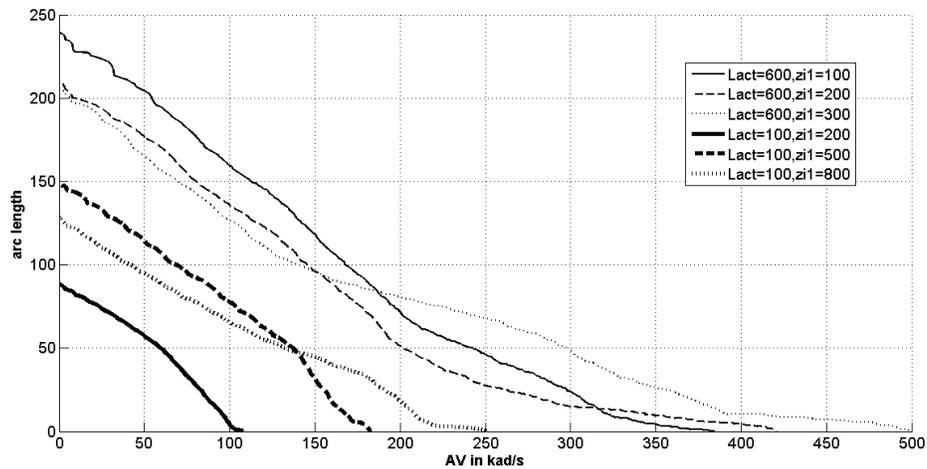


Figure 2.13. The dependence of the arc length on the angular velocity for different lengths of fiber B.

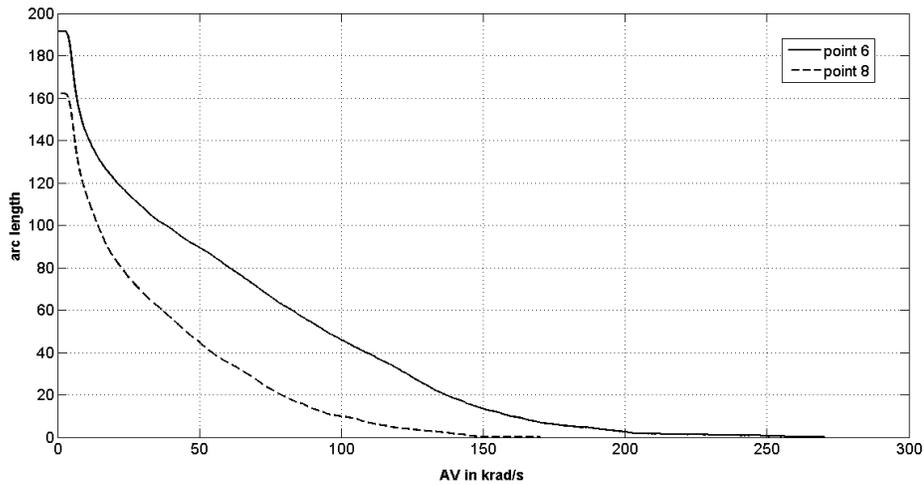


Figure 2.14. The dependence of the experimental arc length on the angular velocity for DCM10 with a ball release height of 9 cm.

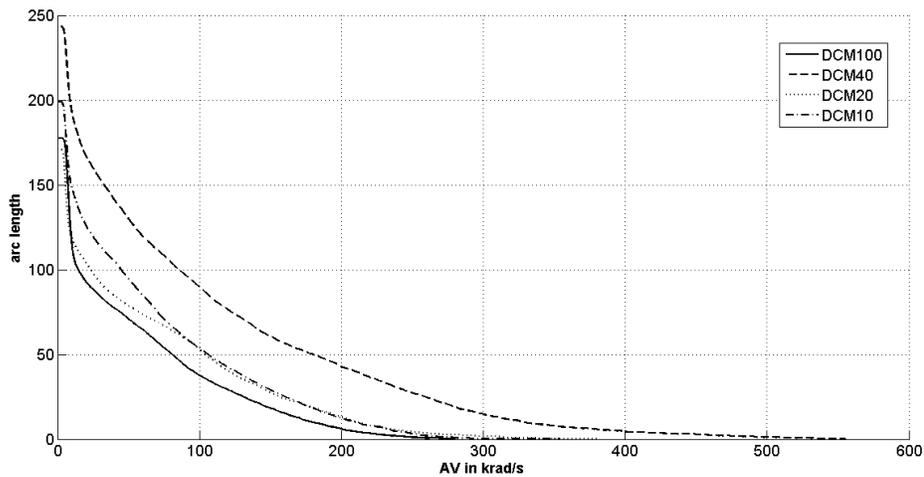


Figure 2.15. The experimental arc length for impact location 1, a release height of 9 cm, and different DCM structures.

2.5 Conclusion

We have developed a model that describes, at least qualitatively, impulsive mechanical impacts on DCM's generated by acoustic perturbations to the fiber birefringence. We have identified the source of rapid changes in polarization as a longitudinal wave propagating outward from a restricted region

of the fiber spool. This model is based on well-established physical considerations and yields results that at least qualitatively agree with our experimental observations which indicate that the maximum value of the AV increases monotonically with the active fiber length. We have employed a simple parameter, the arc length, to characterize the polarization activity on the Poincare sphere as a function of the AV. An examination of this parameter demonstrated that both the active length and the position of the initial excitation affect the polarization activity. Further, the procedure possesses a sufficiently flexible and simple structure adaptable to any DCM device and subsequently can be employed in system evaluations of outage probabilities.

To be able to implement more accurate models, knowledge of the mechanical properties such as the speed of the acoustic wave, the amplitude of the mechanical oscillation, the mean fiber bending radius and the number of wrappings are required in order to investigate the evolution of the propagating acoustic wave inside the DCM in detail. Unfortunately, this is not feasible since the DCM's under consideration cannot be opened non-destructively. As well, the statistical properties of the steady-state fiber birefringence would have to be known, presumably with a back-reflection measurement [39]. Consequently, a heuristic model such as that given here is likely optimal for system simulations.

Chapter 3

Locating Regions of Polarization Activity in WDM Systems

3.1 Introduction

PMD temporal dynamics has been studied extensively in two contexts. Long term measurements of PMD on buried cables have been carried out by several researchers [40], [41] leading to detailed models that have been extensively compared to experiments [42], [43]. It was found that the PMD dynamics of buried cables is well described by a Langevin equation [43]. However, as mentioned in Chapter 1, measurements of installed systems demonstrate that the main sources of time variations arise in regions where the fiber is exposed to the environment or at the location of huts containing devices such as DCM's [44], [45], [46]. Consequently, the hinge model has been put forward. The hinge model incorporating fiber sections with fixed DGD and isotropic polarization scattering at each hinge between these sections demonstrated non-Maxwellian DGD statistics [47]. Further analysis showed that another consequence of the isotropic hinge model is that in WDM systems, some channels remain outage free [48].

Since the DGD statistics are highly dependent on the DGD of the sections between the hinges, knowledge of the hinge positions provides information regarding the temporal behavior of the DGD. Measurements however indicate that DCM's are not entirely isotropic [46] and hence DGD statistics obtained under an isotropic assumption are approximate. In fact, a model with a fixed rotation axis and a sinusoidally varying hinge rotation angle can be employed to approximate the DGD spectra measured over a finite interval of time [46].

A source of time variation that yields rapid changes in polarization and can therefore similarly be considered as a hinge is associated with mechanical impacts on DCM's as we described in Chapter 2. Measurement methods for characterizing and locating such sources of polarization fluctuations include polarization OTDR (optical time domain reflectometry) [49], which however is unreliable for large PMD values [49] and the procedure of [50] for measuring the instantaneous DGD of fiber subsections in a link.

In this chapter, under the assumption that the rapid polarization changes originate from a single spatial location within each time interval, we propose an alternative measurement procedure that yields the amount of DGD in the system preceding and following this position. This technique is particularly applicable to the detection of fast and short-lived polarization changes that are localized to a small area of the link. However, it can also be employed to determine the location of more slowly varying hinges in a several hinge link if only one hinge is active at one time or if one hinge exhibits a far more rapid variation than the other hinges. The method is otherwise independent of the hinge temporal or spectral behavior.

3.2 Theory

A rapid fluctuation of the output polarization resulting from a physical event at a given point along an optical fiber link can be regarded as a polarization hinge activity at this point, as shown in Figure 3.1.

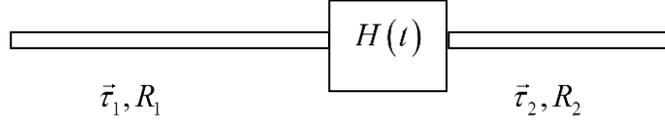


Figure 3.1. A rapidly varying localized polarization variation modeled as a hinge time varying rotation matrix $H(t)$. The fiber before the hinge is section 1, with PMD vector and rotation matrix $\vec{\tau}_1$ and R_1 while section 2 follows the hinge with PMD vector and rotation matrix $\vec{\tau}_2$ and R_2 .

The fiber lengths before and after the hinge will here be termed sections 1 and 2, respectively, where we assume that these fibers are sufficiently long that the mean DGD varies as the square root of their length and is Maxwellian distributed over long periods of time as describe in Chapter 1. If the physical environment of sections 1 and 2 is constant over the characteristic time of the polarization variation at the hinge, the PMD vector and the Stokes vector at the output are respectively described by the concatenation rules

$$\vec{\tau} = \vec{\tau}_2 + R_2 H(t) \vec{\tau}_1 \quad (3.1)$$

$$\hat{s}_{out} = R_2 H(t) R_1 \hat{s}_{in} \quad (3.2)$$

where $\vec{\tau}_1$ and $\vec{\tau}_2$ represent the three-element PMD (column) vectors of sections 1 and 2 respectively, $H(t)$ is the time-dependent rotation matrix of the hinge, t being the time, R_1 and R_2 denote the rotation matrices of sections 1 and 2 respectively \hat{s}_{in} and \hat{s}_{out} are the input and output polarizations respectively.

Assuming that the time scales of the hinge dynamics are much slower than the pulse width, which is clearly valid at normal communication system bit rates, the time correlation functions for the PMD vector and Stokes vector are

$$f = \frac{1}{T} \int_0^T \vec{\tau}(t) \cdot \vec{\tau}(t + \alpha) dt \quad (3.3)$$

$$g = \frac{1}{T} \int_0^T \hat{s}_{out}(t) \cdot \hat{s}_{out}(t + \alpha) dt \quad (3.4)$$

where T is the observation window and α is the time shift. We now compute ensemble averages for f and g . These averages can be obtained in practice by measuring f and g at a multitude of statistically independent frequencies. This is explained in more detail at the end of this section. The ensemble average of f over the different fiber realizations of sections 1 and 2 is obtained by substituting Eq. (3.1) in Eq. (3.3) to obtain

$$\langle f \rangle = \left\langle \frac{1}{T} \int_0^T (\bar{\tau}_2' + \bar{\tau}_1' H'(t) R_2') (\bar{\tau}_2 + R_2 H(t + \alpha) \tau_1) dt \right\rangle \quad (3.5)$$

in which the DGD of the two sections is written as a 3x1 matrix, while the prime here denotes transpose. Exchanging the time and ensemble averages yields, noting that $R_2' R_2 = I$, where I is the identity matrix,

$$\begin{aligned} \langle f \rangle = & \langle \bar{\tau}_2'^2 \rangle + \left\langle \bar{\tau}_1' \left[\frac{1}{T} \int_0^T dt H'(t) \right] R_2' \bar{\tau}_2 \right\rangle + \left\langle \bar{\tau}_2' R_2 \left[\frac{1}{T} \int_0^T dt H(t + \alpha) \right] \tau_1 \right\rangle \\ & + \left\langle \bar{\tau}_1' \left[\frac{1}{T} \int_0^T dt H'(t) H(t + \alpha) \right] \tau_1 \right\rangle \end{aligned} \quad (3.6)$$

The second and third terms yield zero since $\bar{\tau}_1$ and $\bar{\tau}_2$ are independent and the ensemble average of the product of any two of their Cartesian components should be zero. Denoting

$$M = \frac{1}{T} \int_0^T dt H'(t) H(t + \alpha) \quad (3.7)$$

the fourth term can be rewritten as

$$\langle \bar{\tau}_1' M \tau_1 \rangle = \left\langle \begin{pmatrix} \tau_{1x} & \tau_{1y} & \tau_{1z} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} \tau_{1x} \\ \tau_{1y} \\ \tau_{1z} \end{pmatrix} \right\rangle \quad (3.8)$$

With $\langle \tau_{ii} \rangle = 0$, $i = x, y, z$, $\langle \tau_{ii}^2 \rangle = \frac{1}{3} \langle \tau^2 \rangle$ [15] and observing that the elements of the matrix M are independent of the PMD components of section 1, we obtain

$$\langle \tau_1 M \tau_1 \rangle = \frac{1}{3} \langle \tau_1^2 \rangle \langle Tr(M) \rangle \quad (3.9)$$

where $Tr(M)$ is the trace of the matrix M .

Hence Eq. (3.6) reduces to

$$\langle f \rangle = \langle \tau_2^2 \rangle + \frac{1}{3} \langle \tau_1^2 \rangle \langle Tr(M) \rangle \quad (3.10)$$

Before calculating the ensemble average of g , we note that regardless of \hat{s}_m , the transformation $\hat{s} = R_1 \hat{s}_m$ yields a vector that has the properties $\langle s_x \rangle = \langle s_y \rangle = \langle s_z \rangle = 0$; $\langle s_x^2 \rangle = \langle s_y^2 \rangle = \langle s_z^2 \rangle = 1/3$ as long as section 1 is in the long length regime [32]. Employing Eqs. (3.2) and (3.4) and repeating the above procedure, we further obtain for the ensemble average of g

$$\langle g \rangle = \frac{1}{3} \langle Tr(M) \rangle \quad (3.11)$$

Combining Eqs. (3.10) and (3.11) yields

$$\langle f \rangle = \langle \tau_2^2 \rangle + \langle \tau_1^2 \rangle \langle g \rangle \quad (3.12)$$

Since $\langle \tau^2 \rangle = \langle \tau_2^2 \rangle + \langle \tau_1^2 \rangle$, Eq. (3.12) can be rewritten as

$$\langle f \rangle = \langle \tau_2^2 \rangle (1 - \langle g \rangle) + \langle \tau^2 \rangle \langle g \rangle \quad (3.13)$$

Therefore, if the PMD vector and the Stokes parameters are measured for different fiber realizations, $\langle f \rangle$ and $\langle g \rangle$ evaluated, $\langle \tau_2^2 \rangle$ can be obtained by a least squares fit yielding the position of the hinge if the average DGD of the link is known. The least squares fit is performed such that $\langle \tau_2^2 \rangle$ minimizes

$$\text{the error defined by } \int_{-T}^T d\alpha \left(\langle f \rangle - \langle \tau_2^2 \rangle (1 - \langle g \rangle) - \langle \tau^2 \rangle \langle g \rangle \right)^2 .$$

The preceding formulas involve ensemble averages of time correlation functions that can be evaluated in two possible ways. One of these is to observe the system over a time interval that encompasses many average DGD correlation times and then to average over the values obtained over each correlation time. The second is to observe a system over a band of optical frequencies that is broader than a sufficient number of correlation frequencies, such as present in a WDM system and to average the relevant quantities over their uncorrelated values at representative frequencies that differ by at least the correlation frequency. Here the spectral separation of frequencies in the average must accordingly be at least $750 \text{ GHz} / \langle \tau \rangle$, where $\langle \tau \rangle$ is the mean DGD of the link in ps [15] which corresponds to 6 times the correlation frequency of the PMD vector. Further, since the frequency correlation function of the Stokes vector is narrower than that of the PMD vector [51], [52] the Stokes vectors at the employed frequencies will then necessarily also be statistically independent. Thus in a dense WDM (DWDM) system with 80 channels with 50GHz separation [53], we require $\langle \tau \rangle > 15 \text{ ps}$. In a WDM system, each channel's PMD vector and Stokes vector will be monitored. When the rapid changes are detected, the correlation functions are calculated. The value of $\langle \tau^2 \rangle$ is obtained as the value of $\langle f \rangle$ at $\alpha = 0$ and then the value of $\langle \tau_2^2 \rangle$ can be calculated from the least squares fit.

