OPTICAL AND FIBER COMMUNICATIONS REPORTS

OPTICAL AND FIBER COMMUNICATIONS REPORTS

The Optical and Fiber Communications Reports (OFCR) book series provides a survey of selected topics at the forefront of research. Each book is a topical collection of contributions from leading research scientists that gives an up-to-date and broad-spectrum overview of various subjects.
The main topics in this expanding field will cover for example:

- specialty fibers (periodic fibers, holey fibers, erbium-doped fibers)
- broadband lasers
- optical switching (MEMS or others)
- polarization and chromatic mode dispersion and compensation
- long-haul transmission
- optical networks (LAN, MAN, WAN)
- protection and restoration
- further topics of contemporary interest.

Including both general information and a highly technical presentation of the results, this series satisfies the needs of experts as well as graduates and researchers starting in the field. Books in this series establish themselves as comprehensive guides and reference texts following the impressive evolution of this area of science and technology.
The editors encourage prospective authors to correspond with them in advance of submitting a manuscript. Submission of manuscripts should be made to one of the editors. See also http://springeronline.com/series/ 4810.

Andrea Galtarossa Curtis R. Menyuk
Editors

Polarization Mode Dispersion

With 109 Figures

Springer

Andrea Galtarossa
Department of Information
Engineering
University of Padova
Via Gradenigo 6/B
35131 Padova
Italy
andrea.galtarossa@unipd.it

Curtis R. Menyuk
Department of Computer Science
and Electrical Engineering
University of Maryland
Baltimore Country
1000 Hilltop Circle
Baltimore, MD 21250
USA
menyuk@umbc.edu

Library of Congress Control Number: 2005925979

ISBN-10: 0-387-23193-5 Printed on acid-free paper.
ISBN-13: 978-0387-23193-8

Printed in the United States of America. (SPI/EB)

9 8 7 6 5 4 3 2 1

springeronline.com

To Hermann Haus

"For the leader. A Psalm of David. O LORD, Thou hast searched me, and known me."
—Psalm 139

DOI: 10.1007/s10297-004-0001-5
Originally published in J. Opt. Fiber. Commun. Rep. 1, 312–344 (2004)

Preface

Andrea Galtarossa and Curtis R. Menyuk

This volume had its origin in conversations we had with Professor Carlo Someda of the Università di Padova in summer 2001. All three of us have had a long-standing interest in polarization mode dispersion and, more generally, polarization effects in optical fibers. We were all impressed by the great increase in interest in this subject that had occurred in the past two years. This subject has many theoretical and experimental subtleties, and we were concerned by the propagation of misconceptions in the scientific literature. The idea of a summer school arose, focused on polarization mode dispersion—or PMD—as it is usually known. Perhaps, it would be an appropriate topic for Lake Como, where summer schools are held every year. Carlo heartily endorsed the idea of a summer school, but suggested that we hold it in Venice. Thus, the idea of holding a summer school on PMD in Venice was born.

The next year was a flurry of activity in which we decided what topics should be taught and who should be asked to teach them. We asked many of the most distinguished researchers in the field to be instructors, and, to our great delight, everyone that we asked was able to accept. At the same time, Carlo was arranging for the Istituto Veneto di Scienze, Lettere, ed Arti to host the summer school and was finding support from commercial sponsors. This financial support, for which we are very grateful, along with the hard work of the local organizing committee and in particular by Marco Santagiustina, allowed us to invite the instructors gratis and to greatly reduce the cost to the participants from what we had originally anticipated.

During this year, the telecommunications bubble burst, and we were worried that not enough paying participants would come to the school. In the end, we need not have worried. Almost every possible seat was filled. Indeed, the school made a small profit, which was expended as prizes for graduate students and non-tenure-track research faculty at universities.

The linchpin of the school was two-hour lectures. The first, which was delivered by Bob Jopson and Lynn Nelson, gave an introduction to the subject of PMD. The second, which was delivered by Henning Bülow and Stéphanie Lanne, discussed the important topic of PMD mitigation; Andrea Galtarossa and Anna Pizzinat presented

the model for low-PMD fibers; Nicolas Gisin covered the increasingly important topic of the interaction of PMD with polarization dependent loss. Other topics that were included in the school were: "PMD models," which was covered by Antonio Mecozzi and Mark Shtaif; "Interaction of PMD with nonlinearity and chromatic dispersion," which was covered by Curtis Menyuk; "PMD measurement techniques," which was covered by Paul Williams and by Marco Schiano in two separate lectures; "Spatially resolved measurement of fiber polarization properties," which was covered by Luca Palmieri and Andrea Galtarossa; "PMD impact on optical systems," which was covered by Magnus Karlsson and by Francesco Matera in two separate lectures; "Polarization effects in recirculating loops," which was covered by Brian Marks, Gary Carter, and Yu Sun; "PMD Emulation," which was covered by Alan Willner and Michelle Hauer; and, finally, "Applications of importance sampling to PMD," which was covered by Gino Biondini, Bill Kath, and Sarah Fogal. Dipak Chowdhury worked with Artis and VPI—two producers at that time of commercial software for modeling optical fiber communications systems—to present a lecture that covered numerical modeling of PMD. Additionally, we had lectures on special topics by Hermann Haus, Jim Gordon, Herwig Kogelnik, and Carlo Someda. Finally, we had a poster session, which gave the lecturers the opportunity to learn something from our participants.

The feedback that we received from the participants and the lecturers was overwhelmingly positive. This success was due to the great time and energy that all the instructors put into their lectures. At the summer school and thereafter, we continued to receive the suggestion from many of the lecturers and participants that the summer school lectures would be of interest to a broad audience in the optical fiber communications community.

After asking our lecturers to put so much time and energy into their lecturers, we were a little reluctant to request an additional effort. There was also a concern that the material would have to be updated. However, since the material is largely tutorial in nature, little updating was in fact needed. In the end, all the lecturers provided us with contributions, and the result is the volume that you have before you. We hope that it will be of use to researchers and the students in the field of optical fiber communications who want to have an introduction to PMD. If this volume is as successful as the school, it will be due to the hard work of the contributors.

As a final note, we owe a significant debt to Professor Hermann Haus of MIT. He enthusiastically participated in the summer school, sitting in the front row and posing illuminating questions to the lecturers. It was his suggestion to include poster sessions in the school. He passed away suddenly in May 2003. The loss to our community of one of its most distinguished members is still deeply felt.

Padova, Italy — Andrea Galtarossa
Baltimore, Maryland, USA — Curtis R. Menyuk

DOI: 10.1007/s10297-004-0001-5
Originally published in J. Opt. Fiber. Commun. Rep. 1, 312–344 (2004)

Contents

DOI: 10.1007/s10297-004-0001-5
Originally published in J. Opt. Fiber. Commun. Rep. 1, 312–344 (2004)

Introduction to polarization mode dispersion in optical systems

L. E. Nelson[1] and R. M. Jopson[2]

[1]OFS Laboratories, Somerset, NJ USA
[2]Bell Labs, Lucent Technologies
Holmdel, NJ 07733 USA

Abstract. This introduction covers concepts important to the understanding of polarization mode dispersion (PMD), including optical birefringence, mode coupling in long optical fibers, the Principal States Model, and the time and frequency domain behavior of PMD. Other topics addressed include the concatenation rules, bandwidth of the Principal States, PMD statistics and scaling, PMD system impairments, and PMD outage probability calculations.

1. Introduction

Long-haul, fiber-optic transmission systems can be limited by impairments due to polarization mode dispersion (PMD). PMD is caused by the birefringence of optical fiber and the random variation of its orientation along the fiber length. PMD causes different delays for different polarizations, and when the difference in the delays approaches a significant fraction of the bit period, pulse distortion and system penalties occur. Environmental changes including temperature and stress cause the fiber PMD to vary stochastically in time, making PMD particularly difficult to manage or compensate. In addition, although amplifiers or other components such as add-drop multiplexers in an optical system may have constant birefringence, variable polarization rotations between them due to the environment cause these components to add or subtract randomly to or from the PMD of the total system. While improvements in the PMD of optical fiber and optical components have enabled the transmission of ever higher bit-rates, these improvements have been accompanied by requirements of ever more stringent specifications for system PMD.

In this paper, we will review concepts important to the understanding of PMD and its systems impairments. We aim to provide sufficient background for the following, more detailed chapters.

