

Modelling of polarization mode dispersion in optical communications systems

Mark Shtaif¹ and Antonio Mecozzi²

¹ School of Electrical Engineering
Dept. of Physical Electron.
Tel-Aviv University
Tel-Aviv, Israel
Email: shtaif@eng.tau.ac.il

² Department of Electrical Engineering
University of L'Aquila
L'Aquila, Italy
and
Istituto Nazionale di Fisica della Materia
Italy
Email: amecozzi@ing.univaq.it

Abstract. With the rapid increase in the data rates transmitted over optical systems, as well as with the recent extension of terrestrial systems to ultra-long haul reach, polarization mode dispersion (PMD) has become one of the most important and interesting limitations to system performance. This phenomenon originates from mechanical and geometrical distortions that break the cylindrical symmetry of optical fibers and create birefringence. It is the random variations of the local birefringence along the propagation axis of the optical fiber that create the rich and complicated bulk of phenomena that is attributed to PMD. The detailed statistical properties of the local birefringence and its dependence on position are only important as long as the overall system length is comparable with the correlation length of the birefringence in the fiber. In typical systems, however, the latter is smaller by more than three orders of magnitude so that the specific properties of the local birefringence become irrelevant. Instead, the fiber can be viewed as a concatenation of a large number of statistically independent birefringent sections characterized only by the mean square value of their birefringence. This model has been used extensively in the study of PMD and its predictions have been demonstrated to be in excellent agreement with experimental results. This approach opens the door to the world of stochastic calculus, which offers many convenient tools for studying the PMD problem. In this article we review the modelling of PMD and

discuss the properties of this phenomenon as a stochastic process. We explain the use of stochastic calculus for the analysis of PMD and describe the derivation of the frequency autocorrelation functions of the PMD vector, its modulus and the principal states. Those quantities are then related to commonly used parameters such as the bandwidth of the first order PMD approximation, the bandwidth of the principal states and to the accuracy of PMD measurements.

1. What is PMD?

When the cylindrical symmetry of optical fibers is disturbed as a result of mechanical stress, or geometrical distortions, the transmission properties of the fiber become polarization dependent. If the violation of cylindrical symmetry is constant along the fiber, it is possible to identify two orthogonal states of polarization that are characterized by slightly different refractive indices. We will call these states of polarization the eigenpolarizations of the birefringent fiber. When an arbitrarily polarized optical pulse is injected into such a fiber, its components along the two eigenpolarizations are propagated at different group velocities such that the pulse at the output consists of two identically shaped and orthogonally polarized replicas of the original pulse, that are merely delayed relative to each other (see Fig. 1a). One can imagine that such a distortion can be detrimental to optical communications systems where the injected pulses carry actual data. Nevertheless, this case of constant birefringence constitutes, as we shall see, a very degenerate form of PMD, which although potentially detrimental in its raw effect, can be easily avoided. For example, one could consider launching the signal into only one of the eigenpolarizations in which case it will propagate without distortions. A more realistic scenario is that where the birefringence changes both in its strength (the difference between the refractive indices) and orientation (the eigenpolarizations) along the propagation axis of the optical fiber. To picture what happens in this situation we illustrate in Fig. 1b the case in which the birefringence has a step-wise constant behavior. In other words, the fiber consists of many short sections of constant birefringence. In this case the signal splits into two delayed replicas after every individual section such that a very large number of replicas is present at the link output. If instead of an impulse we consider an input pulse of some finite bandwidth, it smears in the process of propagation and becomes distorted in a manner that cannot be easily characterized in the general case. This is in fact the true effect of PMD. It is interesting that when the bandwidth of the pulse is sufficiently narrow, the distortion that it experiences can be described in terms that are equivalent to the case of constant birefringence. In other words, one could replace the link with a fiber of constant birefringence with some well defined delay and eigenpolarizations that would distort the injected signal in a similar manner. The approximate description of PMD in terms of equivalent birefringence is the so called first order PMD approximation. Generally speaking, it results from the first order expansion of the fiber transmission matrix with respect to frequency, but its justification and range of validity will be discussed in what follows.