3.3 Simulation Parameters

In view of the above discussion, we simulate a link by generating 80 random realizations from which we employ the least squares fit described in the previous section to obtain the ratio $X = \langle \tau_2^2 \rangle / \langle \tau^2 \rangle$ which is the fraction of the squared DGD following the hinge. This ratio is then compared to the exact value X_{exact} . We first employ 80 realizations to establish the feasibility of the method in the case that only the channels present in a standard DWDM system are available. The fiber link, which is assumed to possess a mean DGD of $\tau = 20 \text{ ps}$ at a vacuum wavelength of $1.55 \mu\text{m}$, is modeled as a

concatenation of $N = 100$ birefringent sections with a DGD of $\tau_n = \sqrt{\frac{3\pi}{8N}} \langle \tau \rangle + \nu$ in each section where ν is a normally distributed random variable with zero mean and a standard deviation $\sigma_\nu = 0.2 \sqrt{\frac{3\pi}{8N}} \tau$. The Gaussian random variable, ν , randomizes the DGD of each section slightly

and prevents the spectrum of the PMD from becoming periodic with frequency [15]. Each section is chosen to have a random rotation axis. At the same time, the magnitude of the PMD vector of each section is held constant while its direction is varied randomly for each realization. We employed two different models for the hinge. Measurements on DCM's have indicated that both the rotation angle and the rotation axis of the hinge may vary with time. Variation of hinge properties with frequency has been also reported but for large frequency shifts on the order of THz [46]. Therefore, the hinge, which is placed a certain number of sections from the entrance face of the fiber, is then modeled in model 1 by a polarization rotator with zero DGD, a time varying rotation axis given by $\begin{bmatrix} r \cos(2\pi ft) & r \sin(2\pi ft) & r_z \end{bmatrix}$ where $r^2 + r_z^2 = 1$ and a rotation angle given by the decaying sinusoidal function $\varphi_d = 1.5\pi \sin(\omega t + 2\pi z) \exp(-40t)$ where $\omega = 2\pi \times 100$ rad/s and z is a uniformly distributed random variable between 0 and 1. In each realization of the link, the phase $2\pi z$ will take on a different value and this serves to model the variation of the hinge behavior with frequency in a real system. The decaying rotation angle models the rapid but strongly damped excitations that are short lived according to our assumptions. Of course, hinge dynamics are to a certain extent deterministic as, for example, a model based on a prescribed evolution of a few hinges in a link yields results very similar to certain experiments [45], [46]. The present theory, however, does not assume any particular properties of the hinge behavior as a function of time or frequency and therefore our results are valid within wide limits irrespective of this behavior. In model 2, we employ the same form for the rotation angle but in each realization the value of r_z is chosen randomly from the uniform distribution defined over the interval $[-1,1]$ and then r is obtained from $r = \sqrt{1 - r_z^2}$. Model 1 assumes that only the rotation angle is a function of frequency while model 2 assumes that both the rotation axis and the rotation angle of the hinge vary with frequency. The correlation functions f and g are computed over a time after which the rotation angle is effectively zero and are then averaged over all realizations to obtain $\langle f \rangle$ and $\langle g \rangle$. Changing this time interval did not change any of the results significantly. In all simulations, a fixed input polarization $\hat{s}_{in} = [1 \ 0 \ 0]'$ is employed. With the value of $\langle \tau^2 \rangle$ taken as that of $\langle f \rangle$ at $\alpha = 0$, a least squares fit is then employed to determine $\langle \tau_2^2 \rangle$ from Eq. (3.13). The ratio X is then calculated and compared to X_{exact} . Knowledge of the mean DGD of the link together with the value of X then yields an estimate of the DGD preceding the hinge.

3.4 Results and Experimental Setup

In Table 3-1, we display the value of X obtained from our simulations compared to X_{exact} for 80 realizations and two different sequences of random numbers for model 1. The same data is displayed in Table 3-2 for model 2. Except for the data points at $X_{exact} = 0.2, 0.3$ for model 1, sequence 2, the error in all simulation points is less than 15%. This error and error difference between the two sequences demonstrates the statistical limitation of the method as a result of the limited number of

data samples. However, if the rapid PMD changes recur, the statistical error would be decreased. To verify this, we ran our simulations for 1000 realizations instead of 80 for the two different sequences of random numbers. The results displayed in Table 3-3 and Table 3-4 for models 1 and 2 respectively indicate much improved results for the values of X , model 1, sequence 2 hence supporting the theory presented in section 3.2. In addition to that, in Figure 3.2, we show the difference between the values of X obtained from the two sequences for models 1 and 2 for 80 realizations and 1000 realizations. As can be observed, the difference is reduced when more realizations are employed irrespective of the model chosen for the hinge.

Table 3-1. Exact and computed values of X for 80 realizations and two different sequences of random numbers for model 1

Exact	Simulated sequence 1	Simulated sequence 2
0.2	0.2069	0.2786
0.3	0.3575	0.4369
0.4	0.3889	0.4568
0.5	0.5097	0.5384
0.6	0.5937	0.6269
0.7	0.7263	0.737
0.8	0.8150	0.7962

Table 3-2. Exact and computed values of X for 80 realizations and two different sequences of random numbers for model 2

Exact	Simulated sequence 1	Simulated sequence 2
0.2	0.2044	0.2025
0.3	0.27	0.31
0.4	0.37	0.44
0.5	0.45	0.56
0.6	0.5977	0.60
0.7	0.6832	0.72
0.8	0.8044	0.81

Table 3-3. Exact and computed values of X for 1000 realizations and two different sequences of random numbers for model 1

Exact	Simulated sequence 1	Simulated sequence 2
0.2	0.2053	0.2171
0.3	0.3309	0.337
0.4	0.4265	0.4298
0.5	0.5117	0.5194
0.6	0.5914	0.600
0.7	0.7055	0.7052
0.8	0.7926	0.7960

Table 3-4. Exact and computed values of X for 1000 realizations and two different sequences of random numbers for model 2

Exact	Simulated sequence 1	Simulated sequence 2
0.2	0.2115	0.2084
0.3	0.3294	0.3414
0.4	0.4181	0.4216
0.5	0.4961	0.5011
0.6	0.5964	0.6038
0.7	0.7097	0.7075
0.8	0.7926	0.8018

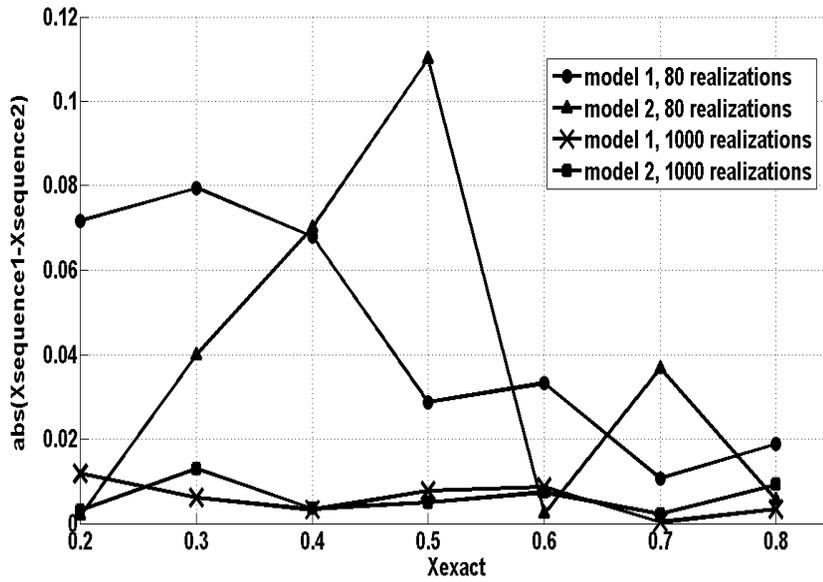


Figure 3.2. Absolute difference between X from sequences 1 and 2 for the two hinge models used and for 80 and 1000 realizations.

The above theory can be implemented through the experimental setup of Figure 3.3. Here a tunable laser is connected a concatenation of two fibers separated by a polarization controller that models a rapidly varying hinge in an actual fiber system. The mean DGD of both fibers 1 and 2 is assumed known or measured beforehand so that the exact value of X can be determined. At the output of fiber 2, a splitter splits the light into PMD and Stokes parameters measuring apparatus. By collecting data for a sufficiently large number of correlation frequencies $\langle f \rangle$ and $\langle g \rangle$ and hence X can be determined and compared to X_{exact} .

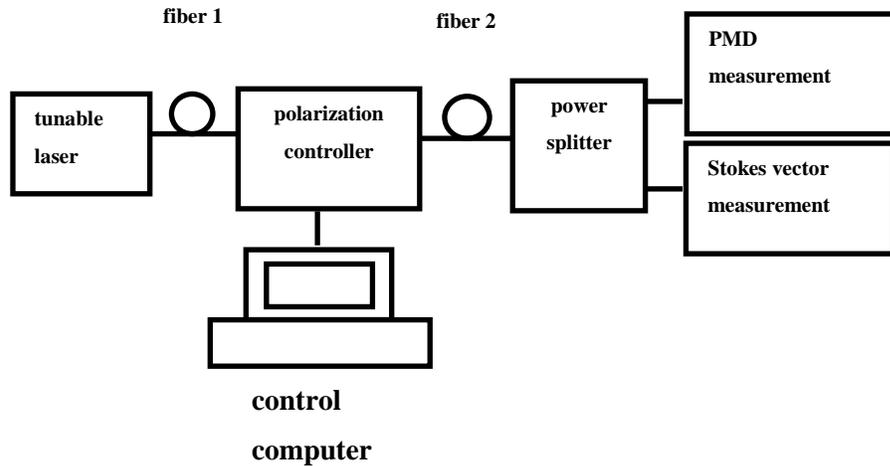


Figure 3.3. Proposed experiment. A tunable laser launches light into a concatenation of two fibers connected through a polarization controller acting as a hinge. A power splitter at the output of fiber 2 splits light into PMD and Stokes vector measurement apparatus for simultaneous measurements. The frequency of the laser is then varied and the process is repeated and ensemble averages are formed. The computer is used to control the time variation of the polarization controller thus implementing the hinge in the link.

3.5 Conclusion

We have demonstrated that the location of a hinge that induces fast polarization changes in WDM communication links together with the amount of DGD before and after the hinge can be determined to within statistical error by averaging over suitable correlation functions. The procedure is not limited to WDM standards. It should be noted here that if sufficient data is collected, incidents in which two or more hinges are active could be separated from the remainder of the events. Statistical error can be reduced either by including more frequencies in the ensemble average, which can be accomplished by including additional diagnostic wavelengths in the system that do not carry data but can be employed in measurements. In a WDM system, such additional wavelength channels could further be periodically modulated at the position of different hinges in order to provide a continuous map of the DGD of the link along its length.

Chapter 4

Differential Group Delay (DGD) Prediction in Optical Fiber Links

4.1 Introduction

Given that the DGD of transmission fiber varies stochastically with time, knowledge of the outage probability for times beyond the present state of a communication system can enable mitigation based on system reconfiguration [54]. For short buried cable routes, characteristic times for the change of DGD have been observed to be of the order of weeks or months. These times are expected to diminish as the reciprocal of the square root of route length [52].

In this chapter, we simulate the time evolution of DGD in a fiber link in the hinge model with time varying hinges. The hinges are assumed here to be deterministic. However, measurements indicate that sometimes hinges show stochastic behavior [55]. We apply time series forecasting procedures, in particular, autoregressive (AR) models with Kalman filter, pattern imitation and a Taylor expansion to the DGD and compare their prediction abilities. Results of this study shed light on the degree to which it is reasonable to reconfigure based on a forecast of possible outage conditions. In these studies, the furthest prediction horizon was attained with the AR accompanied by the Kalman filter procedure. We further discuss the influence of measurement noise on the prediction horizon. We also apply these methods to the measured DGD on 40 Gb/s fiber links and discuss the results.

4.2 Model Parameters

We model a fiber link with a mean DGD of $\langle \tau \rangle = 20$ ps at a vacuum wavelength of 1.55 μm . Five static, highly mode coupled fiber spans each consisting of 50 birefringent segments are connected by four polarization rotating hinges as shown in Figure 4.1. The total number of birefringent segments is

$N = 250$ and the DGD of each section, τ_n , is set to $\tau_n = \sqrt{\frac{3\pi}{8N}} \langle \tau \rangle + v$ in which v is a normally

distributed random variable with zero mean and a standard deviation $\sigma_v = 0.2 \sqrt{\frac{3\pi}{8N}} \langle \tau \rangle$ as described

in Chapter 3. As in [45], [46] hinges are modeled as randomly oriented polarization rotators with frequency-independent axes and a rotation angle given by $\varphi_k = 1.5\pi \sin(2\pi f_k t + (k-1)\pi/8) + 2\pi t/500$, where t is the time, k is the hinge number and the hinge frequency f_k is 0.5 for odd numbered hinges and 0.8 for even numbered hinges. This model has been shown elsewhere to yield results that resemble qualitatively certain experimental observations [45], [46] and in fact our results were very slightly altered for frequencies in the range from 0.5 to 1. The periods $1/f_k$ of the hinge rotation angles are approximately 2 and 1.25 time units. Below, a total time $t = 5$ will be employed to generate the AR models while a time of $t = 50$ is employed to populate the database in the pattern imitation technique. These correspond to 2.5 times

and 25 times the slowest simulated hinge period respectively. The DGD of the link is then determined by applying the concatenation rule of the PMD (polarization mode dispersion) vector at each time instant.

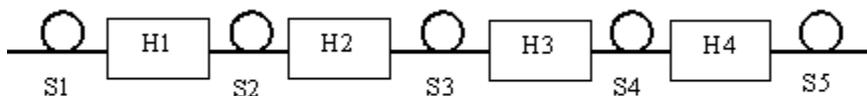


Figure 4.1. Simulated optical link. Five fiber spans S1 to S5 are connected via a series of time varying hinges H1 to H4 modeled as polarization rotators. Each fiber span consists of 50 birefringent sections each having randomly oriented rotation axes.

The simulation yields the three components τ_x, τ_y and τ_z of the PMD vector and τ is then calculated $\tau = \sqrt{\tau_x^2 + \tau_y^2 + \tau_z^2}$ as a function of time. We also simulate measurement noise by adding a noise vector $\vec{\chi}$ to $\vec{\tau}$. The components of $\vec{\chi}$ are independently Gaussian distributed random variables with zero mean and a standard deviation of 1ps. If the signal is associated with τ , the resulting signal to noise ratio S/N is 25 dB over a period of 10 time units. Before applying the methods to the noisy signals, a filter is used to smooth the data. The details of the filter parameters are described in section 4.3.4. We apply the prediction methods to τ , τ_x, τ_y, τ_z , and to τ^2 . In cases when the methods are not applied to τ , τ is derived from the predicted quantities for e.g. from τ^2 by taking the square root. For AR models and the Taylor expansion technique, in the absence of noise, applying the models below to τ_x, τ_y and τ_z individually or to τ^2 and then deriving τ from the resulting signals often yields an improvement over instead modeling τ , except the pattern imitation method. For measured data, only the DGD is available so that our models can only be applied to τ .