2. Basics of PMD

2.1. Representation of Polarization

In free space, light is a transverse wave, where the wave motion is perpendicular to the direction of propagation. Here we choose z as the direction of propagation, such that a continuous wave is described by $E(z,t)e^{j(\omega_0 t-\beta z)}$, where ω_0 is the angular carrier frequency, t is the time, and β is the propagation constant. x, y are the transverse coordinates. The term "polarization" is a vector quantity describing, over the course of an optical cycle, the behavior of the electric field, $E(z,t)$, in the x, y plane centered at time t and propagation distance z. When the phase difference between the x and y components, E_x and E_y, is 0 or an integer multiple of π, the light wave is linearly polarized and the plane of oscillation is constant. When the x and y components have equal amplitude and their phase difference is an odd integer multiple of $\pi/2$, the light is circularly polarized. Coherent light not having linear or circular polarization has elliptical polarization.

The Jones formalism [36] provides a concise representation of polarization using the electric field vector itself. We define $\mid s\rangle$ as the 2D Jones (column) ket vector,

$$\mid s\rangle = \begin{pmatrix} s_x \\ s_y \end{pmatrix}, \tag{1}$$

where s_x, s_y are complex quantities. The bra $\langle s \mid$ indicates the corresponding complex conjugate row vector, i.e., $\langle s \mid = (s_x^*, s_y^*)$. The bra-ket notation is used to distinguish Jones vectors from Stokes vectors. Our Jones vectors are all of unit magnitude, i.e., $\langle s \mid s\rangle = s_x^* s_x + s_y^* s_y = 1$, as we assume coherent light except as noted. This concise formalism provides an intuitive understanding of changes in polarization since "lab-space" coordinates are used. It includes optical phase, which is necessary when combining light, but only fully polarized waves can be described by Jones vectors.

The Stokes formalism [75] is an alternative description of polarization and uses four (real) Stokes parameters, which are functions only of observables of the light wave. The polarization state of any light beam (totally, partially, or not polarized) can be described. For coherent light, the Stokes parameters are

$$s_0 = s_x s_x^* + s_y s_y^* \tag{2a}$$
$$s_1 = s_x s_x^* - s_y s_y^* \tag{2b}$$
$$s_2 = s_x s_y^* + s_x^* s_y \tag{2c}$$
$$s_3 = j(s_x s_y^* - s_x^* s_y). \tag{2d}$$

(The broader definition in terms of intensity differences applicable to incoherent light is discussed in section 2.9.) We define $\hat{s} = (s_1, s_2, s_3)$ as a 3D Stokes vector of unit length indicating the polarization of the field and corresponding to $\mid s\rangle$. By this definition, $s_1 = 1$ for linear polarization aligned with the x-axis; $s_2 = 1$ for linear polarization at 45° to this axis; and $s_3 = 1$ for right-circular polarized light ($s_y = j s_x$) conforming to the traditional optics definition. The assignment of either left-circular or right-circular polarization to $s_3 = 1$ is often implicit in PMD literature and has non-trivial consequences in subsequent analyses. We always use the same letter symbols for corresponding Jones and Stokes vectors. Note that a common phase shift of both components of $\mid s\rangle$ does not change $\hat{s}$.

The locus of Stokes vectors, (s_1, s_2, s_3), representing all possible states of polarization of coherent light forms a unit sphere in Stokes space. This is the Poincaré sphere [64], which provides a unique three-dimensional representation of polarization states (see for example, [18, 33]). Linear polarizations are plotted along the equator; circular polarizations are at the poles. Elliptical polarizations are plotted elsewhere on the surface of the Poincaré sphere. (See Fig. 2.) The Stokes formalism, together with the Poincaré sphere, provides an intuitive understanding of the effects on polarized light of birefringence and hence of PMD. It can describe partially polarized light, but it ignores the optical phase of light.

Partially polarized light is important in a method of PMD monitoring that relies on the degree of polarization (DOP), also discussed in section 2.9. One measure of DOP is the length of the normalized 3D Stokes vector. The DOP of coherent light is unity, whereas the DOP of unpolarized light is zero. DOP is a concept having subtleties not always appreciated. It is necessary to average over time or space to obtain a DOP that differs from unity and the measured DOP will depend on the amount of averaging. In single-mode lightwave systems, there is little spatial averaging, and the time average usually lasts many bit periods.

When light passes through an optical element (e.g., optical fiber, where light is approximately a transverse wave), the input polarization $\hat{s}$ will emerge as a new polarization $\hat{t}$ at the output. The transformation of the polarization can be described using a 3×3 Müller matrix R, through $\hat{t} = R\hat{s}$. Similarly, the Jones vector $\mid s\rangle$ will emerge from the same optical element as $\mid t\rangle$, where the 2×2 unitary transmission matrix T relates output to input via $\mid t\rangle = T \mid s\rangle$. The 2×2 Jones matrix U is related to T by $T = e^{-j\phi_0}U$, where $\det U = 1$ determines ϕ_0, the common phase. In the following, we will consistently use this designation of fiber input and output, as shown in Fig. 1.

Fig. 1. Block diagram of optical fiber under test.

2.2. Birefringence

PMD has its origins in optical birefringence. Although telecommunications fibers are often called "single mode," even in an ideal circularly symmetric fiber, there are two orthogonally polarized HE_{11} modes. In a perfect fiber, these modes have the same group delay. In reality, fibers have some amount of asymmetry due to imperfections in the manufacturing process and/or mechanical stress on the fiber after manufacture. The asymmetry breaks the degeneracy of the orthogonally polarized HE_{11} modes, resulting in birefringence: a difference in the phase velocities of the two modes. Even very small amounts of birefringence can cause evolution of the polarization state as light propagates through fiber.

Both intrinsic and extrinsic perturbations cause birefringence in optical fibers. The manufacturing process can set up permanent, intrinsic perturbations in the fiber. Form (geometric) birefringence arises due to a noncircular waveguide, whereas stress birefringence is due to forces set up by a noncircular core. When fiber is spooled,

cabled, or embedded in the ground, extrinsic perturbations, including lateral stress, bending, or twisting, induce birefringence. These perturbations will change as the fiber's external environment changes.

In a short section of fiber, the birefringence can be considered uniform. The difference between the propagation constants of the slow and fast modes can be expressed as

$$\Delta\beta = \frac{\omega n_s}{c} - \frac{\omega n_f}{c} = \frac{\omega \Delta n}{c}, \tag{3}$$

where ω is the angular optical frequency, c is the speed of light, and $\Delta n = n_s - n_f$ is the differential effective refractive index between the slow (s) and fast (f) modes. Generally, the perturbations discussed above create linear birefringence where there are two linearly polarized waveguide modes whose electric field vectors are aligned with the symmetry axes of the fiber. Circular birefringence results from fiber twist [78].

When an input wave that is linearly polarized at 45° to the birefringent axes is launched into a short fiber, the state of polarization evolves in a cyclic fashion as the light propagates down the fiber, i.e., from linear to elliptical to circular and back through elliptical to a linear state orthogonal to the launch state. Analogously, for a fixed-input polarization state, if the light frequency is varied, the output polarization state from a short length of birefringent fiber will cycle in the same way through the various states. This frequency-domain picture of PMD is illustrated in Fig. 2 for a launch state near the birefringent axis. The output polarization traces a circle on the surface of the Poincaré sphere. Several PMD measurement techniques, including Jones Matrix Eigenanalysis [32], Poincaré Sphere Method [1], and Müller Matrix Method [37] use this frequency-domain picture.

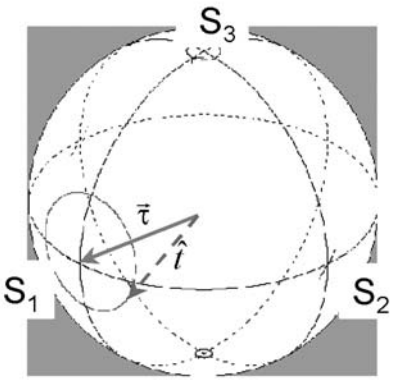

Fig. 2. Frequency-domain behavior of PMD in a short birefringent fiber, where the fiber's birefringent axis $\boldsymbol{\tau}$ is aligned with the S_1 axis.

The differential index, together with the optical wavelength λ, allows us to define a beat length, $L_b = \lambda/\Delta n$, as the propagation distance for which a 2π phase difference accumulates between the two modes or, equivalently, the polarization rotates through a full cycle. Standard telecommunications-type fibers can have beat lengths of ~ 10 m [25], giving $\Delta n \sim 10^{-7}$, which is much smaller than the $\sim 10^{-3}$ index difference between core and cladding.