A formal treatment of PMD requires some mathematical concepts the first of which is the description of polarizations in Stokes space and their representation on the Poincaré sphere. We will now very briefly review the use of these concepts without going into any of the gory details which are very well covered by the known literature

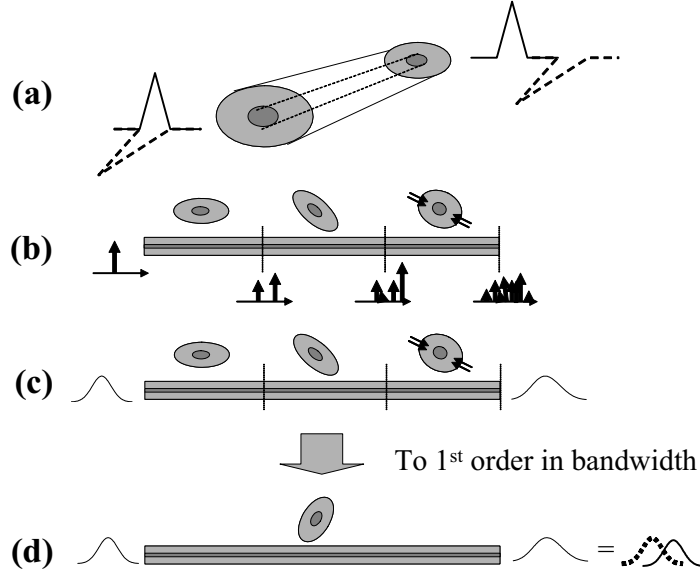


Fig. 1. Constant birefringence introduces a delay between two orthogonal replicas of the transmitted pulse (a). When the fiber consists of many sections with different birefringence, the transmitted pulse splits many times along the propagation and significant distortions to the transmitted waveform may follow as a result of PMD (b). When the injected pulse is sufficiently narrowband (c) the effect of a fiber with changing birefringence becomes equivalent to first order to the effect of constant birefringence (d).

on polarizations [2],[3],[4]. As is well known, an optical signal at any frequency and position along the fiber $\mathbf{E}(\omega, z)$ can be represented by a vector $\mathbf{S}(\omega, z)$ in a three dimensional space called the Stokes space. Unlike the two components of $\mathbf{E}(\omega, z)$ that are complex, the three components of $\mathbf{S}(\omega, z)$ are real, thereby allowing a simple pictorial representation. The vector $\mathbf{S}(\omega, z)$ contains all the information in $\mathbf{E}(\omega, z)$ except for the optical phase of $\mathbf{E}(\omega, z)$ that cannot be extracted from $\mathbf{S}(\omega, z)$.³ The length of the Stokes vector is equal to the optical power $S(\omega, z) = |\mathbf{E}(\omega, z)|^2$ and its orientation $\hat{\mathbf{S}}(\omega, z)$ represents the polarization state of $\mathbf{E}(\omega, z)$. The tip of $\hat{\mathbf{S}}(\omega, z)$ resides on the unit sphere in Stokes space, which is known as the Poincaré sphere, and the various points on its surface uniquely represent the polarization state of $\mathbf{E}(\omega, z)$, as is illustrated in Fig. 2. Notice particularly, that orthogonal polarization states are represented by antipodal points on the sphere. The Stokes space representation has many convenient properties. For example, the effect of constant birefringence is represented in Stokes space by the precession of the vector $\mathbf{S}(\omega, z)$ about some fixed axis, as is illustrated in Fig. 3a. It is not difficult to become convinced that the orientation of that axis must coincide with the eigenpolarizations of the birefringent fiber as they

³ Since the vector $\mathbf{E}(\omega, z)$ has two complex components it consists of 4 degrees of freedom. One of them, the mutual phase of the two components, is lost when shifting to the three-dimensional Stokes space representation.