To quantify the reliability of various prediction methods for the resulting numerical or measured DGD time series we define the metric

$$\Phi = 100 \times \left(1 - \frac{\|y_v(n) - \hat{y}(n)\|}{\|y_v(n) - \bar{y}\|} \right) \quad (4.1)$$

Here n is an integer representing the time index, and the signals are implemented as a discrete process where \hat{y} is the predicted signal, y_v ranges over the validation section of the signal, and \bar{y} is the mean over the entire signal (which in our calculations is nearly independent of the number of data

points for sufficiently large record lengths). The metric thus represents the percentage improvement of the prediction of a given method relative to employing the mean value of the signal as the predicted value. For a signal $\xi(n)$ with $n=1,2,3,\dots,N_L$ the norm, $\|\xi\|$, is defined as

$$\|\xi\| = \left(\sum_{i=1}^{N_L} \xi^2(i) / N_L \right)^{1/2} .$$

At a certain time index N and for a given model, $\hat{y}(N)$ is generally calculated from the values of y_v up to a time index of $N-P$ where P is termed the prediction horizon. To establish the prediction capabilities of each model, the metric Φ is calculated for $N_L = 7000$ corresponding to 4.86 time units and different prediction horizons. The results are expressed in terms of the signal correlation time N_c defined here as the time index halfway between the correlation peak and the minimum of the first oscillation of the correlation function of the DGD as illustrated for our simulated time series in Figure 4.2. Here the time correlation function is defined as

$$R(\alpha) = \frac{1}{T} \int_0^T \tau(t)\tau(t+\alpha)dt \text{ with } T=10 \text{ and } \alpha \text{ the time shift. Increasing } T \text{ does not affect } N_c$$

which equals 97 for our simulated data. The models are also compared by calculating the ratio $\psi = P_{80\%} / N_c$ where $P_{80\%}$ is the prediction horizon for $\Phi = 80\%$. This provides an estimate of the time for which the model can predict future values subject to a given accuracy compared to the decorrelation time of the the DGD signal.

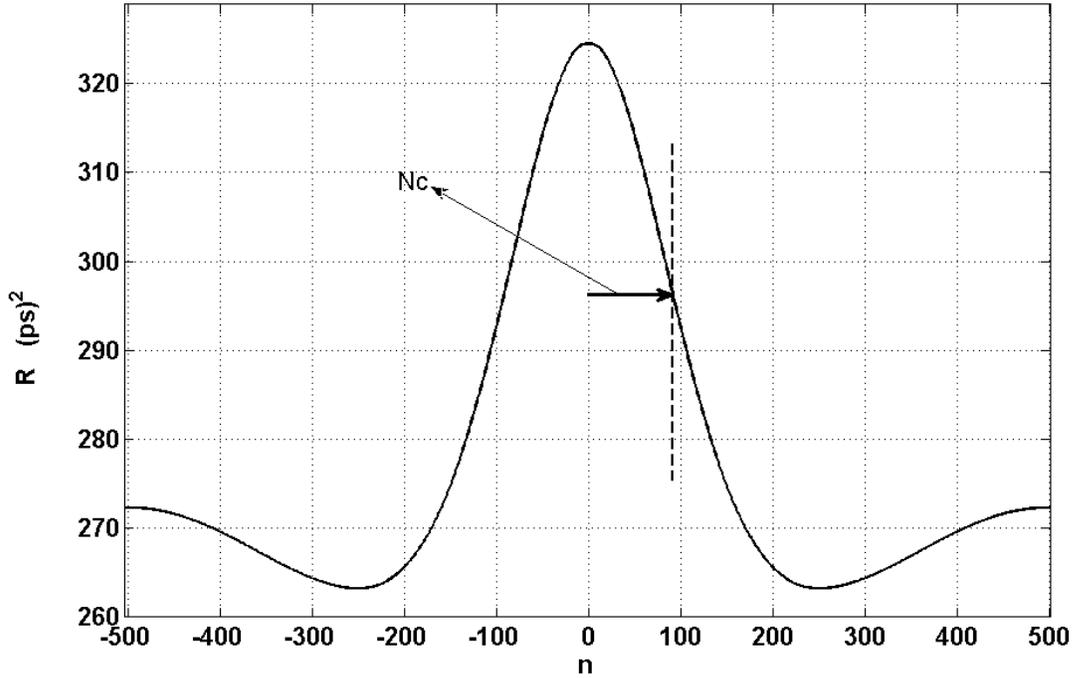


Figure 4.2. Correlation function of simulated DGD time series. The correlation index is defined as the distance between the maximum of the correlation function and the first minimum.

4.3 Prediction Methods

The three prediction algorithms employed in this paper, namely, the AR procedure with a Kalman filter, a pattern imitation technique and a Taylor expansion are now described as well as the filter we employ to reduce the noise.

4.3.1 AR Modeling and Kalman Filters

The first step in implementing an AR model, is to subtract the mean of the data from each data value. The resulting sequence is then divided into estimation and validation sections as shown in Figure 4.3. The estimation section is employed to generate the model that is subsequently analyzed with the validation data. We denote the estimation points by $y(n)$ and the validation points by $y_v(n)$ as detailed in section II. Given k previous values of y , namely, $y(n-1), y(n-2), \dots, y(n-k)$ an estimate of $y(n)$, denoted by $r(n)$ is obtained from the weighted average of the previous k values of y according to

$$r(n) = \sum_{i=1}^k a_i y(n-i) \quad (4.2)$$

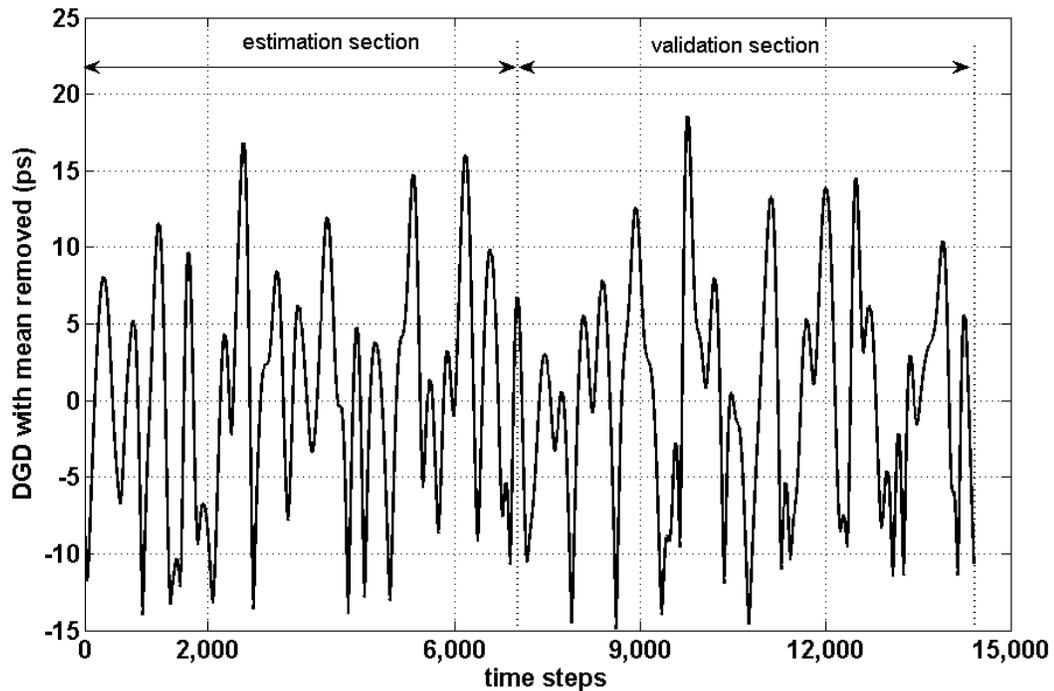


Figure 4.3. DGD time series with mean removed. The first 7200 data points are used to estimate the model and the following 7000 points of the series is employed for validation.

As Eq. (4.2) represents a k^{th} order AR model for $y(n)$ it will be referred to as AR(k) in our following discussion . The coefficients a_i are obtained from a least squares fit of Eq.(4.2) to the true estimation points $y(n)$ for $n = 1, 2, 3, \dots$ up to the end of the estimation record with Burg's method [56]. The resulting model for the estimation section is

$$y(n) = \sum_{i=1}^k a_i y(n-i) + w(n) \quad (4.3)$$

where $w(n)$ represents the error of the fitting procedure. We apply this technique to a record given by a simple sine wave described by $\sin(0.025\pi n)$ with $n = 1, 2, \dots, 75$ which is less than a full period. Thus for example, the least squares fit yields a difference equation $y(n) = a_1 y(n-1) - a_2 y(n-2)$ with $a_1 = 1.994$ and $a_2 = 1.0$ in place of the exact values 2.0, 1.0 for a $k=2$ second order model . To describe the Kalman filter technique, the model of Eq. (4.2) is recast in state space form. For this purpose we define a state vector $\bar{x}(n) = [y(n) \ y(n-1) \ \dots \ y(n-k+1)]^T$. From Eq. (4.3), the state vector at time n is given in terms of the state vector at time $n-1$ through

$$\bar{x}(n) = A \bar{x}(n-1) + \bar{w}(n) \quad (4.4)$$

where A is a $k \times k$ matrix composed of the coefficients a_i

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

and the error vector $\bar{w}(n) = [w(n) \ 0 \ 0 \ 0 \ \dots]^T$.

The current value of y can be obtained from the state vector through

$$y(n) = C \bar{x}(n) \quad (4.5)$$

where C is given by $C = [1 \ 0 \ \dots \ 0]$.

Eqs. (4.4) and (4.5) describe the model developed for the signal from the estimation data points. To predict the value of the signal, $\hat{y}(N)$, for a time N in the validation section, all the $y_v(n)$ with $n \leq N - P$ are employed. In particular, an initial state vector is chosen and the following formula applied [56]

$$\hat{x}(n) = A\hat{x}(n-1) + \vec{K}(n) \left[y_v(n) - CA\hat{x}(n-1) \right] \quad (4.6)$$

where the Kalman filter gain $\vec{K}(n)$ is a $k \times 1$ column vector described in [56]. In this article, $\hat{x}(n)$ is the estimated value of the state vector at time n given the previous value of the state vector $\hat{x}(n-1)$ and the current value of y_v . The Kalman filter gain is successively updated for the next time index [56] and the state vector recalculated with Eq. (4.6) as outlined in [56] up to time index $N - P$. Eq. (4.4) is then applied P times with the error vector absent and $\hat{y}(N)$ is finally obtained from the final state vector according to Eq. (4.5). Note that the actual time series is not stored as only the updated state vector is required for each new value of the validation signal. The results of our simulations are effectively independent of the choice of initial state vector. We have also applied the above method in conjunction with an ARMA (autoregressive moving average) [56] or a state space model [56] in place of the AR model without significantly improving the predicted data values. It is also found that $\vec{K}(n)$ saturates after the first few time updates in a manner such that the predicted value can be just obtained by taking the most recently recorded state vector and applying Eq. (4.5) P times. In other words the AR model described by Eq. (4.3) is sufficient to generate the prediction if applied P times employing $y_v(n-P), y_v(n-P-1), \dots, y_v(n-P-k+1)$.

To select the model order, models of varying orders are first generated from the estimation section that are then verified on the validation section. Since generally increasing the model order beyond a certain value does not enhance the prediction accuracy or horizon, we increase the model order until the value of the metric becomes independent of the order number.

4.3.2 Pattern Imitation [57]

Pattern imitation techniques are based on the assumption of quasiperodicity in a time series that is divided into separate records by matching previous records to the most recent record and then employing the extension of the previous record to predict future values. Thus the time series is recorded over a sufficiently long period of time and stored in a database as a series of state vectors $\bar{x}(n) = [y(n) \ y(n-1) \ \dots \ y(n-(d-1))]$. Each successive set of d measurements in the

validation section yields a new state vector, $\vec{x}_{measured}(n-P) = [y_v(n-P) \ y_v(n-P-1) \ \dots \ y_v(n-P-(d-1))]$. The database is searched for the nearest neighbor state vector, $\vec{x}_{nn}(l)$, that minimizes the norm of the difference vector $\vec{x}_{measured} - \vec{x}$ after which the predicted value of the signal is given by $\hat{y}(N) = y(l+P)$. Alternatively, several closely matching state vectors can be identified and the average of the corresponding values of $y(l+P)$ of these vectors employed as the prediction.

To implement this technique, we increase the size of the database until no further improvement is recorded in the metric, Eq. (4.1). The number of nearest neighbors and the size of the state vector d are optimized iteratively by fixing one variable and varying the other variable to maximize the metric and repeating this for each variable in turn until convergence is achieved. From the numerical data we have determined that a database of 72,000 points is sufficient for our problem with while $d = 10$ and 10 nearest neighbors supply optimal results.

4.3.3 Taylor Expansion

A straightforward procedure evaluates the derivatives of the time series up to a certain order. The predicted signal $\hat{y}(N)$ after a time step $\Delta t = Pt_s$ for a sampling time t_s , is then given by the Taylor expression

$$\hat{y}(N) = y_v(N-P) + \Delta t y_v'(N-P) + \frac{(\Delta t)^2}{2} y_v''(N-P) + \frac{(\Delta t)^3}{3!} y_v'''(N-P) + \dots \quad (4.7)$$

As in AR modeling, the order of the derivative at which the series is truncated is chosen such that increasing the order does not noticeably affect the metric, Eq. (4.1).

4.3.4 Filter Description

Before applying prediction methods to the noisy signal, the time series is smoothed with a filter such that numerical differentiation can be applied in our final method. We employ local regression in which successive sets of N_{fit} data points are fitted to a quadratic function with a least squares technique along the lines of the Savitzky Golay filter [58]. The derivative of the noisy signal at a point is then estimated from the derivative of the resulting quadratic function at the same point. Subsequently N_{fit} is optimized by calculating a predicted signal and comparing it to τ which is noise free. However, to avoid significantly distorting the signal as a result of the smoothing operation, we have employed $N_{fit} = 250$ points for the numerical data as our results are then not appreciably changed by increasing N_{fit} by 100 points.

4.4 Results

We first investigate the three methods outlined above in the absence of noise in Figure 4.4 for which the quantity Φ defined in Eq. (4.1), is plotted against the prediction horizon P . Where this metric is negative, indicating that the mean value is a better predictor than the model, it is replaced by 0%. The lines in the figures are the results of (1) an AR(5) model applied to τ (dotted line), (2) an AR(8) model applied to τ^2 (dashed line), (3) a combination of three separate AR(9), AR(9) and AR(8) models applied to τ_x, τ_y and τ_z respectively (circle markers), (4) a seventh-order Taylor expansion for τ_x, τ_y and τ_z (square markers), and (5) the pattern imitation procedure applied to τ (triangle markers). The results of Figure 4.4 indicate that models (2) and (3) yield the longest prediction horizon of the models surveyed subject to a given level of accuracy. Indeed, the individual PMD components and τ^2 do not experience the nonlinearity introduced by taking the square root of τ^2 , improving the performance of the linear AR model and the Taylor expansion technique. The pattern imitation method, however, displays equal accuracy when applied either to the individual components or to τ^2 . As indicated in Figure 4.5, which employs model (3) for $P = 130$, predicted values with $\Phi \geq 80\%$, can be obtained even for horizons larger than the correlation time. For this model, $\psi = 1.44$.