In the time-domain picture, for a short section of fiber, the differential group delay (DGD), $\Delta\tau$, is defined as the group-delay difference between the slow and fast modes. This $\Delta\tau$ can be found from the frequency derivative of the difference in propagation constants [Eq. (3)]:

$$\frac{\Delta\tau}{L} = \frac{d}{d\omega}\left(\frac{\Delta n\omega}{c}\right) = \frac{\Delta n}{c} + \frac{\omega}{c}\frac{d\Delta n}{d\omega}. \tag{4}$$

This "short-length" or "intrinsic" PMD, $\Delta\tau/L$, is often expressed in units of picoseconds per kilometer of fiber length. The linear length dependence of DGD applies when the birefringence can be considered uniform, as in a short fiber. For the "long-length" PMD regime, the orientation of the fast and slow modes becomes a function of location and the DGD has a square-root-of-length dependence.

Figure 3 shows the time-domain effect of PMD in a short fiber, where a pulse launched with equal power on the two birefringent axes results in two pulses at the output, separated by the DGD, $\Delta\tau$. From Eq. (4) and ignoring the dispersion of Δn, we can then see that the DGD for a single beat length, L_b, is equal to an optical cycle:

$$\Delta\tau_b = L_b\frac{\Delta n}{c} = \frac{\lambda}{c} = \frac{1}{\nu}, \tag{5}$$

which is 5.2 fsec at 1550 nm. The polarization-dependent signal delay method [58] for measuring PMD relies on the time-domain picture of PMD.

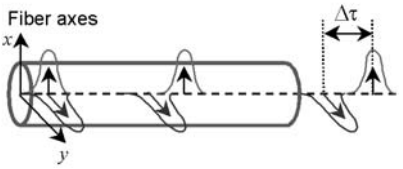

Fig. 3. Time-domain behavior of PMD in a short birefringent fiber.

2.3. Polarization Mode Coupling

While DGD in the short-length regime is predictable because the birefringence is inherently additive, fiber lengths in today's terrestrial and submarine transmission systems are 100s or 1000s of km, and the birefringence is no longer additive. Random variations in the axes of the birefringence occur along the fiber length, causing polarization-mode coupling wherein the fast and slow polarization modes from one segment decompose into both the fast and slow modes of the next segment. Polarization-mode coupling results from localized stress during spooling/cabling/deployment, from splices and components, from variations in the fiber drawing process, and from intentional fiber "spinning" during drawing, which induces mode coupling at "meter" lengths [40]. A commonly used model for long fibers is a concatenation of birefringent sections whose birefringence axes (and magnitudes) change randomly along the fiber.

Owing to mode coupling, the birefringence of each section may either add to or subtract from the total birefringence, and therefore the DGD does not accumulate linearly with fiber length. Instead, it has been shown that in long fiber spans, the DGD

accumulates as a three-dimensional random-walk, and on average increases with the square root of distance [66,71]. Although mode coupling helps to reduce the DGD of a fiber span, because the mode coupling is determined by the fiber's environment, variations in, for example, external stresses will change the mode coupling and thus the fiber's DGD. Therefore, a statistical approach for PMD must be adopted, as discussed in Section 3.

The categorization of a fiber in the short- or long-length regime is determined by a parameter called the correlation length L_c, also referred to as the coupling length [41]. This parameter describes weak random coupling between two waveguides or the equivalent random coupling between the two polarization modes of a fiber with uniform birefringence subject to random perturbations. One considers the evolution of the polarization modes as a function of length in an ensemble of fibers with statistically equivalent perturbations. While the input polarization is fixed, it is equally probable to observe any polarization state at large lengths. The evolution is characterized by the difference $\langle p_x \rangle - \langle p_y \rangle$ of the ensemble averages of the power in the x and y polarizations. Assuming $\langle p_x \rangle = 1$ and $\langle p_y \rangle = 0$ at the input, this difference evolves from a value of 1 at the input to a value of zero at large lengths. L_c is defined as that length where the power difference has decayed to $\langle p_x \rangle - \langle p_y \rangle = 1/e^2$ (for further detail see [71]). Correlation lengths can be less than 1 m when fiber is spooled (due to large amounts of polarization-mode coupling); conversely, L_c can be $\sim$ 1 km when fiber is cabled. The actual fiber behavior is affected by fiber "spinning" during draw, temperature, spool diameter and tension, cable design, installation conditions, and fiber relaxation.

The correlation length thus defines the two different PMD regimes. When the fiber transmission distance L satisfies $L << L_c$, the fiber is in the short-length regime and the DGD increases linearly with distance. When $L >> L_c$, the fiber is considered to be in the long-length regime and the mean DGD, $\overline{\Delta\tau}$, increases with the square root of distance. Transmission systems are generally in the long-length regime, so fiber PMD is often specified using a PMD coefficient having units of ps/(km)$^{1/2}$. While fibers manufactured today can have mean PMD coefficients less than 0.05 ps/(km)$^{1/2}$, "legacy" fibers installed in the 1980s may exhibit PMD coefficients higher than 0.8 ps/(km)$^{1/2}$ [63].

The statistical theory of PMD [21,81] has provided an elegant expression linking the mean square DGD of the fiber to L_b and L_c:

$$\langle \Delta\tau^2 \rangle = 2 \left(\Delta\tau_b \frac{L_c}{L_b} \right)^2 (L/L_c + e^{-L/L_c} - 1). \tag{6}$$

This expression is valid for both regimes and also for the transition region. For $L << L_c$, the above relation simplifies to

$$\sqrt{\langle \Delta\tau^2 \rangle} = \Delta\tau_{\mathrm{rms}} = \Delta\tau_b L/L_b, \tag{7}$$

whereas for $L >> L_c$,

$$\Delta\tau_{\mathrm{rms}} = (\Delta\tau_b/L_b)\sqrt{2LL_c}, \tag{8}$$

reflecting the length dependence discussed earlier. Single-ended backscattering measurement techniques [16] can determine birefringence (L_b) and the mean square DGD. Then the correlation length (L_c) can be inferred from the fundamental Eq. (6) relating the three quantities.

2.4. Principal States Model

The propagation of a pulse through a long length of fiber is very complicated due to random mode coupling and pulse splitting at every change in the local birefringence axes. But a surprising aspect of PMD is that even for long fibers, one can still find two orthogonal-polarization launch states at the fiber input that result in an output pulse that is undistorted to first order. Figure 4 shows the resulting output pulse shapes when a 10-Gb/s, 50% duty-cycle return-to-zero signal was launched with various polarizations through a 48-km fiber with large PMD. The figure shows the output pulse for each of two polarization launches that minimize the bit-error rate (BER) at the fiber output. Note the difference in arrival times, the DGD. Also shown is the output pulse shape for a polarization launch that maximizes the BER at the output. It is apparent that the two launches minimizing the BER result in fairly undistorted pulses, while the pulse from the third launch is significantly broadened. The two least distorted pulses are the fastest and the slowest pulses of all the polarizations launched.

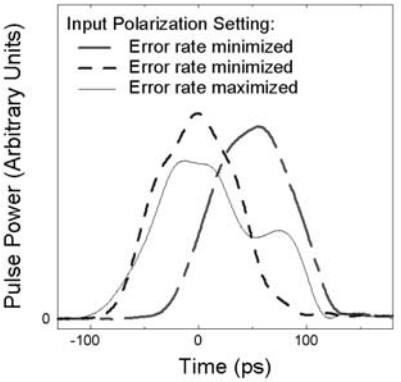

Fig. 4. Output pulse shapes for three polarization launches of a 10Gb/s, 50% duty cycle signal through a 48-km fiber with ∼60 ps PMD.

The Principal States Model, originally developed by Poole and Wagner [65] was the first to describe this phenomenon and is still in common use today for the characterization of PMD. The model provides both a time domain and a frequency domain characterization of PMD. Figure 4 illustrates the time-domain picture. The frequency-domain picture allows a very simple definition. It states that, for a length of fiber, there exists for every frequency a special pair of polarization states, called the Principal States of Polarization (PSPs). A PSP is defined as that input polarization for which the output state of polarization is independent of frequency to first order, i.e., over a small frequency range. The Principal States Model applies when the bandwidth of the pulses is small, with a pulse length greater than the PMD-induced DGDs. This is because PMD is intrinsically an interference phenomenon, caused by the coherent addition of the complex amplitudes of the multiplicity of pulses created by the repeated pulse splitting mentioned above.