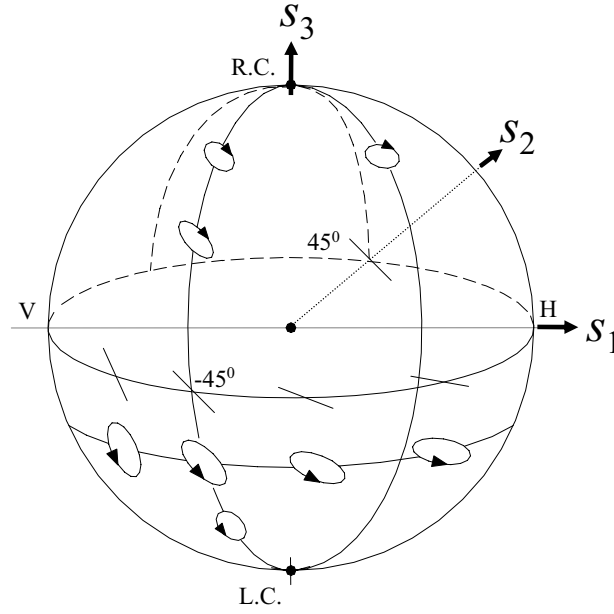


Fig. 2. The relation between the actual polarizations and the position on the Poincaré sphere. The abbreviations H, V, R.C. and L.C. stand for horizontal, vertical, right-circular and left-circular, respectively. After [4].

are represented on the Poincaré sphere. It is convenient to represent birefringence by a vector β in Stokes space whose orientation coincides with the slow (or fast in the formalism of [5]) eigenpolarization state and whose magnitude is equal to the difference between the wave vectors of the two eigenpolarizations. Neglecting the dispersion of the birefringence, one may write $|\beta| = \Delta n \omega / c$ with Δn being the difference in refractive indices and with c being the speed of light. The phase difference acquired between the eigenpolarizations when propagating through a distance z is equal to $|\beta|z$ and it is also equal to the angle that the vector S rotates about the birefringence vector β . It is now instructive to consider the way in which a change in the optical frequency of the launched signal affects the polarization state of the signal at the output of a fiber with constant birefringence. Having neglected the dispersion of the birefringence, the dependence of the output polarization on frequency is identical to its dependence on position because the rotation angle $|\beta|z = \Delta n \omega z / c$ is symmetric with respect to its dependencies on frequency ω and position z . In the following we will refer to the delay vector $\beta'(\omega, z) = \partial \beta(\omega, z) / \partial \omega$ that is equal to the time delay between the two eigenpolarizations per unit of fiber length. The evolution of the output polarization in response to a continuous change in the optical frequency is illustrated in Fig. 3b.

The spectral evolution of the output state of polarization becomes more complicated when the birefringence vector changes along the propagation axis of the fiber. Once again we consider for simplicity the stepwise constant evolution of the birefringence vector that corresponds to the concatenation of many constant birefringence sections with different birefringence vectors. This situation is illustrated in Fig. 4a.

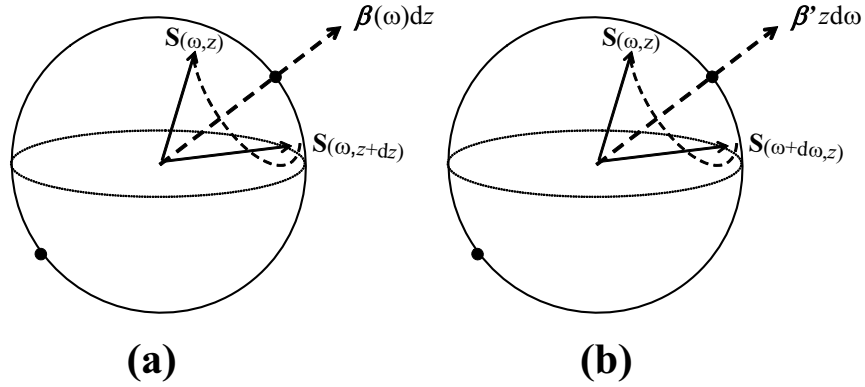


Fig. 3. The effects of changing position (a) along a constant birefringence fiber and of changing the optical frequency (b) are both reflected through a rotation of the polarization around the birefringence vector in Stokes space.