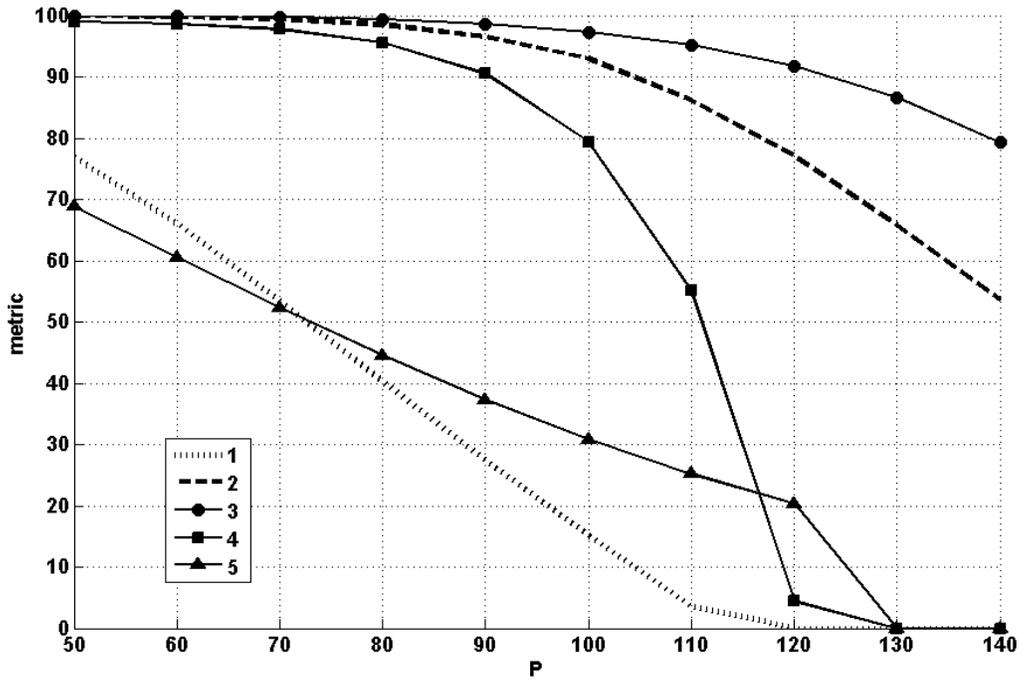


Figure 4.4. The metric versus prediction horizon for the different methods employed. The legend indicates model numbers as described in section 4.4. (1) an AR(5) model applied to τ (dotted line), (2) an AR(8) model applied to τ^2 (dashed line), (3) a combination of three separate AR(9), AR(9) and AR(8) models applied to τ_x, τ_y and τ_z respectively (circle markers), (4) a seventh-order Taylor expansion for τ_x, τ_y and τ_z (square markers), and (5) the pattern imitation procedure applied to τ (triangle markers).

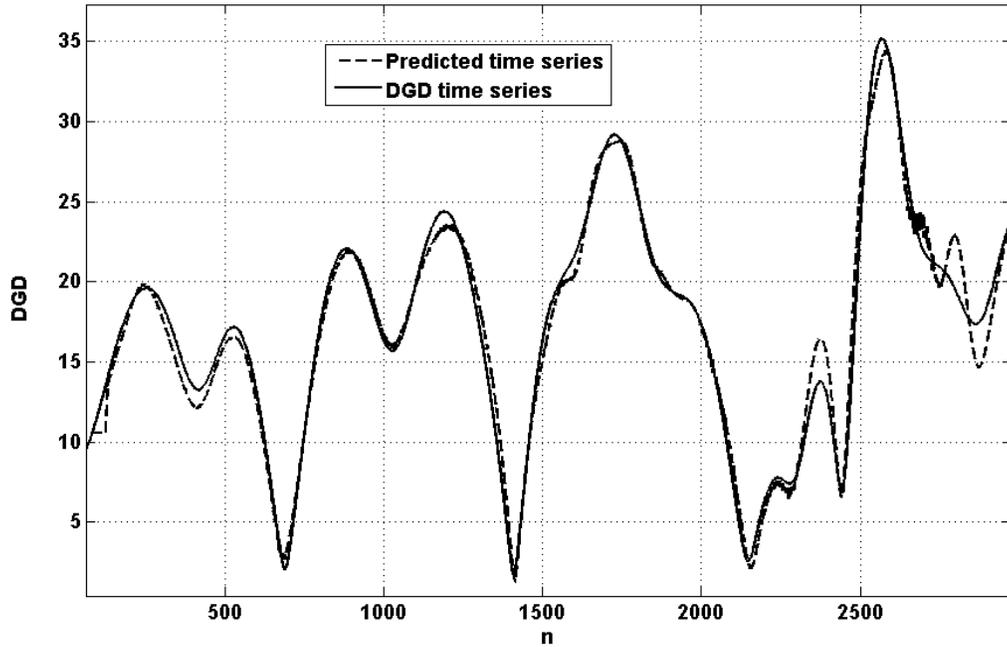


Figure 4.5. Simulated and predicted DGD time series for model (3) for $P = 130$ and $\Phi = 86\%$.

In Figure 4.6, we display the same results after noise addition and filtration. Graph (1) is an AR(10) model of τ (dashed line), (2) for AR(9), AR(9), AR(8) models applied to the PMD vector components just as in the noiseless case, (3) Taylor expansion up to second order and (4) the pattern imitation technique. The figure indicates that the prediction horizon for the same Φ is reduced for filtered noisy data. $\psi = 0.77$ for model (1) which is the best among the models considered . It is also evident from the figure that when the model order in the DGD implementation is sufficiently large, the resulting Φ can be better in the presence than in the absence of noise. In fact, filtering removes not only high frequency noise but also the high-frequency harmonics generated when a square root operator is applied to τ^2 . On the other hand, the pattern imitation technique is unaffected by filtering, while higher order Taylor expansion models instead diverge rapidly even for small prediction horizons. Indeed, a derivative introduces high pass filtering of the low pass prefiltered signal so that the higher order derivatives are affected by the prefiltering. In Figure 4.7, we accordingly compare the filtered DGD time series (solid curve) and the predicted time series (dashed curve) for a prediction horizon of 75 points and $\Phi = 81\%$.

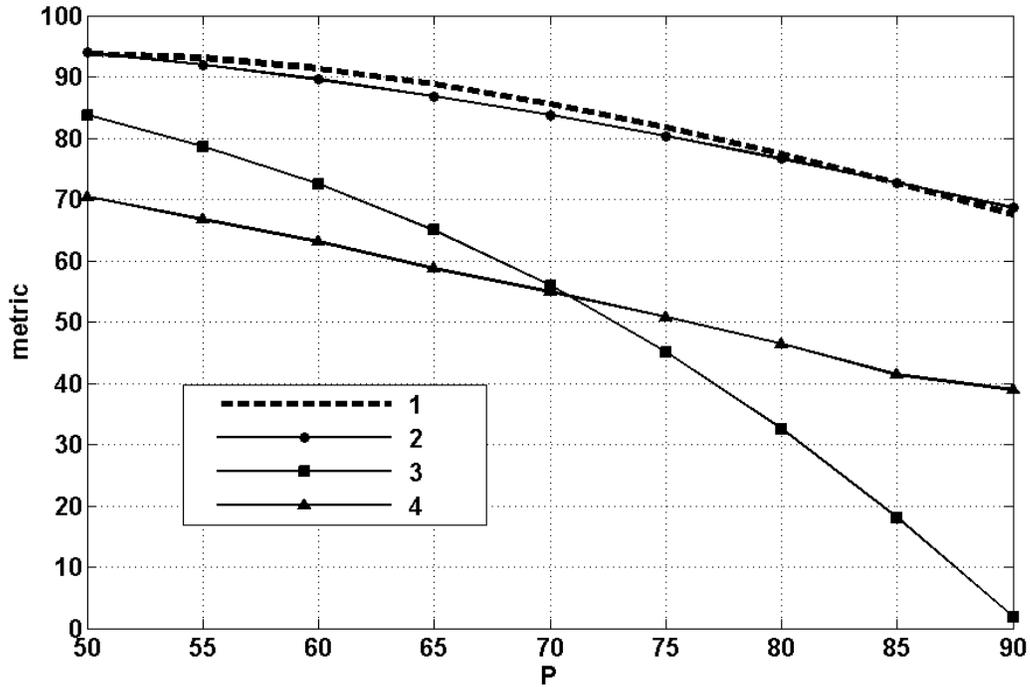


Figure 4.6. Metric plotted as a function of the prediction horizon after filtering.

We further applied our procedures to the measured DGD for two different fibers employed in 40 Gb/s systems. The data were obtained from field measurements spanning 46 and 89 correlation times for fibers 1 and 2 respectively. The DGD was sampled every 60s and was measured over a period of 11 days for fiber 1 and 21 days for fiber 2. A total of 16984 samples were collected for fiber 1 and 28359 for fiber 2. The smoothing filter described in section 4.3.4 with $N_{fit} = 300$ was employed to reduce noise. Since the speed of variation of the DGD increases with both the number of hinges in the link and the rate at which each hinge varies, our lack of information on the distribution and time variation of the hinges in the link precludes establishing a relationship between the filter integration time and the hinge dynamics. However, we observe that the signal to noise ratio is 20dB for fiber 1 and 25dB for fiber 2. Fitting an AR(8) model to the resulting DGD time series in which half of the points are employed for estimation and the other half for validation, we found ψ equal 0.25 and 0.38 for fibers 1 and 2 respectively to an accuracy of $\pm 2.5\%$ as determined by varying the validation record length. For fiber 1, Figure 4.8 demonstrates the predictive accuracy of the model, which yields a horizon of 90 points with $\Phi = 81\%$. The 90 point horizon successfully predicts future DGD changes of up to 50% as in the region between $n = 2017$ to $n = 2107$ where the DGD changes from 19ps to 30ps. In Figure 4.9 we display the corresponding prediction and measured time series for fiber 2. Here the horizon is 120 points while $\Phi = 81\%$.

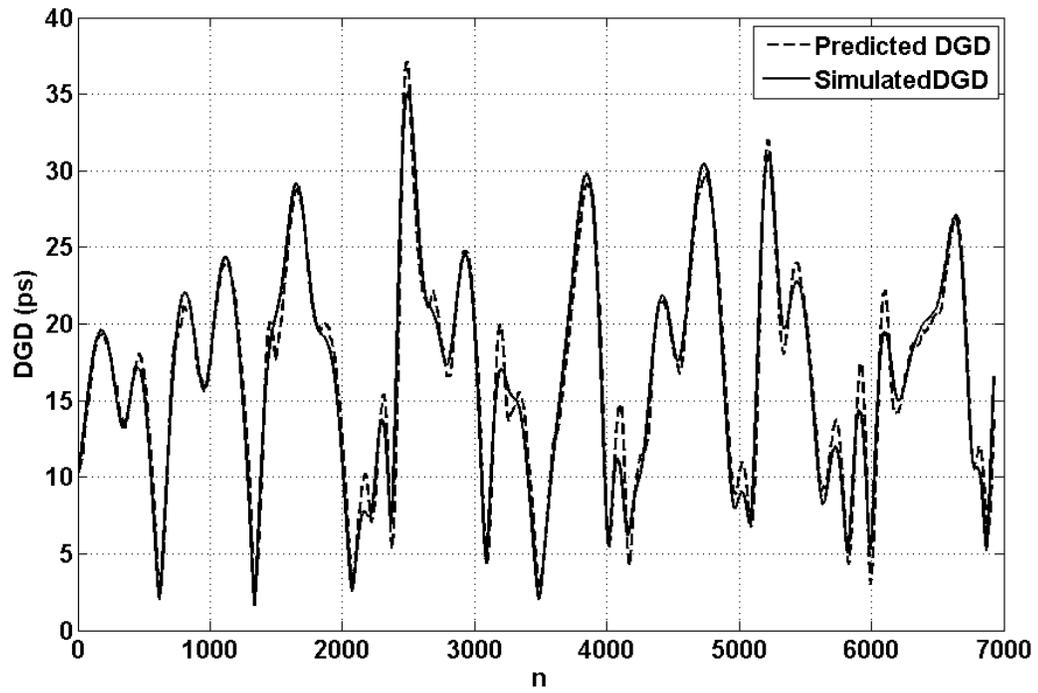


Figure 4.7. Simulated DGD and predicted DGD time series for noisy data for AR(10) model applied to the DGD.

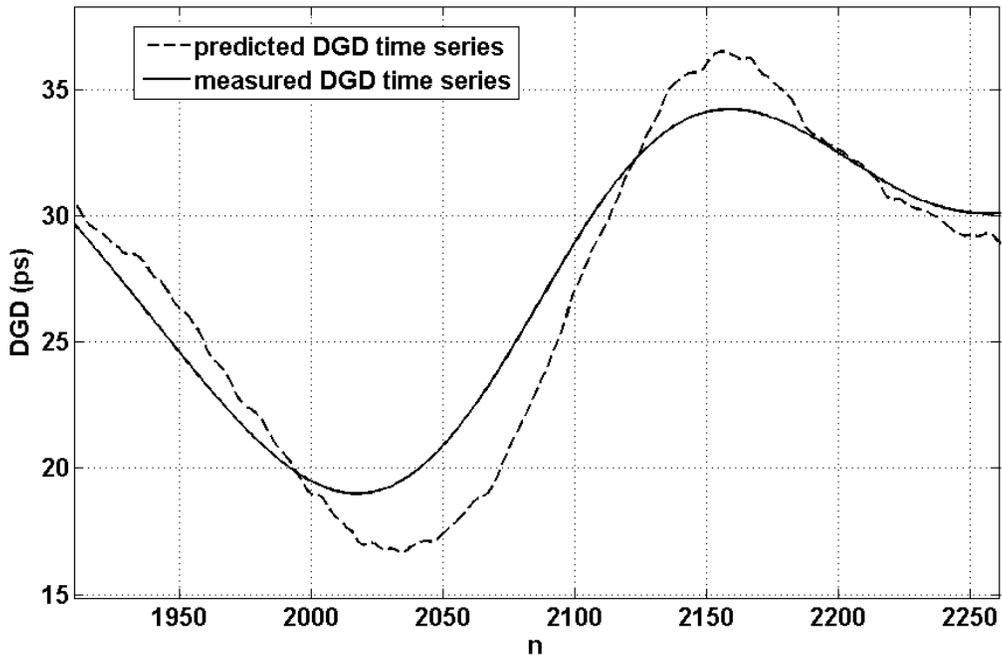


Figure 4.8. Predicted and measured DGD.

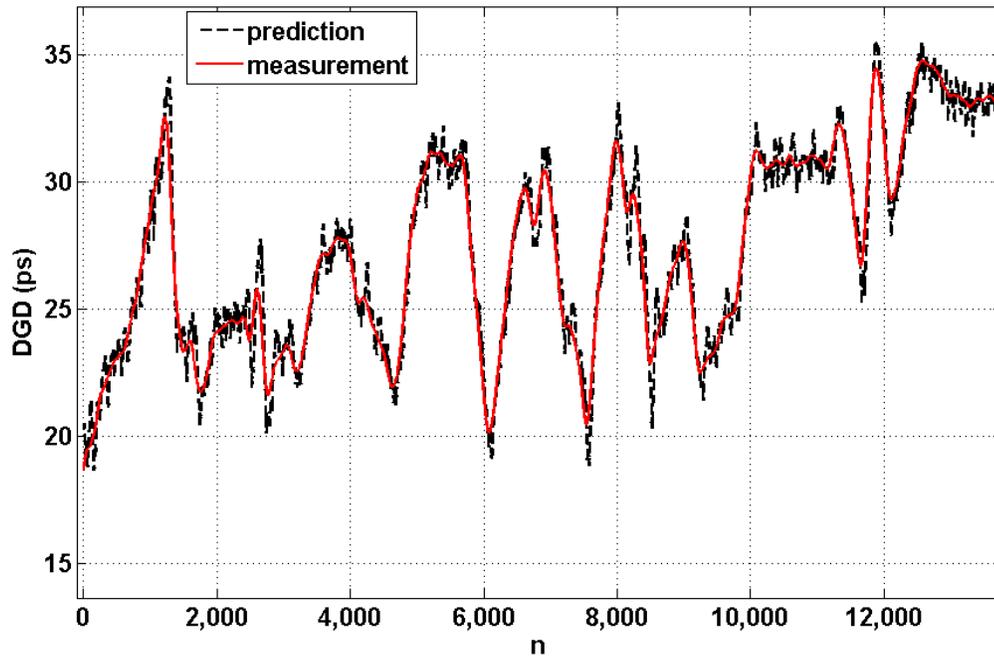


Figure 4.9. Predicted and measured DGD for fiber 2.