In the absence of polarization-dependent loss, the PSPs are orthogonal. For each pair of input PSPs, there is a corresponding pair of orthogonal PSPs at the fiber output. The input and output PSPs, designated by the unit Stokes vectors $\hat{p}_s$ and $\hat{p}$, respectively,

are related by the fiber's transmission matrix R, via $\hat{p} = R\hat{p}_s$, just as any input polarization is related to a polarization at the fiber output.

2.5. PMD Vector

Using the Principal States Model, PMD can be characterized by the PMD vector:

$$\boldsymbol{\tau} = \Delta\tau\hat{p}, \tag{9}$$

a vector in three-dimensional Stokes space, where the magnitude, $\Delta\tau$, is the DGD. The unit vector $\hat{p}$ points in the direction of the slower PSP, whereas the vector $-\hat{p}$ indicates the orthogonal, faster PSP. In Stokes space, $-\hat{p}$ is 180° from $\hat{p}$. Note that the definition used here is in right-circular Stokes space (with S_3 denoting right-circular polarization), whereas the original PMD vector $\boldsymbol{\Omega}$ of Poole et al. [67] was defined in left-circular Stokes space. See [48] for further explanation of the relation between the two.

The PMD vector at the fiber input $\boldsymbol{\tau}_s$ is related to the output PMD vector $\boldsymbol{\tau}$ by $\boldsymbol{\tau} = R\boldsymbol{\tau}_s$. One can then show that the frequency derivative of $\hat{t} = R\hat{s}$ leads directly to the law of infinitesimal rotation

$$\hat{t}_\omega = \frac{d\hat{t}}{d\omega} = \boldsymbol{\tau} \times \hat{t}, \tag{10}$$

where $\boldsymbol{\tau}\times = R_\omega R^T$ and R^T is the transpose of R. Here, the PMD vector describes how, for a fixed input polarization, the output polarization $\hat{t}$ will precess around $\boldsymbol{\tau}$ as the frequency is changed. The direction of $\hat{t}$ relative to $\boldsymbol{\tau}$ determines the angle of precession, whereas the magnitude, $\Delta\tau$, determines the rate at which $\hat{t}$ precesses around $\boldsymbol{\tau}$. For example, if $\hat{s}$ is launched with equal power along the PSPs, $\boldsymbol{\tau} \times \hat{t}$ will have its largest value, and the largest change in the output polarization will occur for a frequency change $\Delta\omega$. The precession has magnitude $\phi = \Delta\tau\Delta\omega$, where ϕ is the rotation angle on the Poincaré sphere. If $\hat{t}$ is aligned with $\pm\boldsymbol{\tau}$, then there is no precession and no change of the output polarization with frequency. This is, of course, the postulate for a PSP. The rotation law, Eq. (10) thus provides a precise mathematical definition of the PSP and of its length, $\Delta\tau$, the DGD. The law also says that there are only two PSPs, corresponding to the two alignments, $\hat{t} = \pm\boldsymbol{\tau}$.

A length of polarization-maintaining fiber (PMF) has a constant PMD vector whose length, the DGD, and direction $\hat{p}$ do not change with frequency. For this simple case, the output vector $\hat{t}$ will trace a circle on the Poincaré sphere as the frequency is varied, as illustrated in Fig. 2. In real fibers, however, both the magnitude and the direction of $\boldsymbol{\tau}$ change with frequency, as shown in Fig. 5.

The rotation law still applies locally in this case, describing $\hat{t}(\omega)$ as a circular arc for a small range of frequencies. In this range, characterized by first-order PMD, the behavior of the real fiber resembles that of PMF. The DGD at an instant in time for such a small frequency range is often called the "instantaneous DGD" to distinguish it from a mean DGD obtained by averaging over time or frequency. The longer-range motion of $\hat{t}(\omega)$ around $\boldsymbol{\tau}(\omega)$ is more complicated, reflecting higher-order PMD.

Whereas the frequency domain provides a continuous wave, single-frequency view of PMD, the time domain involves pulses. This allows an alternative physical interpretation of the DGD parameter, $\Delta\tau$, to the speed of precession identified previously. The

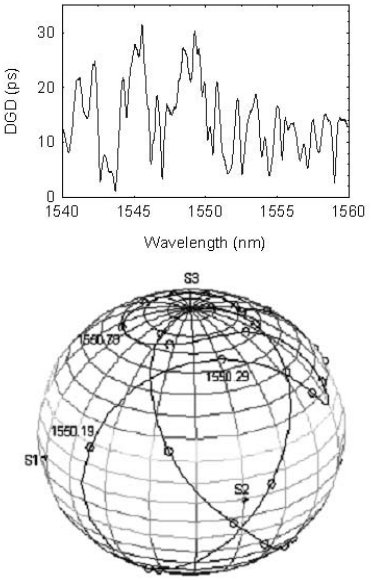

Fig. 5. PMD measurement of a 14.7ps mean DGD fiber. Top: Magnitude of $\boldsymbol{\tau}$. Bottom: Direction of $\boldsymbol{\tau}$ (markers indicate 0.1-nm intervals).

time-domain view uses laboratory coordinates, Jones vectors to characterize polarization, and a 2×2 unitary complex transmission matrix T to relate the input and output Jones vectors, via $\mid t\rangle = T \mid s\rangle$. In this framework, the PSPs are characterized by the unit Jones vectors, $\mid p\rangle$ and $\mid p_-\rangle$, corresponding to the Stokes vectors, $\hat{p}$ and $-\hat{p}$, discussed previously. Using the PSPs as an orthogonal basis set, any input or output polarization can be expressed as the vector sum of two components, each aligned with a PSP. Within the realm of first-order PMD, the output electric field from a fiber with PMD has the form

$$\mathbf{E}_{\text{out}}(t) = a \mid p\rangle E_{\text{in}}(t - \tau_0 - \Delta\tau/2) + b \mid p_-\rangle E_{\text{in}}(t - \tau_0 + \Delta\tau/2), \qquad (11)$$

where $E_{\text{in}}(t)$ is the input electric field, a and b are the complex weighting coefficients indicating the field amplitude launched along the slow and fast PSPs, $\mid p\rangle$ and $\mid p_-\rangle$, and τ_0 is the polarization-independent transmission delay. In this formulation, $\Delta\tau$ is identified as the difference in arrival times between the two principal states, explaining its designation as the DGD. It is usually stated in picoseconds (ps). It is apparent from Eq. (11) that PMD can cause pulse broadening due to the DGD, and that there is no pulse broadening when the input is aligned with a PSP, i.e., when a or b is zero. Note that the simple PMD Stokes vector, $\boldsymbol{\tau}$, does not have a vector analog in the laboratory frame where Eq. (11) separates the DGD from the PSP polarizations.

The precise definition of PSPs and DGD in the time domain can be obtained from an expression for polarization-dependent group delay τ_g experienced by a signal launched with polarization $\hat{s}$:

$$\tau_g = \tau_0 + \frac{1}{2}\boldsymbol{\tau} \cdot \hat{s}, \qquad (12)$$

where τ_0 is the polarization-independent delay of the fiber (whose wavelength dependence leads to chromatic dispersion). This simple, yet powerful relation was originally derived and further substantiated by analysis based on first moments of the transmitted pulses [28, 42, 53, 73]. This time-domain definition identifies the PSPs as those input polarizations ($\hat{s} = \pm\hat{p}$) that maximize or minimize the signal delay ($\tau_g = \tau_0 \pm \Delta\tau/2$). This definition is equivalent to the earlier frequency-domain definition and is the basis of the polarization-dependent signal delay method for PMD measurement.

The traditional Jones Matrix Eigenanalysis (JME) measurement approach provides yet another equivalent definition bridging the time and frequency domain. Haus [31] observes that the Hermitian appearing in the JME is connected to the energy of the light stored in the fiber. The light in the slow PSP spends more time in the fiber and maximizes the stored energy, whereas transmission along the fast PSP minimizes the stored energy.

The evolution of the PMD vector with fiber length is described by the dynamical equation for PMD [70],

$$\frac{d\boldsymbol{\tau}}{dz} = \frac{d\boldsymbol{\beta}}{d\omega} + \boldsymbol{\beta} \times \boldsymbol{\tau}, \tag{13}$$

relating the PMD vector to the microscopic birefringence. Here z is the position along the fiber. $\boldsymbol{\beta}$ is the three-dimensional, local birefringence vector of the fiber [19] pointing in the direction of the birefringence axis and having magnitude $\Delta\beta$ proportional to Δn (see [28]). This equation is the basis for the statistical theory of PMD [21].