We applied the pattern imitation technique to the DGD time series of fiber 2 since more data points are present in this record. Approximately half of the samples were employed as the database (14179) points. For $P = 120$, we obtained $\Phi = 55\%$. While increasing the database size generally improves Φ , the extent of this improvement cannot be estimated here because of the limited number of points in the data record. Applying the Taylor series technique with three derivatives to the same time series (additional terms generated divergences) yielded $\Phi = 45\%$ for $P = 120$. In Table 4-1 we summarize the results of the different models for simulated data and measured data of fiber 2 as a function of the parameter ψ . The table demonstrates that for both simulated and measured data, the AR model yields the longest prediction horizons. Although the AR models perform better, the Taylor expansion technique is less complex. AR models require a suitable amount of data points to be able to build a reliable model that minimizes the least squares error. On the other hand, a Taylor expansion technique employing third order derivatives requires only four measurement points which can be a highly desirable feature, depending on system requirements.

Table 4-1. ψ compared for the three methods applied to simulated and measured data.

$\Phi = 80\%$	Prediction Method		
	AR model	Taylor Expansion	Pattern Imitation
Simulated data	0.77	0.57	0.41
Measured data	0.38	0.28	0.19

4.5 Conclusion

In this chapter, we have examined methods that could potentially be employed to predict PMD impairments sufficiently long in advance to enable rerouting of the affected traffic. We considered in particular three different time series prediction techniques for both simulated and measured data. Of these, applying the AR model to each of the three PMD vector components followed by a Kalman filter yielded the longest predictions horizons for a given level of accuracy in the absence of noise. However, we did not observe an advantage in applying the model to the components of the PMD rather than the DGD directly when noise was present. Our results with simulated Gaussian noise yielded a prediction horizon equal to about 77% of the correlation time while for a time series taken from actual measurement data the horizon was greater than 25% of the correlation time. While the AR/Kalman filter method will not necessarily be optimal for arbitrary systems, other modeling methods may be employed such as ARMA which we found to be more suitable for signals with discontinuities.

Chapter 5

Temporal Autocorrelation Functions of the PMD and Stokes Vectors in the Anisotropic Hinge Model

5.1 Introduction

PMD measurements of buried cables over multiple days indicate that to a reasonable approximation, the time dependence of the PMD of each cable section is described by assuming a Langevin equation for the birefringence [43]. Accordingly, such sections can be modeled as a concatenation of numerous segments each of which possesses a birefringence vector with a fixed orientation but with a magnitude that drifts randomly with time. Averaging over different realizations of the segments yields the temporal autocorrelation function (ACF) of the PMD vector [43], [52]

$$\langle \vec{\tau}(0) \cdot \vec{\tau}(\alpha) \rangle = \langle \tau^2 \rangle \frac{1 - \exp(-|\alpha|/t_d)}{|\alpha|/t_d} \quad (5.1)$$

where t_d is a decorrelation time characteristic of the fiber. In this chapter, we derive analytical expressions for the temporal ACF's of the SOP and the PMD vector in the hinge model. Our studies demonstrate that for large time offsets, the ACF of the PMD vector approaches a constant value that depends principally on the DGD of the last fiber section but is also affected by increasingly diminishing amounts by the DGD of the preceding fiber sections. The accuracy of the procedure is further established through comparison with numerical simulations.

5.2 ACF of the Hinge Matrix

We examine PMD in optical fiber links, illustrated by Figure 5.1., that contain n hinges separated by $n+1$ time-invariant fiber spans. Denoting the time-dependent rotation matrices of the hinges by $H_i(t)$ with $i=1,2,\dots,n$, and the PMD vectors and rotation matrices of the fiber sections by $\vec{\tau}_i$ and R_i respectively, from the PMD concatenation rule

$$\vec{\tau} = \vec{\tau}_{n+1} + R_{n+1}H_n(t)\vec{\tau}_n + R_{n+1}H_n(t)R_{n-1}H_{n-1}(t)\vec{\tau}_{n-1} \cdots + R_{n+1}H_n(t) \cdots R_3H_2(t)R_2H_1(t)\vec{\tau}_1 \quad (5.2)$$

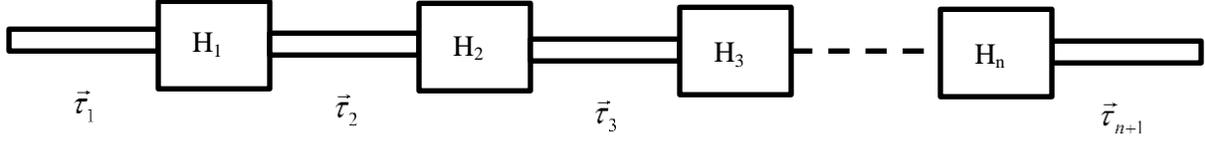


Figure 5.1. Optical fiber link with hinges.

The anisotropic hinges considered here possess invariant rotation axes and a rotation angle $\varphi_i(t)$ that evolves in time according to

$$\frac{\partial \varphi_i}{\partial t} = g(t) \quad (5.3)$$

where $g(t)dt$ is a Wiener process with $\langle g(t) \rangle = 0; \langle g(0)g(\alpha) \rangle = \sigma^2 \delta(\alpha)$. This model is motivated by the measurements of [59] in which the frequency average of the autocovariance function of $\psi(t) = \varphi(t) - \varphi(0)$ displayed an almost linear decay with the time offset. Such behavior is a characteristic of the Wiener process as illustrated in Figure 5.2, which plots the autocovariance of $\psi(t)$ obtained after solving Eq. (5.3) numerically for $\sigma = 1$, averaging over 5000 realizations and normalizing to unity at zero time offset.

In the following analysis, the hinge angles are statistically independent of the rotation axes orientations. Also, these orientations are set to different values uniformly distributed on the Poincare sphere [60] for each realization as this simplifies the derivation of the analytical expressions. However, the case in which the axes remain identical for all realizations will be subsequently examined numerically. The ACF of the rotation matrix of a single hinge is obtained by averaging over realizations

$$\langle H(0)H(\alpha) \rangle \quad (5.4)$$

The ACF can in practice be determined by averaging over values obtained at different light frequencies spaced by more than a frequency correlation length. Hence the assumption of identical axes for all realizations would correspond physically to the axes remaining invariant with a change of frequency. With $H(t) = e^{\varphi(t)\hat{r} \times}$, $\hat{r} = (r_1, r_2, r_3)^T$ where \hat{r}

and $\varphi(t)$ represent the rotation axis and angle respectively, since $\langle \sin(\psi(\alpha)) \rangle = 0$ and $\langle \cos(\psi(\alpha)) \rangle = \exp\left(-\frac{\sigma^2}{2}|\alpha|\right)$ [32],

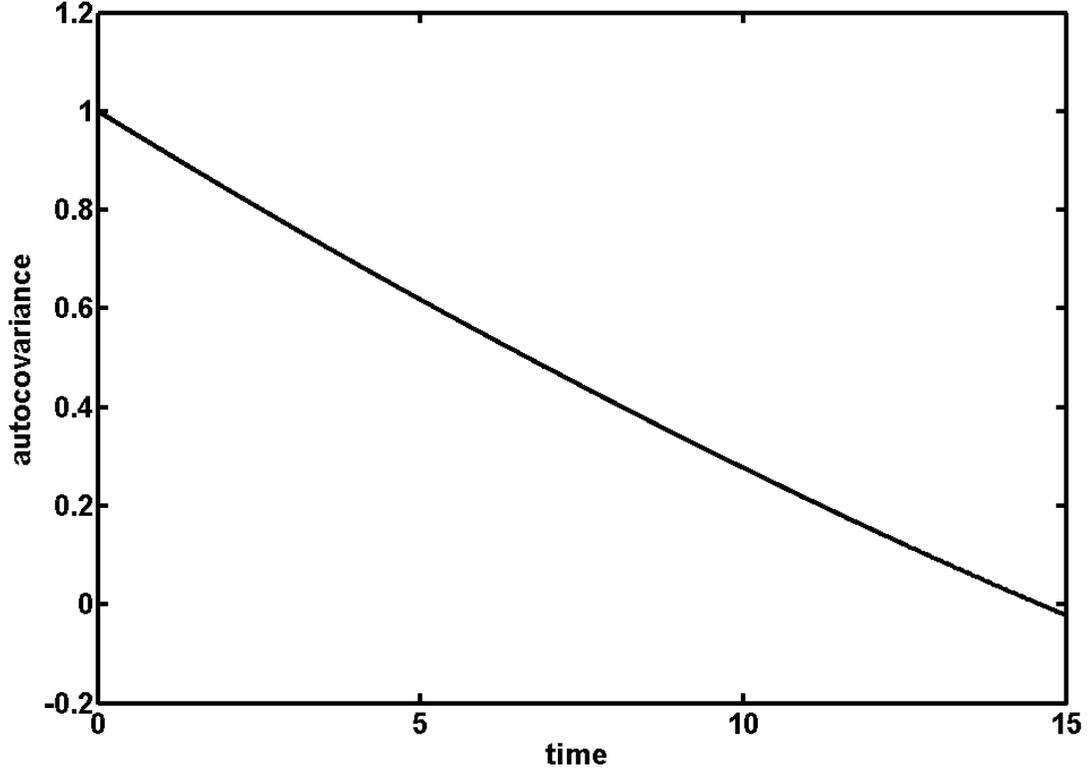


Figure 5.2. Averaged autocovariance of $\psi(t)$ for 5000 realizations.

$$\langle H^T(0)H(\alpha) \rangle = \langle \hat{r}\hat{r} \rangle - \langle (\hat{r} \times)(\hat{r} \times) \rangle \exp\left(-\frac{\sigma^2}{2}|\alpha|\right) \quad (5.5)$$

From $\langle \hat{r}\hat{r} \rangle = \mathbf{I}/3$, where \mathbf{I} is the identity matrix, and $\langle (\hat{r} \times)(\hat{r} \times) \rangle = -2\mathbf{I}/3$ [52],

$$\langle H^T(0)H(\alpha) \rangle = \frac{1}{3}\mathbf{I}\left(1 + 2\exp\left(-\frac{\sigma^2}{2}|\alpha|\right)\right) \quad (5.6)$$

Hence for $\alpha \gg 2/\sigma^2$, the correlation function approaches $\mathbf{I}/3$. Accordingly, a characteristic correlation time, $t_c = 2/\sigma^2$, which corresponds to the variance of the Weiner process, can be associated with the hinge.

5.3 ACF for the PMD Vector

If $\vec{\Omega}_n$ denotes the PMD vector of a fiber link before to the n^{th} hinge, the PMD vector after the n^{th} hinge and the $n+1$: th fiber section is obtained from

$$\vec{\Omega}_{n+1} = \vec{\tau}_{n+1} + R_{n+1} H_n(t) \vec{\Omega}_n \quad (5.7)$$

Following the procedure of [51], Eq. (7) leads to a recurrence relation for the ACF after it is averaged over all fiber and hinge realizations (or equivalently over uncorrelated frequencies) according to

$$C_{n+1}(\alpha) = \langle \vec{\Omega}_{n+1}(0) \cdot \vec{\Omega}_{n+1}(\alpha) \rangle = \left\langle \left(\Omega_n^T(0) H_n^T(0) R_{n+1}^T + \tau_{n+1}^T \right) \left(\tau_{n+1} + R_{n+1} H_n(\alpha) \Omega_n(\alpha) \right) \right\rangle \quad (5.8)$$

Here the PMD vector is written as column matrix and the dot product is replaced by the matrix product of $\Omega_n^T(0)$ and $\Omega_n(\alpha)$. With ensemble averaging, the components $\vec{\tau}_i$'s become statistically independent Gaussian random variables with zero mean that are further independent of the hinge variables. Therefore, the cross terms and products of $\vec{\Omega}_n$ and $\vec{\tau}_{n+1}$ in Eq. (5.8) vanish and

$$C_{n+1}(\alpha) = \langle \tau_{n+1}^2 \rangle + \langle \Omega_n^T(0) H_n^T(0) H_n(\alpha) \Omega_n(\alpha) \rangle \quad (5.9)$$

In the same manner as for the frequency ACF of the PMD vector [51], Eq. (5.9) can be written as a recurrence relation

$$C_{n+1}(\alpha) = \langle \tau_{n+1}^2 \rangle + M_n C_n(\alpha) \quad (5.10)$$

where $C_n(\alpha) = \langle \vec{\Omega}_{n+1}(0) \cdot \vec{\Omega}_{n+1}(\alpha) \rangle$ and $M_n = \text{tr} \left(\langle H_n^T(0) H_n(\alpha) \rangle \right) / 3$. From Eq. (5.6),

$$M_n = \frac{1}{3} \left(1 + 2 \exp \left(-\frac{|\alpha|}{t_{cn}} \right) \right) \quad (5.11)$$

where t_{cn} is the correlation time of n^{th} hinge. By repeatedly applying (5.10) with $C_0 = \langle \tau_1^2 \rangle$, the ACF of an arbitrary number of hinges is found to be

$$\begin{aligned} \langle \vec{\tau}(0) \cdot \vec{\tau}(\alpha) \rangle &= \langle \tau_{n+1}^2 \rangle + \frac{1}{3} \langle \tau_n^2 \rangle \left(1 + 2 \exp\left(-\frac{|\alpha|}{t_{cn}}\right) \right) \\ &+ \frac{1}{3^2} \langle \tau_{n-1}^2 \rangle \left(1 + 2 \exp\left(-\frac{|\alpha|}{t_{cn}}\right) \right) \left(1 + 2 \exp\left(-\frac{|\alpha|}{t_{cn-1}}\right) \right) + \dots \end{aligned} \quad (5.12)$$

Where t_{ci} , $i = 1$ to n are the correlation times of the hinges. For fiber sections with comparable mean DGD, the ACF of the PMD vector is primarily affected by the time correlation properties of the last hinge in the link as the contribution to the ACF decreases by a factor of $1/3$ for each preceding hinge.

As $|\alpha| \rightarrow \infty$, $\langle \vec{\tau}(0) \cdot \vec{\tau}(\alpha) \rangle \rightarrow \sum_{i=0}^n \frac{E[\tau_{i+1}^2]}{3^{n-i}}$. The ACF for the output Stokes vector $\hat{s}_{out}(t)$ is obtained in a similar manner. For n hinges

$$\langle \hat{s}_{out}(0) \cdot \hat{s}_{out}(\alpha) \rangle = \frac{1}{3^n} \prod_{i=1}^n \left(1 + 2 \exp\left(-\frac{|\alpha|}{t_{ci}}\right) \right) \quad (5.13)$$

Accordingly, unlike the PMD vector, each hinge contributes equally to the ACF of the SOP. As $|\alpha| \rightarrow \infty$

$$\langle \hat{s}_{out}(0) \cdot \hat{s}_{out}(\alpha) \rangle \rightarrow \frac{1}{3^n} \quad (5.14)$$

5.4 Simulations

We simulate a 3 hinge optical fiber link with correlation times of $t_{c1} = 2$, $t_{c2} = 5$, $t_{c3} = 6.6$ days. For each realization, the PMD vector of each fiber section is selected from a Gaussian distribution with zero mean and a standard deviation determined by the requirement that the mean DGD of each fiber section equals 10 ps. The hinge rotation angle is obtained from the numerical solution of Eq. (5.3) and rotation axes are randomly oriented on the Poincare sphere. The ACF of the PMD vector, the SOP and the PSP's is determined by averaging over 5000 statistically independent realizations of the fiber link, hinge angles and axes and are subsequently normalized to unity at $\alpha = 0$ and compared as a function of α in days with the analytic results of Eq. (5.12) and (5.13) in Figure 5.3.

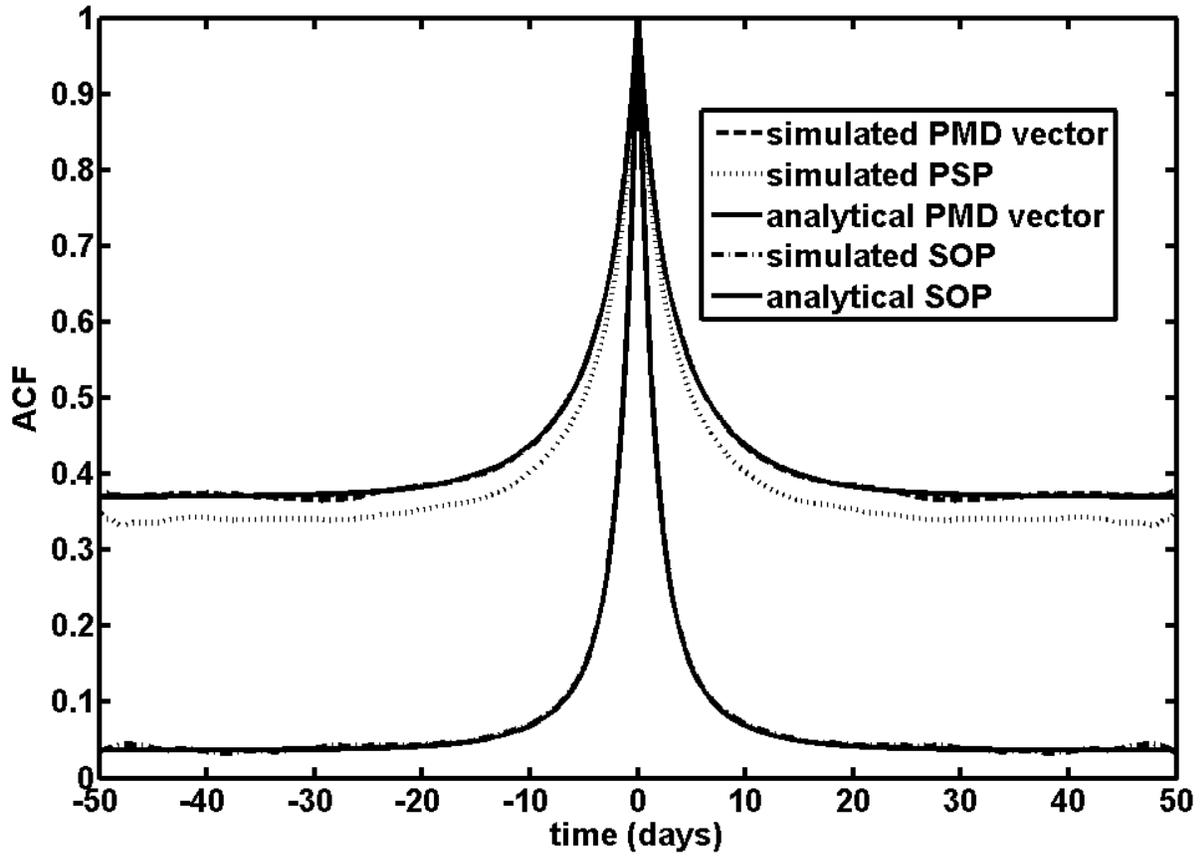


Figure 5.3. Normalized ACF for the PMD vector and the SOP as a function of α in days.

The numerical and the analytic results of Eq. (5.12) and (5.13) displayed as a function of time offset in days nearly coincide as illustrated in Figure 5.3. Further, the ACF of the PSP is very similar to that of the PMD vector, as noted for the frequency ACF in [61].

In Figure 5.4, we plot the simulated ACF of the PMD vector for three different sets of hinge rotation axes that are identical for every numerical realization together with the result of Eq. (5.12).

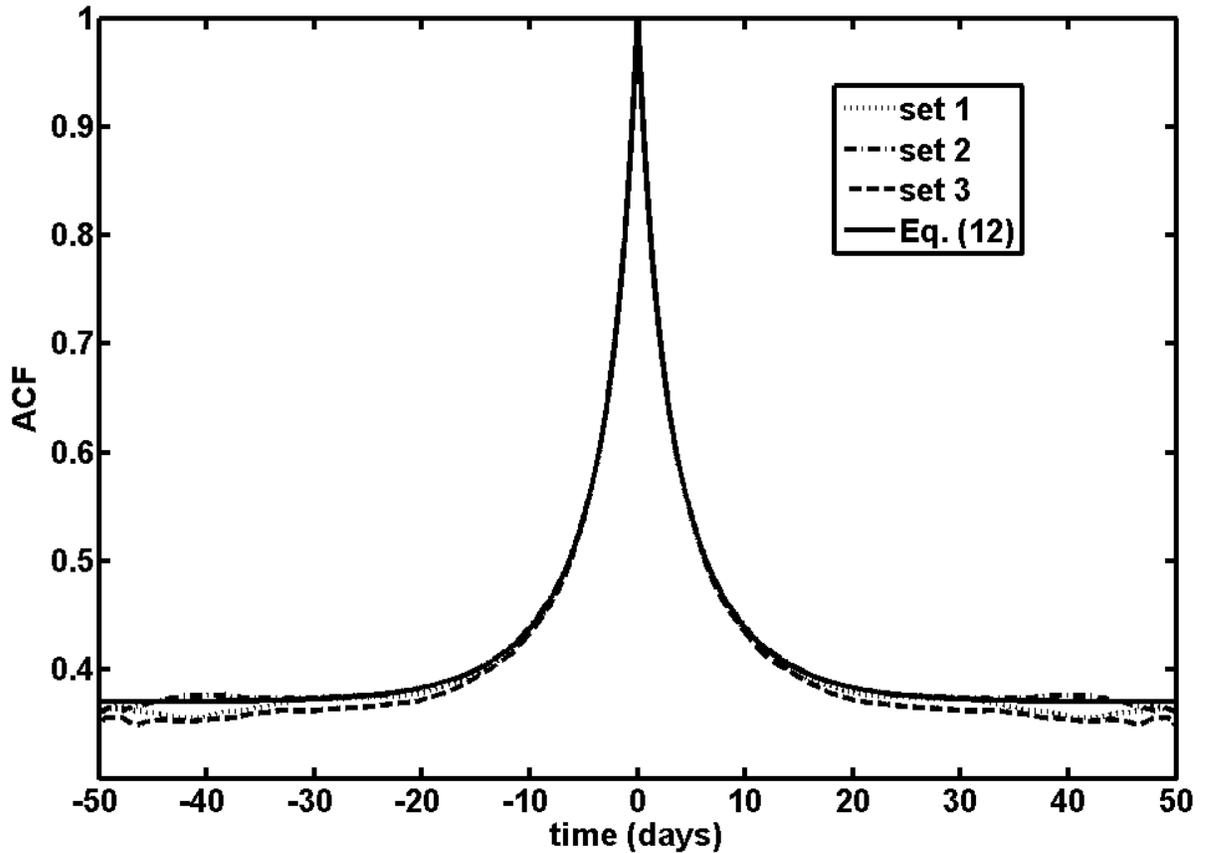


Figure 5.4. Simulated and analytic ACF's of the PMD vector for three different sets of invariant hinge rotation axes.

Figure 5.4 indicates that since the ACF is not affected significantly if the hinge axes remain invariant, the analytical result remains approximately valid.

5.5 Conclusion

We have presented analytical expressions for the ACF of the state of polarization and the PMD vector in the hinge model. Our analysis incorporates a model in which the time evolution of the polarization rotation properties of each hinge is completely described by a single correlation time. While this model is consistent with experimental observations only for certain fiber links, our results may be qualitatively applicable to a wider range of link behavior.

Chapter 6

Simulation of a WDM Polarization Multiplexed Quadrature Phase Shift Keyed (PM-QPSK) System

6.1 Introduction

System performance is judged through the dependence of the SER on the SNR. An improvement in the SNR implies a reduced SER. The magnitude of the improvement enables the system designer to evaluate the importance and cost of using for example better receiver architecture with lower noise or increased input power. The purpose of this chapter is to present a full performance simulation of a WDM PM-QPSK system. This study serves as a starting point to evaluate the impact of PMD on such systems. A schematic diagram of the system is shown in Figure 6.1.

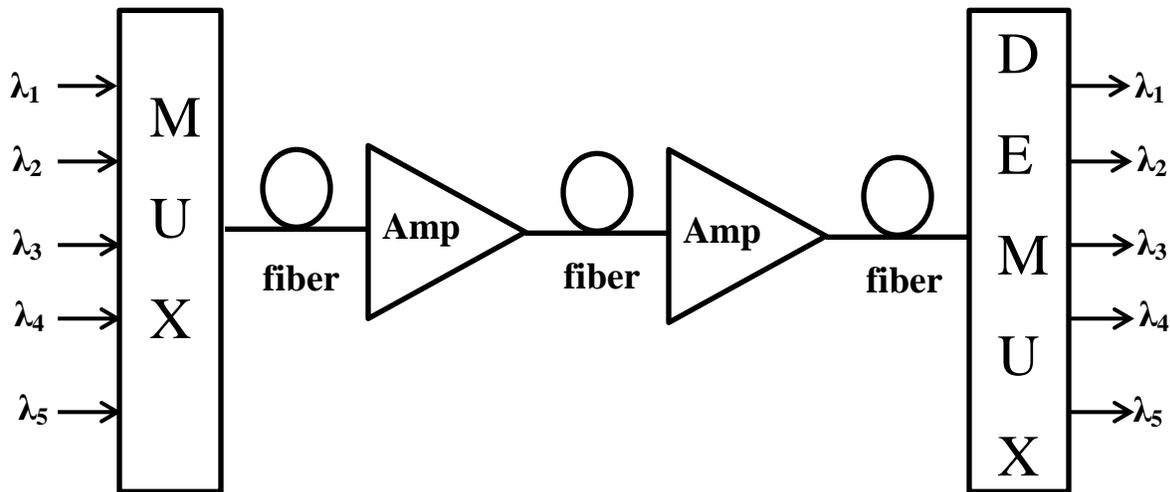


Figure 6.1. A schematic diagram of the simulated system.

Several channels are modulated using an electrical bit stream and multiplexed together then launched into an optical fiber. Amplifiers restore the optical signal to the initial power level periodically along the line. At the receiver the channels are demultiplexed and the electrical bit stream is restored through a receiver bank.

6.2 System Description and Parameters

The system of interest incorporates $N = 5$ channels separated by 50 GHz at a symbol rate of 10.7 GBaud/s per polarization, resulting in a 42.8 Gb/s transmission bit rate. The channel power is set to 4

dBm. The transmission medium consists of eight, 80 km standard single mode fiber (SMF) or large effective area fiber (LEAF) fiber spans making the total length of the link 640km. Both fibers have the following parameters: nonlinear refractive index , $n_2 = 2.64 \times 10^{-20} m^2 / W$, $A_{eff} = 79.6 \mu m^2$, $D_{PMD} = 0.05 ps / \sqrt{km}$, $\alpha = 0.21 dB / km$ but differ on the dispersion coefficient $D = -2\pi c \beta_2 / \lambda_0^2$ which is $D = 17 ps / km / nm$ for the SMF and $D = 4 ps / km / nm$ for the LEAF fibers. After each fiber span, the signal is amplified by an amount equal to its power loss in the span. Once each channel of the input field is propagated through the fiber by solving the Manakov PMD equation numerically, and detected with a polarization diversity optical coherent receiver, the cumulative effect of ASE is modeled by adding a Gaussian random value of zero mean and standard deviation consent with the specified OSNR (optical SNR) to the center of each symbol in the probe channel. That the ASE can be loaded at the receiver end rather than distributed along the link without appreciably affecting the accuracy of the model was established for PM-QPSK systems in [62]. Subsequently, an electronic post compensation filter removes the total accumulated chromatic dispersion and PMD. Sampling the resulting optical field $|A_{RX}(t)\rangle$ at the transmission rate yields the received digital symbols $|r_m\rangle = |A_{RX}(mT)\rangle$. In the polarization multiplexed WDM system of interest, the optical field can be expressed as $|A(z,t)\rangle = \sum_{n=0}^{N-1} |A_n(z,t)\rangle e^{i\Omega_n t}$ which superimposes the fields of five optical channels with center frequency differences Ω_n .

6.3 Initial Field Temporal Profile

Consider one of the channels' Jones vectors $|A_i(z,t)\rangle$. A random symbol stream is generated for A_{ix} and A_{iy} and for all channels. Initially at the input, the pulse shape for each symbol is a RZ square root raised cosine filter profile with a unity rolloff factor chosen because it minimizes intersymbol interference [63]. Each symbol is assigned one of the allowed phases at random. Referring back to Eq. (1.1), we can write for channel i ,

$$A_{ix} = \sqrt{2} \sum_{m=0}^M a(t - mT_s) e^{-j\theta_m} \quad (6.1)$$

where $a(t)$ and θ_m are the pulse shape and the symbol phase respectively and $M + 1$ is the number of symbols. Expanding the complex phase, we can rewrite Eq. (6.1) as

$$A_{ix} = \sum_{m=0}^M (xI_m + jxQ_m) a(t - mT_s) \quad (6.2)$$

where $xI_m, xQ_m = \pm 1$ depending on the symbol m . The I and Q components were mentioned in section 1.2. Hence to generate the random symbol stream, we generate random streams of ± 1 and $\pm j$ for the required number of symbols and multiply them by the time shifted version of $a(t)$. We choose for our simulations a pattern length of $M + 1 = 16$ symbols. The real part of A_{ix} is displayed in Figure 6.2 .

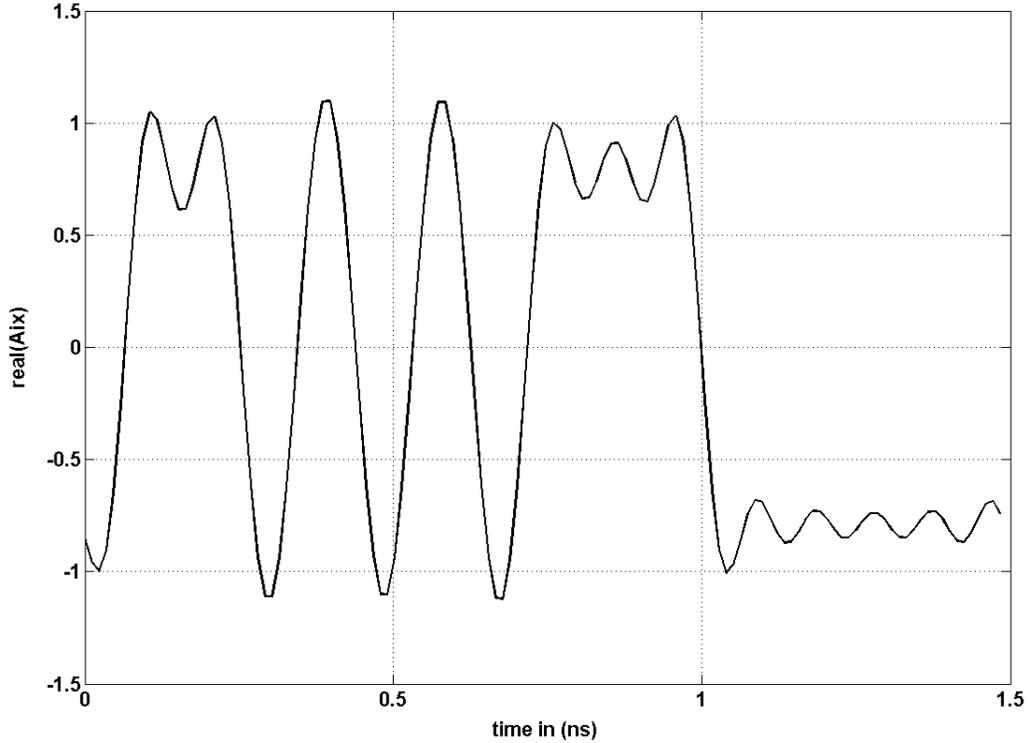


Figure 6.2. Real part of A_{ix} versus time in ns.