2.6. Second-order PMD

Because the fiber PMD vector varies with optical angular frequency ω, a Taylor-series expansion of $\boldsymbol{\tau}(\omega)$ with $\Delta\omega$ about the carrier frequency ω_0 has typically been used for larger signal bandwidths [8, 21, 27],

$$\boldsymbol{\tau}(\omega_0 + \Delta\omega) = \boldsymbol{\tau}(\omega_0) + \boldsymbol{\tau}_\omega(\omega_0)\Delta\omega + \cdots . \tag{14}$$

So-called second-order PMD is then described by the derivative,

$$\boldsymbol{\tau}_\omega = \frac{d\boldsymbol{\tau}}{d\omega} = \Delta\tau_\omega \hat{p} + \Delta\tau \hat{p}_\omega, \tag{15}$$

where the subscript ω indicates differentiation. Second-order PMD thus has two terms. Since $\hat{p}_\omega$, which is not a unit vector, is perpendicular to $\hat{p}$ (i.e., $\hat{p} \cdot \hat{p}_\omega = 0$), the first term on the right-hand side of Eq. (15) is $\boldsymbol{\tau}_{\omega\parallel}$, the component of $\boldsymbol{\tau}_\omega$ that is parallel to $\boldsymbol{\tau}$, whereas the second term, $\boldsymbol{\tau}_{\omega\perp}$, is the component of $\boldsymbol{\tau}_\omega$ that is perpendicular to $\boldsymbol{\tau}$. Figure 6 shows a vector diagram of the principal parameters and their interrelationships.

The magnitude of the first term, $\Delta\tau_\omega$, is the change of the DGD with wavelength and causes polarization-dependent chromatic dispersion (PCD) [22,68], resulting in polarization-dependent pulse compression and broadening. It can be viewed as a polarization-dependent change in the chromatic dispersion, DL, of the fiber, described by an effective dispersion,

$$(DL)_{\text{eff}} = DL \pm \tau_\lambda. \tag{16}$$

In accordance with the customary dispersion measure, DL, the PCD is defined as,

$$\tau_\lambda = -(\pi c/\lambda^2)\Delta\tau_\omega = \frac{1}{2}\frac{d\Delta\tau}{d\lambda}, \tag{17}$$

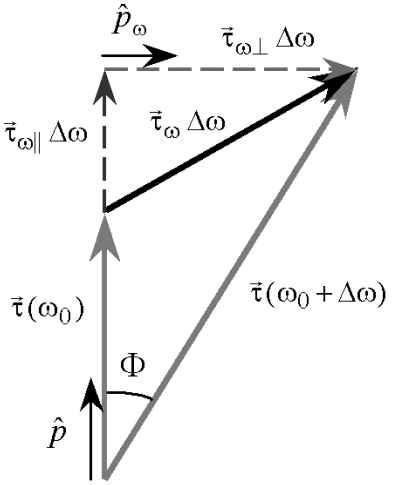

Fig. 6. Schematic diagram of the PMD vector $\boldsymbol{\tau}(\omega)$ and the second-order PMD components showing the change of $\boldsymbol{\tau}(\omega)$ with frequency. Note that $\hat{p}_\omega$ is perpendicular to $\hat{p}$. The angular rotation rate, $d\Phi/d\omega$, of the PMD vector $\boldsymbol{\tau}(\omega)$ with ω is described by $\mid \hat{p}_\omega \mid$.

where c is the velocity of light, λ is the wavelength, and τ_λ is usually expressed in ps/nm. The PCD is proportional to the wavelength derivative of the DGD spectrum. The plus and minus signs in Eq. (16) correspond to alignment with the two PSPs. Note that the magnitudes of $\boldsymbol{\tau}_{\omega\parallel}$ and $\Delta\tau_\omega$ are equal. Figure 7(a) shows the PCD for data depicted in Fig. 5. The DGD data were numerically differentiated to obtain the PCD. It is apparent that PCD causes the effective dispersion to fluctuate rapidly with wavelength.

The second term in Eq. (15), $\Delta\tau\hat{p}_\omega$, describes PSP depolarization, a rotation of the PSPs with frequency. As shown in Fig. 6, the angular rate of rotation, $d\Phi/d\omega = \mid \hat{p}_\omega \mid$, of the PMD vector $\boldsymbol{\tau}(\omega)$ is measured by the magnitude $\mid \hat{p}_\omega \mid$, which we express in ps. Note that $d\Phi/d\nu$[mrad/GHz] $= 2\pi \mid \hat{p}_\omega \mid$ [ps], where ν is the optical carrier frequency and $\omega = 2\pi\nu$. We have already seen in Fig. 5 the rapid motion of $\hat{p}$ for the 14.7-ps mean DGD fiber. Figure 7(b) is a plot of $\mid \hat{p}_\omega \mid$ for this same fiber and wavelength range. Pulse distortions caused by depolarization include overshoots and generation of satellite pulses. PSP depolarization can also have a detrimental effect on first-order PMD compensators.

The input and output second-order PMD vectors ($\boldsymbol{\tau}_{s\omega}$ and $\boldsymbol{\tau}_\omega$, respectively) transform the same way as the first-order PMD vector, so that

$$\boldsymbol{\tau}_\omega = R\boldsymbol{\tau}_{s\omega}, \tag{18}$$

where, again, R is the Müller rotation matrix [28]. For the third-order PMD vectors, one can show that

$$\boldsymbol{\tau}_{\omega\omega} = R\boldsymbol{\tau}_{s\omega\omega} + \boldsymbol{\tau} \times \boldsymbol{\tau}_\omega. \tag{19}$$

The statistical theory of second-order PMD [21,22] has provided probability density functions for the various second-order components that have been experimentally confirmed [23, 39], as has their scaling with mean DGD [57]. These results will be outlined

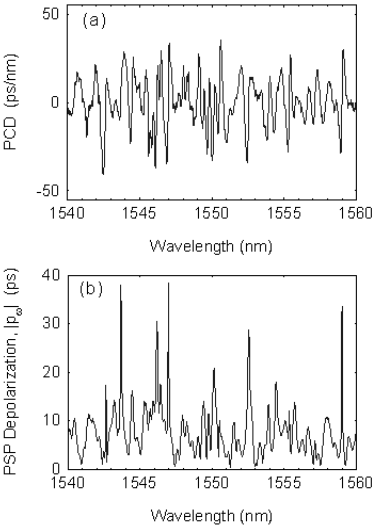

Fig. 7. (a) Plot of the polarization-dependent chromatic dispersion and (b) plot of PSP depolarization, for the same 14.7-ps mean DGD fiber as in Fig. 5.

in Section 3. Higher-order PMD has also been described using other formulations [5, 20, 73] rather than the Taylor-series expansion described above. These formulations attempt to better describe how the PMD vector changes with optical frequency. However, the statistics have not been completely derived for these approaches.

2.7. The Bandwidth of the Principal States

The bandwidth of the principal state provides guidance on the change of the PMD vector $\boldsymbol{\tau}(\omega)$ of the fiber with frequency (or wavelength). It is the bandwidth, $\Delta\omega_{\mathrm{PSP}} = 2\pi\Delta\nu_{\mathrm{PSP}}$, or the corresponding wavelength range, $\Delta\lambda_{\mathrm{PSP}}$, over which the PMD vector is reasonably constant. This concept is important for frequency-domain measurement of PMD vectors, where measurements of polarization rotations at two or more frequencies are required. These frequencies have to be confined to the range $\Delta\lambda_{\mathrm{PSP}}$ in order to reduce inaccuracy caused by higher-order PMD. On the other hand, in statistical PMD measurements, the samples of $\boldsymbol{\tau}(\lambda)$ are deemed to be statistically independent if their wavelengths are at least $6\Delta\lambda_{\mathrm{PSP}}$ apart.

While different constants have been reported [3, 5], studies of the accuracy of measurements of PMD provide a good practical estimate for $\Delta\omega_{\mathrm{PSP}}$ given by the relation [37]

$$\Delta\omega_{\mathrm{PSP}} \cdot \overline{\Delta\tau} = \pi/4, \tag{20}$$

where $\overline{\Delta\tau}$ is the mean DGD of the fiber. This implies a frequency band $\Delta\nu_{\mathrm{PSP}} = 1/(8\overline{\Delta\tau})$, or

$$\Delta\nu_{\mathrm{PSP}} = 125\ \mathrm{GHz}/\overline{\Delta\tau}, \tag{21}$$

when $\overline{\Delta\tau}$ is expressed in ps. For wavelengths near 1550 nm, the corresponding wavelength range $\Delta\lambda = \Delta\nu \cdot \lambda^2/c$ can be written in the simple form $\Delta\lambda_{\mathrm{PSP}} = 1\mathrm{nm}/\overline{\Delta\tau}$.