After the $z = 0$ temporal profile of each channel and polarization is determined, a random timing delay is applied to each channel because in practice they are not synchronized. Also, a random rotation matrix is applied to the Jones vector of each channel to ensure that the channels start off with random polarization states. The resulting superposition of the electric fields is propagated through the fiber by solving the Manakov PMD equation numerically. The method used to achieve this is the split step Fourier method (SSFM) and is described briefly in the next section.

6.4 SSFM for Numerical Solutions to the Manakov PMD equation

In the split step procedure we rewrite Eq. (1.67) as

$$\frac{\partial}{\partial z}|A\rangle = (\hat{N} + \hat{D})|A\rangle \quad (6.3)$$

where the nonlinear operator \hat{N} and the dispersion operator \hat{D} are given by

$$\hat{N} = -j\gamma \frac{8}{9} \langle A|A \rangle \quad (6.4)$$

$$\hat{D} = \frac{j\beta_2}{2} \frac{\partial^2}{\partial t^2} - \vec{b} \cdot \vec{\sigma} \frac{\partial}{\partial t} - \frac{\alpha}{2} \quad (6.5)$$

For a step h , the solution to Eq. (6.3) is simply

$$|A(z+h, t)\rangle = e^{(\hat{N}+\hat{D})h} |A(z, t)\rangle \quad (6.6)$$

The exponential is difficult to evaluate, hence it's approximated by

$$e^{(\hat{N}+\hat{D})h} \approx e^{h\hat{N}} e^{h\hat{D}} \quad (6.7)$$

Such an approximation is exact if \hat{N} and \hat{D} commute. For a small enough step h , Eq. (6.7) is used to solve the Manakov equation by treating each operator separately. The dispersion step can be solved entirely in the frequency domain and the result is

$$|A_D(z, t)\rangle = IFFT \left(\exp \left(-h \left(\frac{j\beta_2 \omega^2}{2} + j\omega \vec{b} \cdot \vec{\sigma} + \frac{\alpha}{2} \right) \right) FT(|A(z, t)\rangle) \right) \quad (6.8)$$

where FT and $IFFT$ denote the Fourier transform and inverse Fourier transform respectively. After performing the dispersion step, the nonlinear step is performed as

$$|A(z+h,t)\rangle = \exp\left(-j\gamma h \frac{8}{9} \langle A_D | A_D \rangle\right) |A_D(z,t)\rangle \quad (6.9)$$

Therefore, the fiber is divided into small steps of size h and the initial field is propagated from the input to the output.

6.5 Simulation Results

Our simulations were performed with two constant values of the step sizes, h , of the split step Fourier routine. The first of these is given by the smaller of the h that insures a nonlinear phase rotation per step below 0.01 and the h that restricts the walk-off between the two edge channels to less than 16% of the symbol period [64]. Subsequently we repeated the calculations with h increased from the value above by a factor of 10. The number of birefringent segments in each span is given by its length divided by h . The PMD vector of each segment, $h\omega\vec{b}(z)$ of Eq.(6.8), was selected from three independent zero mean Gaussian random variables with standard deviation $D_{pmd}(h/3)^{1/2}$. The xI and xQ components of the output electric field of the eighth symbol of the central channel were finally employed as the control quantities for the multicanonical algorithm which evaluates the joint density function conditional on the initial transmission of a (11) symbol.

As described in section 1.5, we apply the multicanonical method to obtain the pdf. The present simulation employs the following random variables: 1) the value of each bit in the bit streams of each channel [65], 2) the phase shift between channel symbol streams 3) the angles of the random rotation matrices for each channel, 4) the PMD vector components of each section and 5) the ASE noise variables.

In a proposed transition, each bit (except for the test bit) is inverted with a probability such that on average ≈ 2.5 bits are changed while symmetric random values are added to the random time delay of each channel and to the rotation angles. The normalized ASE variables are each modified by adding a Gaussian variable with zero mean and unity standard deviation, g . Subsequently, as in [21] and [26], g is replaced by $g + v$, where v is given by a small uniformly distributed random value. The new noise values are accepted with probability $\text{minimum}\left\{1, \exp\left(-\frac{(g+v)^2}{2}\right) / \exp\left(-\frac{g^2}{2}\right)\right\}$. The same method is employed to perturb the normalized PMD vector components which are also selected from a Gaussian distributed random distribution with zero mean. Finally, the modified PMD variables are multiplied by $D_{pmd}(h/3)^{1/2}$ and the ASE variables are scaled to reproduce the physical OSNR.

We display in Figure 6.3, the contours of the base 10 logarithm of the joint pdf of the xI, xQ components in a single polarization of a single mode fiber for the central bit when this bit is launched with $xI = xQ = 1$ for a OSNR of 15dB. These contours, which describe pdf values $\geq 10^{-11}$ were obtained after four iterations of the multicanonical procedure each of which employed the split step

Fourier method with 3×10^5 samples that were propagated through the entire link. The Monte Carlo result additionally is present in Figure 6.3 as the dotted contours.

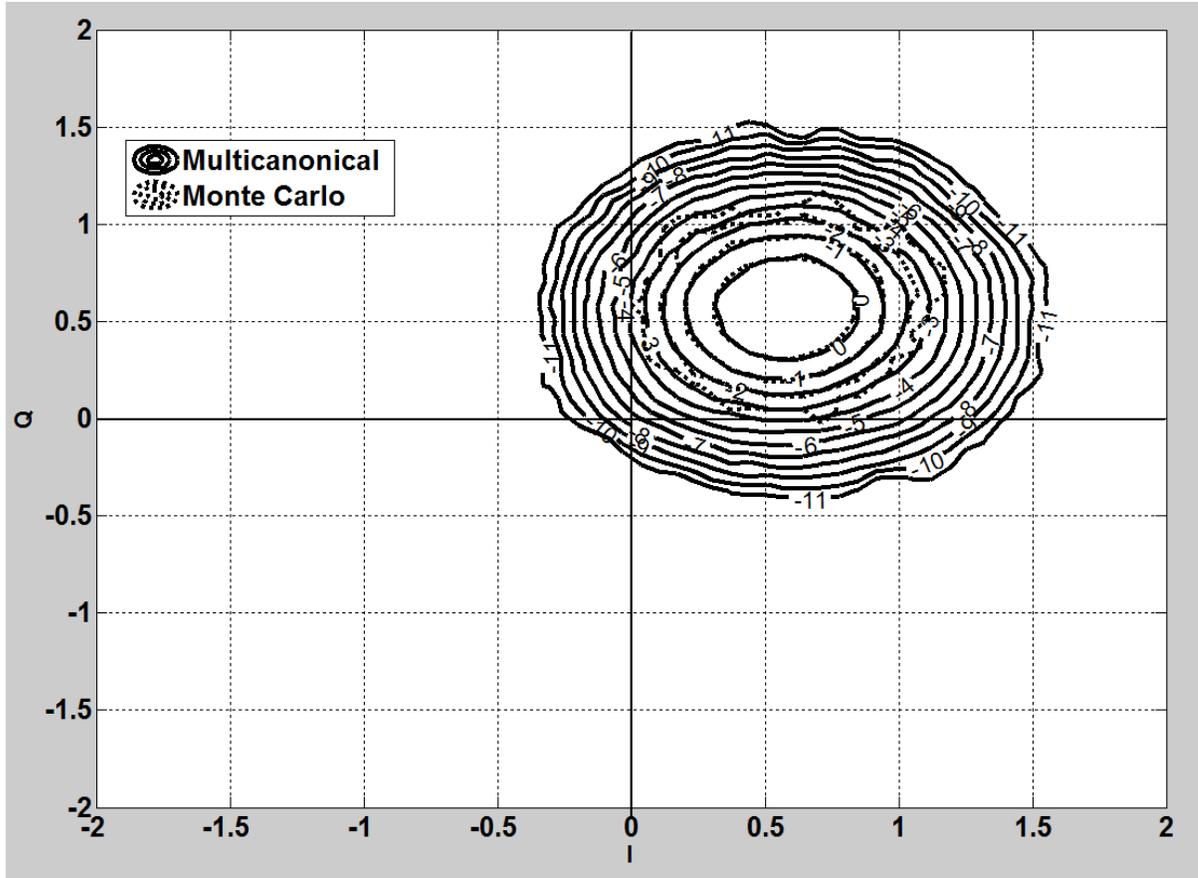


Figure 6.3. Log_{10} of the joint pdf of the xI, xQ components for a $xI = xQ = 1$ symbol launched into a single polarization in the center channel, for four 3×10^5 multicanonical samples per iteration and the Monte Carlo procedure. The smallest pdf value shown is 10^{-11} .

Figure 6.4 and Figure 6.5 plot the SER vs the OSNR for both fiber types with both $h_s = 0.69 \text{ km}$ and $h_t = 6.9 \text{ km}$, where h_s is the step size employed in the SSFM routine. For both fibers, the larger step size underestimates the SER for all OSNR values; however, the error is less than $\approx 0.5 \text{ dB}$ over most of the range investigated. While this value, as expected, increases for larger nonlinearities, even for a SMF at an OSNR of 20 dB, the larger step size calculations only underestimate $\text{log}_{10}(\text{SER})$ by 1.2 dB. Hence, especially for analyzing trends, larger step sizes can often be employed to reduce

calculation times. We have verified that the results converged for h_s , as for an OSNR of 20 dB, the ratio of the SER's obtained using h_s and $h_s/2$ is 1.1.

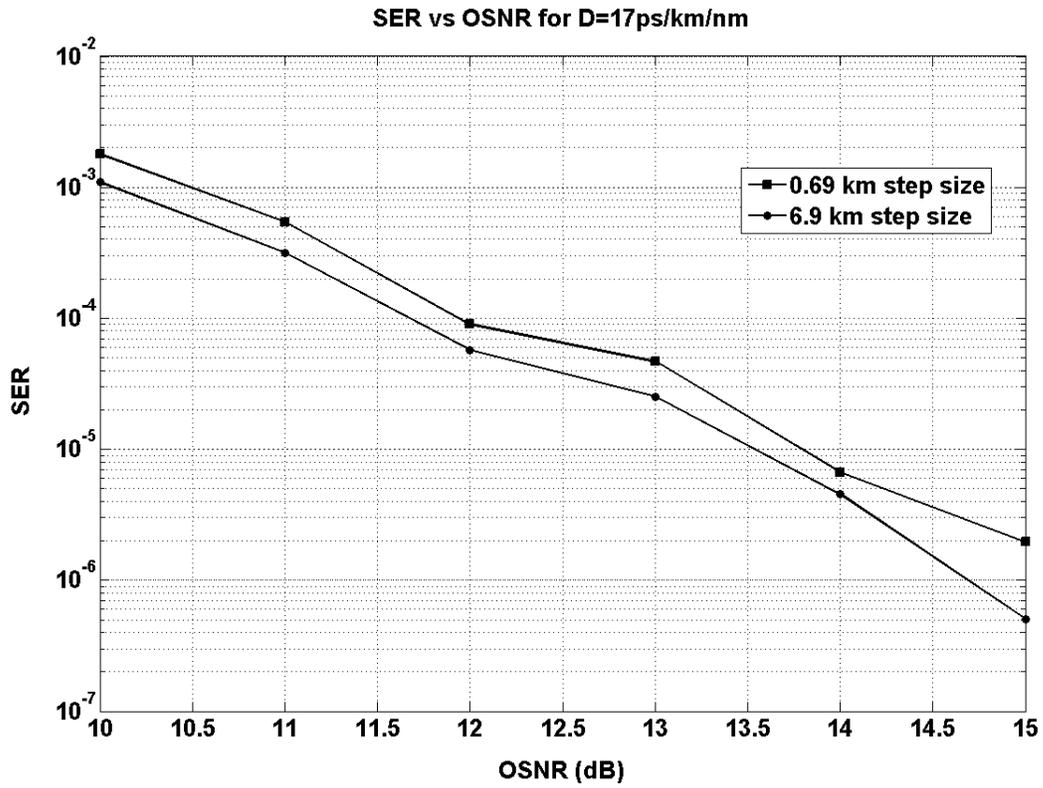


Figure 6.4. The variation in the SER with OSNR for a SMF for two step sizes.

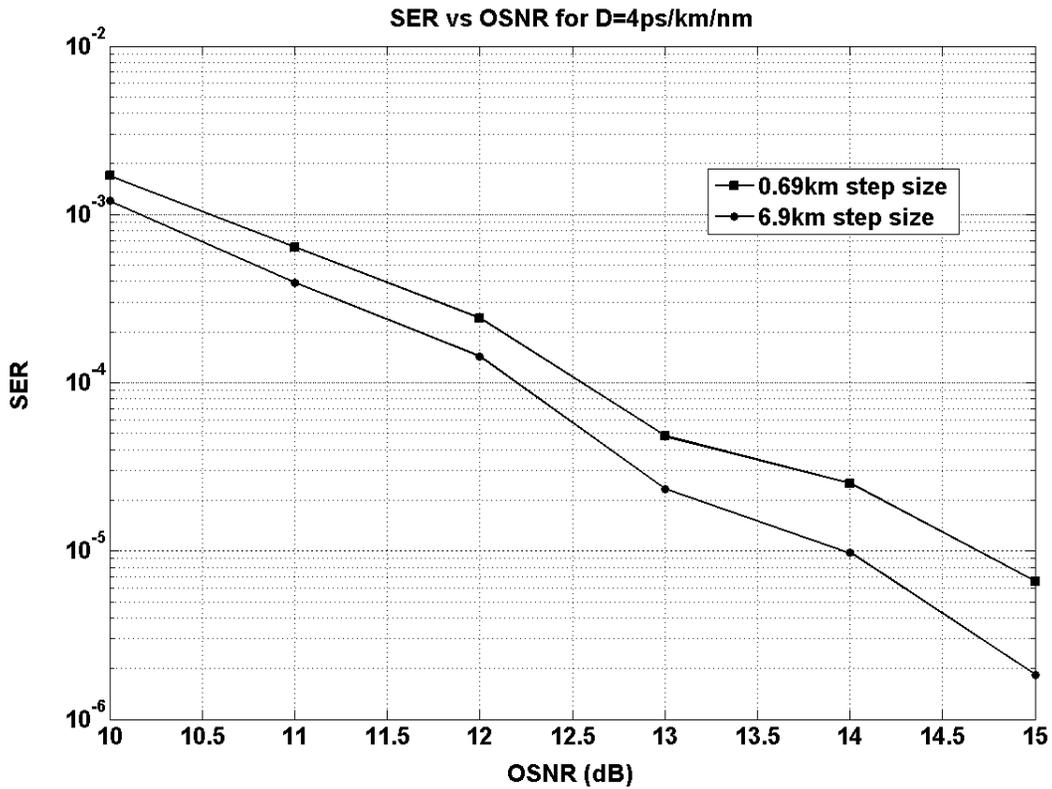


Figure 6.5. As in Figure 6.4 but for a LEAF.

6.6 Conclusion

We have analyzed a WDM-PM-QPSK system in the presence of both fiber nonlinearities and PMD with the multicanonical procedure. Our studies demonstrate that even for large propagation step sizes the errors in the calculated SER are less than 1.2dB for an OSNR range of 10 to 20dB for SMF. The associated reduction in computation time considerably extends the applicability of the method in the context of numerous practical system models.

Chapter 7

Conclusions and Future Work

In this thesis, we have examined several problems related to dynamics of polarization and PMD in optical communication systems. They fall under basically three categories: modeling and measurement of high speed polarization transients, modeling time varying PMD effects and system performance assessment.

Mechanical shocks imparted by the free fall of a steel ball generate rapid polarization variations, the largest reported to date. Given their reproducibility, these measurements serve as diagnostic tools for both the DCM devices and the electronic polarization tracking system. Our proposed model shows that the active length of the fiber is the main contributor to the speed of the transients. These measurements can be extended by repeating the same experiments for fiber spools. More acoustic measurements on the device could lead to a better model and hence better understanding of what other properties of the DCM come into play.

We have shown that it is possible to determine the location of temporal variation of the polarization through a statistical measurement. The accuracy of the method increases if more wavelengths are employed. The technique is attractive because it is independent of the hinge behavior. The described technique holds the potential for locating any number of hinges in a link provided accurate characterization of the temporal and frequency characteristics of hinges become available.

The hinge model has been employed to study the time variation of PMD. We have shown that it might be possible to predict an outage due to a high DGD value using a linear prediction model. Modern systems however, rely on DSP to generate a time varying model for the channel to eliminate the effect of PMD. A combination between the simple model presented in this thesis and the current state of the art techniques might result in relaxing the complexity and speed of the currently employed compensators. Also, applying these linear prediction techniques to the elements of the Jones matrix or the rotation matrix can replace the current algorithms employed in the DSP processor as a simpler alternative for inverting the channel PMD characteristics. Another important extension would be applying more complicated nonlinear models based on, for example, neural networks and assessing the predictive ability and whether the complexity added by employing these models is warranted.

Using a physically reasonable stochastic model for the hinges, analytical expressions for the ACF of the PMD vector and the SOP have been derived as a function of a single correlation time associated with each hinge. Our results provide insight into which parameters affect the ACF the most. Provided with field measurements, the results we obtained may be employed to determine the locations of the hinges in links which exhibit stochastic hinge behavior. ACF's of the squared DGD and the PSP's can also be studied.

We have assessed the performance of a state-of-the-art PM-QPSK system through the evaluation of the SER vs the OSNR at the receiver. We have demonstrated that over a range of 10dB, the SER varies by no more than 1.2dB when the step size is increased by a factor of 10. Hence, for an order of

magnitude estimate, we can speed up computation by relaxing the requirements on the step size. Further studies of interest on such a system would be analysis of the effect of the symbol pattern length on the SER. The system examined can be further studied to map the impact PMD has on the performance. In particular, since PMD introduces some dispersion, it is expected that higher values of PMD might actually be more desired in a system as it helps mitigate the nonlinear effects. Optimal values for PMD are expected for such systems.

Bibliography

- [1] G. P. Agrawal, "Coherent Lightwave Systems," in *Fiber-Optic Communication Systems, third edition*, John Wiley and Sons, 2002, pp. 478-517.
- [2] M. Nakazawa, K. Kikuchi and T. Miyazaki, "Coherent Optical Communications: Historical Perspectives and Future Directions," in *High Spectral Density Optical Communication Technologies*, Springer, 2010, pp. 11-49.
- [3] L. Guifang, "Recent advances in coherent optical communication," *Advances in optics and photonics*, vol. 1, no. 2, pp. 279-307, 2009.
- [4] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems*, John Wiley & Sons, 2005.
- [5] I. Fatadin, S. J. Savory and D. Ives, "Compensation of quadrature imbalance in optical QPSK coherent receiver," *Photonics Technology Letters*, vol. 20, no. 20, pp. 1733-1735, 2008.
- [6] C. Laperle, B. Villeneuve, Z. Zhang, D. McGhan, H. Sun and M. O'Sullivan, "WDM and PMD performance of a coherent 40Gbits/s dual polarization QPSK transceiver," *J. Lightwave Technology*, vol. 26, no. 1, pp. 168-175, 2008.
- [7] K. Roberts, M. O'Sullivan, K.-T. Wu, H. Sun, A. Awadalla, D. J. Krause and C. Laperle, "Performance of Dual-Polarization QPSK for Optical Transport Systems," *J. Lightwave Technology*, vol. 27, no. 16, pp. 3546-3559, 2009.
- [8] C. D. Poole and N. Jonathan, "Polarization Effects in Lightwave Systems," in *Optical fiber telecommunications IIIA*, Academic Press, 1997, pp. 114-161.
- [9] J. P. Gordon and H. Kogelink, "PMD fundamentals: Polarization mode dispersion in optical fibers," *PNAS*, vol. 97, no. 9, pp. 4541-4550, 2000.
- [10] A. Yariv and P. Yeh, "Photonics: Optical Electronics in modern communications", Oxford University Press, 2006.
- [11] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single mode fibres," *Electronics Letters*, vol. 22, no. 19, pp. 1113-1114, 1986.
- [12] C. D. Poole, J. H. Winters and J. A. Nagel, "Dynamical equation for polarization dispersion," *Optics Letters*, vol. 16, no. 6, pp. 372-374, 1991.

- [13] G. P. Agrawal, "Nonlinear Fiber Optics", Academic Press, 2007.
- [14] C. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," *Journal of Engineering Mathematics*, vol. 36, no. 1, pp. 113-136, 1999.
- [15] H. Kogelnik, R. M. Jopson and L. E. Nelson, "Polarization mode dispersion," in *Optical fiber telecommunications, IVB*, Academic Press, 2001, pp. 725-861.
- [16] F. Curti, B. Diano, G. DeMarchis and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single mode fibers," *J. Lightwave Technology*, vol. 8, no. 8, pp. 1162-1166, 1990.
- [17] C. D. Poole and G. J. Foschini, "Statistical theory of polarization dispersion in single mode fibers," *J. Lightwave Technology*, vol. 9, no. 11, pp. 1439-1456, 1991.
- [18] M. Brodsky, N. J. Frigo and M. Tur, "Polarization Mode Dispersion," in *Optical Fiber Telecommunications V A*, Academic Press, 2008, pp. 605-669.
- [19] D. Yevick, "Multicanonical communication system modeling - application to PMD statistics," *Photonics Technology Letters*, vol. 14, no. 11, pp. 1512-1514, 2002.
- [20] T. Lu and D. Yevick, "Efficient multicanonical algorithms," *Photonics Technology Letters*, vol. 17, no. 4, pp. 861-863, 2005.
- [21] R. Holzlohner and C. R. Menyuk, "Use of multicanonical monte carlo simulations to obtain accurate bit error rates in optical communication systems," *Optics Letters*, vol. 28, no. 20, pp. 1894-1896, 2003.
- [22] Y. Yadin, M. Shtaf and M. Orensteun, "Bit error rate of optical DPSK in fiber systems by multicanonical Monte Carlo simulations," *Photonics Technology Letters*, vol. 17, no. 6, pp. 1355-1357, 2005.
- [23] W. Pellegrini, J. Zweck, C. R. Menyuk and R. Holzlohner, "Computation of bit error ratios for a dense WDM system using the noise covariance matrix and multicanonical Monte Carlo methods," *Photonics Technology Letters*, vol. 17, no. 8, pp. 1644-1646, 2005.
- [24] A. Bononi and L. A. Rusch, "Multicanonical Monte Carlo for simulation of optical links," *Optical and Fiber Communications Reports*, vol. 7, pp. 373-413, 2011.
- [25] D. Yevick, "Multicanonical evaluation of joint probability density functions in communication system modeling," *Photonics Technology Letters*, vol. 15, no. 11, pp. 1540-1542, 2003.

- [26] B. Tromborg, M. Reimer and D. Yevick, "Multicanonical evaluation of the tails of the probability density function of semiconductor optical amplifier output power fluctuations," *Journal of Quantum Electronics*, vol. 46, no. 1, pp. 57-61, 2010.
- [27] B. Koch, A. Hidayat, H. Zhang, V. Mirvoda, M. Lichtinger, D. Sandel, and R. Noe, "Optical endless polarization stabilization at 9krad/s with FPGA based controller," *Photonics Technology Letters*, vol. 20, no. 12, pp. 961-963, 2008.
- [28] P. Krummrich and K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," in *Optical fiber communication conference*, Anaheim, 2004.
- [29] M. Reimer, D. Dumas, G. Soliman, D. Yevick and M. O'Sullivan, "Polarization evolution in dispersion compensation modules," in *Optical Fiber Communication Conference*, San Diego, 2009.
- [30] K. Roberts, M. O'Sullivan, K. Wu, H. Sun, A. Awadalla, D. J. Krauss, and C. Laperle, "Performance of dual polarization QPSK optical transport system," *J. Lightwave Technology*, vol. 27, no. 16, pp. 3546-3559, 2009.
- [31] W. Shieh and H. Kogelnik, "Dynamic Eigenstates of Polarization," *Photonics Technology Letters*, vol. 13, no. 1, pp. 40-42, 2001.
- [32] P. K. A. Wai and C.R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 19, no. 19, pp. 1517-1519, 1994.
- [33] J. N. Damask, "Statistical properties of polarization in fiber," in *Polarization optics in telecommunications*, Springer, 2005, p. 391.
- [34] H. E. Engan, B. Y. Kim, J. N. Blake, and H. J. Shaw, "Propagation and optical interaction of guided acoustic waves in two-mode optical fibers," *Lightwave technology*, vol. 6, no. 3, p. 428-436, 1988.
- [35] Karl E Graff, "Wave Motion in Elastic Solids", Dover, 1975.
- [36] B. M. Beadle and J. Jarzynski, "Measurement of speed and attenuation of longitudinal elastic waves in optical fibers," *Optical engineering*, vol. 40, no. 10, p. 2115-2119, 2001.
- [37] Chin-Lin Shen, "Birefringence in single mode fibers," in *Foundations for guided wave optics*, Wiley, 2007.

- [38] R. Ulrich, S. C. Rashleigh and W. Eickhoff, "Bending-induced birefringence in single mode fibers," *Optics Letters*, vol. 5, no. 6, pp. 273-275, 1980.
- [39] A. Galtarossa, L. Palmieri, M. Schiano and T. Tambosso, "Measurements of beat length and perturbation length in long single mode fibers," *Optics Letters*, vol. 25, no. 6, pp. 384-386, 2000.
- [40] R. Caponi, B. Riposati, A. Rossaro and M. Schiano, "WDM system impairments due to highly correlated PMD spectra of buried optical cables," *Electronic Letters*, vol. 38, no. 14, pp. 737-738, 2002.
- [41] C. T. Allen, P. K. Kondamuri, D. L. Richards and D. C. Hague, "Measured temporal and spectral PMD characteristics and their implications for network-level mitigation approaches," *J. Lightwave Technology*, vol. 21, no. 1, pp. 79-86, 2003.
- [42] A. Mecozzi, C. Antonelli and M. Brodsky, "Statistics of the polarization mode dispersion dynamics," *Optics Letters*, vol. 32, no. 20, pp. 3032-3034, 2007.
- [43] C. Antonelli, C. Colamarino, A. Mecozzi and M. Brodsky, "A model for temporal evolution of PMD," *Photonics Technology Letters*, vol. 20, no. 12, pp. 1012-1014, 2008.
- [44] M. Brodsky, P. Magill and N. J. Frigo, "Polarization mode dispersion of installed recent vintage fiber as a parameteric function of temperature," *Photonics Technology Letters*, vol. 16, no. 1, pp. 209-211, 2004.
- [45] M. Brodsky, M. Boroditsky, P. Magill and N.J. Frigo, "Persistence of spectral variations in DGD statistics," *Optics Express*, vol. 30, no. 11, pp. 4090-4095, 2005.
- [46] M. Brodsky, N. J. Frigo, M. Broditsky and M. Tur, "Polarization mode dispersion of installed fibers," *J. Lightwave Technology*, vol. 24, no. 12, pp. 4584-4559, 2006.
- [47] C. Antonelli and A. Mecozzi, "Theoretical characterization and system impact of the hinge model of PMD," *J. Lightwave Technology*, vol. 24, no. 11, pp. 4064-4074, 2006.
- [48] M. Boroditsky, M. Brodsky, N. J. Frigo, P. Magill, C. Antonelli and A. Mecozzi, "Outage probabilities for fiber routes with finite number of degrees of freedom," *Photonics Technology Letters*, vol. 17, no. 2, pp. 345-347, 2005.
- [49] A. Galtarossa, D. Grosso, L. Palmieri and L. Schenato, "Reflectometric characterization of hinges in optical fiber links," *Photonics Technology Letters*, vol. 20, no. 10, pp. 854-856, 2008.
- [50] A. Galtarossa, D. Grosso, L. Palmieri and L. Schenato, "Distributed polarization mode

- dispersion measurement in fiber links by polarization sensitive reflectometric techniques," *Photonics Technology Letters*, vol. 20, no. 23, pp. 1944-1946, 2008.
- [51] M. Karlsson and J. Brentel, "Autocorrelation function of the polarization mode dispersion vector," *Optics Letters*, vol. 24, no. 14, pp. 939-941, 1999.
- [52] M. Karlsson, J. Brentel and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *J. Lightwave Technology*, vol. 18, no. 7, pp. 941-951, 2000.
- [53] G. Keiser, *Optical Communication Essentials*, McGraw Hill, 2003.
- [54] R. Cigliutti, A. Galtarossa, M. Giltrelli, D. Grosso, A.W.R. Leitch, L. Palmieri, S. Santoni, L. Schenato and D. Waswa, "Design, estimation and experimental validation of optical polarization mode dispersion compensator in 40 Gbit/s NRZ and RZ optical systems," *Optical fiber technology*, vol. 15, no. 3, pp. 242-250, 2009.
- [55] J. Li, G. Biondini, W. L. Kath and H. Kogelnik, "Anisotropic hinge model for polarization-mode dispersion in installed fibers," *Optics Letters*, vol. 33, no. 16, pp. 1924-1926, 2008.
- [56] Monson H. Hayes, "Lattice Filters and Wiener Filtering," in *Statistical Digital Signal Processing and Modeling*, John Wiley and Sons, 1996, pp. 289-370.
- [57] J.D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, no. 8, pp. 845-848, 1987.
- [58] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627-1639, 1964.
- [59] M. Brodsky, J. Martinez, N. J. Frigo and A. Sirenko, "Dispersion compensation module as a polarization hinge," in *ECOC*, Glasgow, 2005.
- [60] J. Li, G. Biondini, W. L. Kath and H. Kogelnik, "Outage statistics in a waveplate hinge model of polarization mode dispersion," *J. Lightwave technology*, vol. 28, no. 13, pp. 1958-1968, 2010.
- [61] M. Shtaif and A. Mecozzi, "Study of the frequency autocorrelation function of the differential group delay in fibers with polarization mode dispersion," *Optics letters*, vol. 25, no. 10, pp. 707-709, 2000.
- [62] A. Bononi, P. Serena and N. Rossi, "Nonlinear signal-noise interactions in dispersion managed links with various modulation formats," *Optical Fiber Technology*, vol. 16, no. 2, pp. 73-85, 2010.

- [63] S. Haykin, "Communication Systems", John Wiley and Sons, 2001.
- [64] O. V. Sinkin, R. Holzlohner, J. Zweck and C. Menyuk, "Optimization of the Split-Step Fourier Method in Modeling Optical-Fiber Communications Systems," *J. Lightwave Technology*, vol. 20, no. 1, pp. 61-68, 2003.
- [65] L. Gerardi, M. Secondini and E. Forestieri, "Performance evaluation of WDM systems through multicanonical Monte Carlo simulations," *Lightwave Technology*, vol. 29, no. 6, pp. 871-879, 2011.