Figure 8 shows measured DGD data for a fiber with a mean DGD $\overline{\Delta\tau} = 40$ ps, shown over a range of 2 nm with 0.1 nm markers. Here, the values of $\Delta\tau$ appear reasonably constant over the calculated $\Delta\lambda_{\mathrm{PSP}}$ of 0.025 nm. Note that the bandwidth of the PSP is consistent with the concept of second-order PMD, $\boldsymbol{\tau}_{\omega}$, as explained in [48].

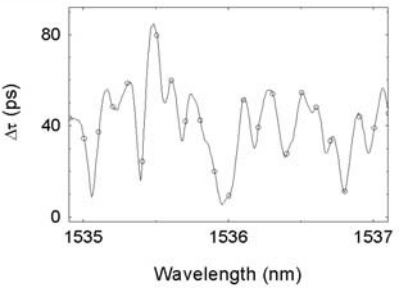

Fig. 8. DGD measurement of a fiber with mean DGD of 40 ps. The measured values of $\Delta\tau$ appear reasonably constant over the bandwidth of the PSP, 0.025-nm. Markers indicate 0.1-nm intervals.

The correlation function of the PMD vectors $\boldsymbol{\tau}(\omega_0)$ and $\boldsymbol{\tau}(\omega_0 + \Delta\omega)$ recently reported by Karlsson and Brentel [43] and Shtaif et al. [74] (and discussed further in Section 3) provides an elegant confirmation and interpretation of the $\Delta\omega_{\mathrm{PSP}}$ concept and its practical implications. Figure 9 shows a normalized plot of this correlation as a function of the frequency separation, $\Delta\nu = \Delta\omega/2\pi$, for a mean DGD of $\overline{\Delta\tau} = 1.25$ ps. For this value, the bandwidth of the PSP is $\Delta\nu_{\mathrm{PSP}} = 100$ GHz. At this frequency spacing the correlation is seen to drop from 1 to 0.89, supporting the idea that $\boldsymbol{\tau}$ is essentially constant over the 100-GHz width. At $6\Delta\nu_{\mathrm{PSP}} = 600$ GHz, the correlation drops to 0.11, indicating that PMD vectors at that frequency spacing are essentially uncorrelated.

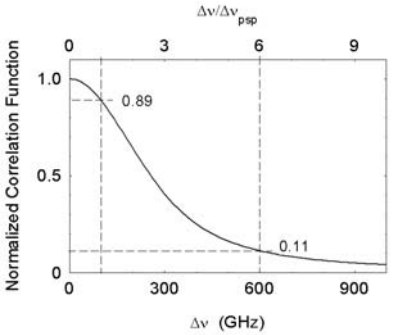

Fig. 9. Plot of the (normalized) correlation function, $\langle \boldsymbol{\tau}(\nu_0) \cdot \boldsymbol{\tau}(\nu_0 + \Delta\nu) \rangle / \langle \Delta\tau^2 \rangle$, as a function of the frequency separation, $\Delta\nu$, for a mean DGD of $\overline{\Delta\tau} = 1.25$ ps and bandwidth of the PSP, $\Delta\nu_{\mathrm{PSP}} = 100$ GHz. The top x axis shows the normalized frequency separation, $\Delta\nu/\Delta\nu_{\mathrm{PSP}}$, allowing general use of the plot.

2.8. Concatenation Rules

The total PMD vector of a series of two or more elements with known PMD vectors can be determined using the simple, but powerful concatenation rules [17, 21, 26, 28, 53, 70]. The concatenation rules have been used in the analysis of how the PMD vector grows with fiber length [17] and for statistical PMD modeling [21]. They are also useful for PMD simulation and in the design of multi-section PMD compensators. Although the concatenation rules have appeared in sum, differential, and integral formulations for both first- and second-order PMD vectors, this section will concentrate on the sum rules. See [28] for the other formulations.

The concatenation rule for first-order PMD is similar to that for transmission-line impedances: To obtain the PMD vector of an assembly, transform the PMD vectors of each individual section to a common reference point and take the vector sum (in three-dimensional Stokes space). This vector can then be transformed to any other location in the system using the known rotation matrices of the different sections. For example, for the two sections shown in Fig.10, the PMD vector, $\boldsymbol{\tau}_m$, at the midpoint is

$$\boldsymbol{\tau}_m = \boldsymbol{\tau}_1 + \boldsymbol{\tau}_{s2} = \boldsymbol{\tau}_1 + R_2^T \boldsymbol{\tau}_2, \tag{22}$$

where all $\boldsymbol{\tau}_i$ and R_i are functions of frequency. To find the total PMD vector at the output, $\boldsymbol{\tau}$, we must then transform it by R_2, so that

$$\boldsymbol{\tau} = R_2 \boldsymbol{\tau}_1 + \boldsymbol{\tau}_2, \tag{23}$$

since $R_2 R_2^T \boldsymbol{\tau}_2 = \boldsymbol{\tau}_2$. The corresponding PMD vector diagram is also shown in Fig. 10. It provides a simple geometrical interpretation of the concatenation.

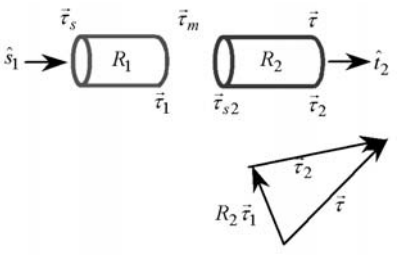

Fig. 10. Diagram of the concatenation of PMD vectors for two sections. The PMD vector $\boldsymbol{\tau}$ at the output is equal to the sum of $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$, after transforming by R_2.

Equation (23) is the basic concatenation rule. We can use it similarly to find the total PMD vector at the input, $\boldsymbol{\tau}_s$, by the transformation

$$\boldsymbol{\tau}_s = R_1^T \boldsymbol{\tau}_m = R_1^T (\boldsymbol{\tau}_1 + R_2^T \boldsymbol{\tau}_2). \tag{24}$$

The rule can be generalized to multiple sections as well as to differentially small sections. Second-order PMD can also be concatenated. By differentiating Eq. (23) and making the proper substitutions, one can show that

$$\boldsymbol{\tau}_\omega = \boldsymbol{\tau}_2 \times \boldsymbol{\tau} + R_2 \boldsymbol{\tau}_{1\omega} + \boldsymbol{\tau}_{2\omega}. \tag{25}$$

The first- and second-order PMD vectors for many sections can be determined by repeated application of the two section rules in Eqs. (23) and (25). For the fiber in Fig. 11 consisting of m sections, each with known rotation matrix R_n and output PMD vector $\boldsymbol{\tau}_n$, the sum rules of the assembly are for first-order PMD

$$\boldsymbol{\tau} = \sum_{n=1}^{m} R(m, n+1)\boldsymbol{\tau}_n, \tag{26}$$

and for second-order PMD

$$\boldsymbol{\tau}_\omega = \sum_{n=1}^{m} R(m, n+1)\{\boldsymbol{\tau}_{n\omega} + \boldsymbol{\tau}_n \times \boldsymbol{\tau}(n)\}, \tag{27}$$

where we define the rotation matrix of the last $m - n + 1$ sections as $R(m, n) = R_m R_{m-1} \cdots R_n$, where $R(m, m) = R_m$ and $R(m, m + 1)$ is the identity matrix. The differential concatenation rule for PMD shows how $\boldsymbol{\tau}(z)$ changes due to the differential addition of length Δz [28] and is equivalent to Eq. (13), the dynamical PMD equation.

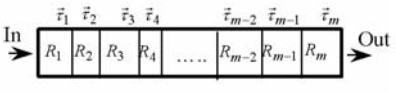

Fig. 11. Concatenation of m sections of PMD, each with known rotation matrix R_n and output PMD vector $\boldsymbol{\tau}_n$. [See Eqs. (26) and (27).]

2.9. Degree of Polarization

The concept of the degree of polarization (DOP) characterizes the average polarization state of light over a broad spectral range. For time-dependent signals it is also defined as an average over a specified time period. The definition of the DOP is based on the Stokes parameters routinely measured by a polarimeter-based instrument (containing polarizers and four photodiodes) or DOP monitor.

A coherent sinusoidal optical carrier (or a narrow spectral component) has a well-defined polarization, described by a unit Stokes vector denoted by a lowercase $\hat{s}$. The spectrally averaged Stokes parameters, S_i, are denoted by capital letters, where S_0 is the total intensity of the light. The other three parameters are the differences between the measured intensities of pairs of orthogonal polarizations, where S_1 refers to the vertical/horizontal polarizations, S_2 to the $\pm 45^\circ$ polarizations, and S_3 to right/left circular polarizations. The definition of the DOP is

$$\text{DOP} = \sqrt{S_1^2 + S_2^2 + S_3^2}/S_0. \tag{28}$$

A narrow spectral component then has DOP $= 1$.

If the filter of the monitor passes several WDM channels (or spectral components), the measured intensity is the sum of all channel intensities for any given polarizer setting. The monitor, therefore, measures Stokes parameters that are the sum of the

Stokes parameters of the individual WDM channels. For the illustrative example of two channels with equal power, one finds

$$\text{DOP} = \sqrt{(1 + \hat{s}_1 \cdot \hat{s}_2)/2} = \cos(\phi/2), \tag{29}$$

where $\hat{s}_1$ and $\hat{s}_2$ are the unit Stokes vectors of the two channels and ϕ is the Stokes angle between the two channels. If the channels have parallel polarizations, the DOP $= 1$; for antiparallel polarizations in Stokes space (orthogonal polarizations in the lab), the DOP $= 0$.

When two channels (or spectral components) having a channel (or spectral) spacing of $\Delta\omega$ are launched with identical polarizations into a fiber with DGD, $\Delta\tau$, the fiber PMD rotates $\hat{s}_2$ relative to $\hat{s}_1$. This reduces the measured DOP at the output. The law of infinitesimal rotation [Eq. (10)] characterizes this relative rotation and leads to the approximate expression

$$\text{DOP} = 1 - \frac{1}{8}(\Delta\omega\Delta\tau \sin\theta)^2, \tag{30}$$

valid for small $\Delta\omega\Delta\tau$. Here θ is the angle between the launched Stokes vector and the PSP of the fiber.

Owing to its dependence on the DGD, $\Delta\tau$, the DOP has been successfully used as a monitor for PMD compensators [72]. Advantages are that the DOP is bit-rate independent and that it can provide high speed, e.g., kHz response time, without requiring electronics operating at frequencies as high as the bit rate.

3. Statistics and Scaling

3.1. Statistics of PMD

PMD differs from many lightwave system properties in that the impairment caused by PMD usually changes stochastically with wavelength and time. This means that one cannot predict the impairment of the system at any particular wavelength and time, but must instead resort to a statistical description. It also means that, in general, one cannot characterize the PMD of a system to arbitrary accuracy, since a measurement will sample only a portion of the statistical ensemble. Most other impairments, such as those due to chromatic dispersion, system attenuation, or self-phase modulation, will have some uncertainty caused by manufacturing tolerances, imprecise wavelength control, aging, and the like. However, in most cases, the probability densities for these impairments will vanish beyond some value, allowing a system to be designed to tolerate the worst-case impairment. Tolerance of worst-case impairment is rarely possible when PMD is a significant source of impairment. The probability densities for PMD in most situations have asymptotic tails extending to unacceptably large impairment. Hence, the system cannot be designed to handle the worst-case PMD impairment and must instead be designed for a specified outage probability. For similar reasons, the goal of PMD compensation cannot be to eliminate the impairment, but rather to reduce the PMD outage probability. To understand and predict system outage probabilities, to design PMD compensators, and to accurately measure PMD in systems, one must understand the statistics of the phenomena associated with PMD.

Interesting statistical characteristics of PMD include the probability densities of first- and higher-order PMD, the scaling of PMD phenomena with changes in the mean DGD, various correlation functions, and characteristics associated with the accuracy of PMD measurements. The first statistical property of PMD to attract interest was the mean DGD [65,1].

Since the DGD usually changes quite slowly with time, it is difficult to obtain an accurate estimate of the mean DGD by repeatedly measuring the DGD at one wavelength. Instead, the DGD is averaged over wavelength and the result is assumed to be equal to the time average. This assumption is of crucial importance. Virtually every measured mean DGD value quoted in the literature was obtained by wavelength averaging. Although some data has been taken covering a broad span of time and wavelength [45, 65], the issue has not been fully explored. Instead, this equality, upon which the accuracy of much experimental PMD work relies, is assumed.

Once the validity of wavelength scanning is assumed, experimental attention can be focused on probability densities. Figure 12 shows the probability density of the DGD for a system having a mean DGD of 1 ps. As shown below, this density and all PMD densities can be easily scaled for a different mean DGD. Measurements, together with simulation results, can be used to establish the veracity of analytic derivations of probability densities and the validity of the PMD models. The densities are then combined with an understanding of PMD-induced system penalties (see Section 4) to predict system outage probability. Powerful scaling laws emerge from the densities and their derivation. With these laws, one can obtain extensive understanding of the statistical properties of the PMD in a particular system through the measurement of a single system characteristic, the mean DGD.

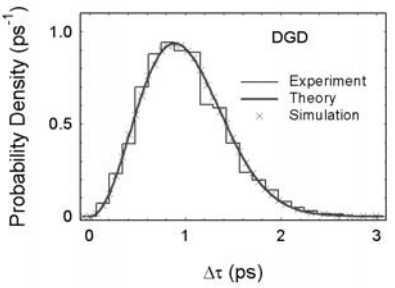

Fig. 12. Maxwellian probability density of DGD for a fiber having 1-ps mean DGD.

3.2. Probability Densities

The probability density of the DGD was the first PMD density to attract attention [17, 70]. For fully random PMD, the problem can be reduced to that of the probability density of the magnitude of the sum of 3-dimensional vectors having random orientation and length. The result is the well-known Maxwellian distribution shown in Fig. 12. The simulation used a quarter million instances of 1200 randomly oriented linearly birefringent plates. The measured data was scaled from a 120-nm scan of a fiber having a mean DGD of 14.7 ps. Using the bandwidth of the principal states discussed above

in section 2.7, we expect the 6000-point scan to contain only about 300 statistically independent measurements.

Replotting with a logarithmic vertical scale yields Fig. 13, which starts to reveal the range of density that is more important for the estimation of outage probabilities. The deviations of the simulation values for small probabilities are caused by the small number of simulations achieving the higher values of DGD. The experimental values demonstrate what appears to be a systematic deviation from the expected curve for the high values of DGD. Since only two statistically independent points are involved in the deviation, it is likely a statistical fluctuation. The smoothness of the deviation arises from the data being 20-fold over sampled.

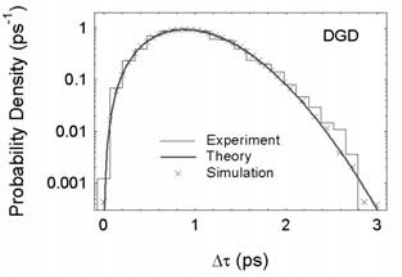

Fig. 13. Maxwellian probability density of DGD for a fiber having mean DGD of 1 ps.

Figure 14 shows the density for the second-order quantity, $\hat{p}_\omega$, which characterizes the depolarization of the PSPs. (As mentioned below, $\hat{p}_\omega$ scales as a first-order quantity.) To obtain the measured data for Fig. 14, the Müller Matrix Method was first used to obtain the spectrum of $\boldsymbol{\tau}$ used for Figs. 12 and 13. From this, $\boldsymbol{\tau}_\omega$ was obtained by numerical differentiation and $\hat{p}_\omega$ obtained by vector algebra. Figures 13 and 14 illustrate the difficulty of verifying densities through measurement or simulation. The interesting regions of the densities are the asymptotic tails containing an integrated probability of 10^{-4} to 10^{-7}. The 300 statistically independent measurement points do not illuminate this region at all and the 250,000-point simulation barely reaches it. One approach to the problem of simulating the important, but improbable regions of the densities is importance sampling, a detailed description of which appears elsewhere [4] and in this tome.

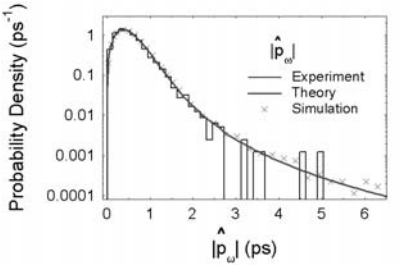

Fig. 14. Probability density for the depolarization of the principal states of polarization for a fiber with mean DGD of 1 ps.

Analytic expressions for probability densities have been obtained for first- and second-order PMD [14, 17, 21, 22, 23, 39, 57, 70]. In particular, densities for both the magnitude and an individual component of first-order PMD are known. For second-order PMD, they are known for the magnitude and a component of the vector as well as the component of the second-order vector that is aligned with the first order vector (the PCD component). The densities are known in integral form for the component of the second-order vector that is perpendicular to the first order vector and for this component normalized by the DGD (the PSP depolarization). Once the densities are known, they can be used to obtain the mean and mean square (or second moment or variance) of each of the variables of interest.

Appendix A summarizes the results. For simplicity, τ rather than $\overline{\Delta\tau}$ is used in the appendix to represent the mean DGD. The estimation of system outage probabilities often requires complementary cumulative probability distributions, which are

$$CC(y) = \int_y^\infty dx\, p_X(x),$$

where $p_X(x)$ could be one of the densities in Appendix A. These integrals provide the probability of exceeding some value, y, which is typically the threshold for unacceptable impairment. Figure 15 shows the complementary cumulative probability distribution together with the cumulative probability distribution for the DGD, once again using a mean DGD of 1 ps. The markers show results from the simulation described previously. The deviation of the simulation from theoretical predictions for $\Delta\tau > 3.1$ ps is likely a result of the small number of realizations (five) that generated a DGD larger than 3.1 ps.

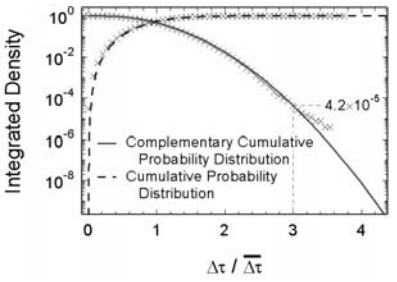

Fig. 15. Integrals of the Maxwellian probability density of DGD for a mean DGD of 1 ps.

4. System Effects of PMD

Fiber PMD can cause a variety of impairments in optical fiber transmission systems. For a single digital channel, intersymbol interference (ISI) impairment is caused by the differential group delay, $\Delta\tau$, between the two pulses propagating in the fiber when the input polarization, $\hat{s}$, of the signal does not match the PSP of the fiber, $\hat{p}$. In this first-order PMD effect the fractional powers launched into the PSPs are $\gamma = \langle s \mid p\rangle\langle p \mid s\rangle = (1 + \hat{p}\cdot\hat{s})/2$ and $(1-\gamma) = (1 - \hat{p}\cdot\hat{s})/2$. For larger signal bandwidths system impairments can occur due to second- and higher-order PMD,

particularly when these PMD components combine with chromatic fiber dispersion or signal chirp.

Power penalties and system outage in digital systems due to first-order PMD are described in the succeeding sections. We also mention the effects of PMD on multichannel systems. PMD can cause severe impairments in analog systems; the reader is referred to [71] and to [14] for information on this topic. While the combined effects of PMD and polarization dependent loss is now an active area of investigation [30, 34, 82], the following sections assume absence of PDL.

4.1. Power Penalties due to First-order PMD

In the first-order picture, PMD splits the input signal entering the fiber into two orthogonally polarized components that are delayed by $\Delta\tau$ relative to each other during transmission. The impairment caused by this effect can be expressed as a power penalty ϵ of the form [69]

$$\epsilon(\text{dB}) = (A/T^2)\Delta\tau^2\gamma(1-\gamma) = A(\Delta\tau/2T)^2 \sin^2\theta, \tag{31}$$

where the penalty, expressed in dB, is assumed to be small. Here, T is the bit interval, $0 \leq \gamma \leq 1$ is the power-splitting ratio, and θ is the angle between the input polarization, $\hat{s}$, and the input PSP, $\hat{p}$. The DGD value, $\Delta\tau$, appearing in Eq. (31) is the "instantaneous" DGD value, assumed to be constant during the penalty measurement. The $\gamma(1-\gamma)$ dependence shown in this expression has been verified by experiment in [47]. The dimensionless A-parameter depends on pulse shape, modulation format, and specific receiver characteristics such as the detailed response of the electrical filter and whether optical or thermal noise predominates. For pin receivers the reported values for A range from 10 to 40 for non-return-to-zero (NRZ) modulation format and from 20 to 40 for return-to-zero (RZ) modulation format, whereas the A ranges for optically preamplified receivers are 10 to 70 for NRZ and 10 to 40 for RZ.

Jopson et al. [38] report the penalty measurements, emulations, and simulations for optical preamplifiers shown in Fig. 16. For these specific receiver types, the data for NRZ transmission show a good fit to the penalty formula (31) for an A-parameter of about 60 to 70. For RZ transmission the measured A-parameters range from about 15 to 25. Sunnerud et al. [76] use a different approach, reporting simulations for PMD-induced RZ and NRZ system degradation.

At this time the penalty formula [Eq. (31)] should be regarded as a semi-empirical rule. The receiver parameter, A, in particular, should be determined by simulation and experiment [83]. However, the general nature of the formula, particularly the dependence of the penalty on the DGD and the direction of the launch polarization relative to the PSP, is in good agreement with analysis of the moments of the received signal [28, 42, 73].

As reflected in Eq. (31) the power penalty, ϵ, changes with the DGD and with the launch penalty factor, g, given by

$$g = \sin^2\theta = [1 - (\hat{p}\cdot\hat{s})^2] = (\hat{p}\times\hat{s})^2. \tag{32}$$

This factor depends on the alignment between the Stokes vectors, $\hat{s}$, of the input polarization and the instantaneous PSP, $\hat{p}$. We have $g = 0$ for perfect alignment, where there is no power penalty, and a maximum of $g = 1$ when the Stokes vectors $\hat{s}$ and

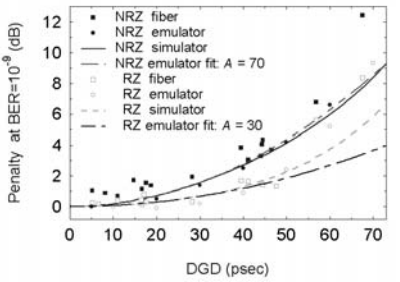

Fig. 16. NRZ and RZ penalties for measurements of real fiber and a first-order PMD emulator, as well as for simulations and optically preamplified receivers plotted as a function of instantaneous DGD. Worst-case polarization launch was used in all cases. Emulator fits use Eq. (31).

$\hat{p}$ are perpendicular (i.e., when equal powers are launched on the PSPs). The latter is the worst-case launch resulting in the maximum power penalty, and is often used in receiver penalty experiments such as those shown in Fig. 16.

4.2. System Outage Due to PMD

Outage specifications for optical fiber transmission systems depend on the application. Usually the power penalty contributions of PMD are required to be less than 1 dB for all but a specified cumulative probability, often a fraction of time ranging from 10^{-4} and 10^{-8}. This translates to average cumulative outages ranging from fractions of a second to 60 minutes per year.

To specify this limit one requires knowledge of the probability density, $p_\epsilon(\epsilon)$, for the occurrence of a power penalty ϵ. The penalty formula (31) shows that ϵ is proportional to the product of the launch penalty g and a DGD term, $\Delta\tau^2$. The probability density of g for uniform distribution over the Poincaré sphere is known (see for example, [48]). The density of $\Delta\tau$ is Maxwellian (Appendix A) from which the density of the $A(\Delta\tau/2T)^2$ can be derived. From these densities, $p_\epsilon(\epsilon)$ can be deduced using standard probability theory for the density of a product [71].

However, there is another, more direct, approach shedding further light on the mechanisms of PMD penalty statistics [48]. In view of

$$\Delta\tau^2(\hat{p}\times\hat{s})^2 = (\boldsymbol{\tau}\times\hat{s})^2 = (\boldsymbol{\tau}_\perp\times\hat{s})^2 = \Delta\tau_\perp^2,$$

we can rewrite the first-order penalty formula as

$$\epsilon = (A/4T^2)\Delta\tau_\perp^2, \tag{33}$$

where $\boldsymbol{\tau}_\perp$ is the component of $\boldsymbol{\tau}$ perpendicular to $\hat{s}$, and $\Delta\tau_\perp$ is its magnitude. The penalty ϵ is caused only by the components of the PMD vector $\boldsymbol{\tau}$ that are perpendicular to $\hat{s}$. The statistics of the two components of $\boldsymbol{\tau}_\perp$ are described by two independent Gaussians. The magnitude $\Delta\tau_\perp$, therefore, follows a Rayleigh distribution, with the probability density

$$p_{\Delta\tau_\perp}(x) = \frac{x}{\sigma^2}\cdot e^{-x^2/2\sigma^2}, \tag{34}$$

where