Figures 4b and 4c describe the evolution of the polarization state at two different optical frequencies when the input state of polarization is kept constant. In the process of passing through each one of the constant birefringence sections the state of polarization rotates around the corresponding birefringence vectors and draws an arch on the

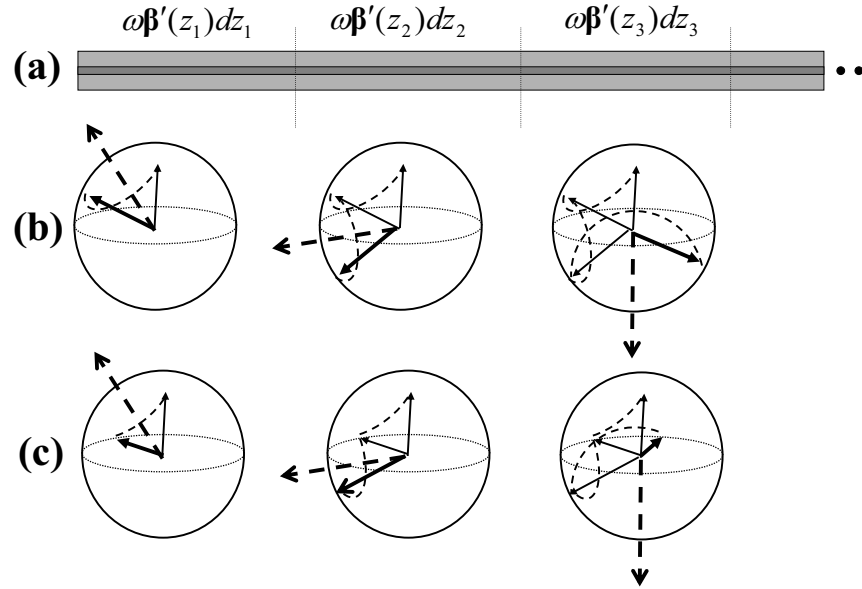


Fig. 4. (a) is an example of a case where the birefringence changes with position along the link. (b) and (c) show what happens to the polarization Stokes vector of the signal when passing through the individual sections. The optical frequency in (c) is slightly lower than in (b) such that all the arches are proportionally shorter. The figure shows that there is no simple relation between the output polarization states that correspond to different frequencies in a complex fiber.

Poincaré sphere. The lengths of the arches are different at the two different frequencies and as a result, the dependence of the output state of polarization on frequency is rather complex. When the number of birefringent sections becomes large the output state of polarization draws a seemingly random curve on the surface of the Poincaré sphere in response to a continuous change in the optical frequency, as is illustrated in Fig. 5. Notice, that locally, in the immediate vicinity of any optical frequency ω_0 the evolution of the curve can be described as precession⁴ about some axis $\tau(\omega_0)$. It is customary to set the length of this vector $\tau(\omega_0)$ to be equal to the rotation angle of the output state of polarization per unit of frequency in the vicinity of ω_0 . When defined in this way $\tau(\omega)$

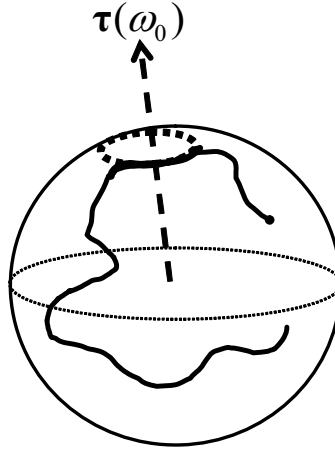


Fig. 5. An illustration of a trajectory that the polarization state at a given position z along the link draws on the Poincaré sphere. Locally the variation of the state of polarization with frequency can be described as rotation around the PMD vector τ .

is called the PMD vector, its orientation $\hat{\tau}(\omega)$ defines the so-called principal states of polarization (PSP)⁵ and its magnitude $\tau(\omega_0) = |\tau(\omega)|$ is the differential group delay (DGD). The physical meaning of $\tau(\omega)$ can be understood through analogy with the birefringence vector. Specifically, if we consider a narrowband optical signal such that within its bandwidth the frequency dependence of τ can be neglected, then the frequency components of that signal lie on an arch that is part of a circle surrounding the vector $\tau(\omega_0)$ with ω_0 being the central optical frequency. In this case the dependence on frequency is identical to what would result from propagation in a constant birefringence fiber whose birefringence vector satisfies $\beta'z = \tau(\omega_0)$. It is therefore clear by analogy that the output pulse will consist of two orthogonally polarized replicas of the original pulse that are delayed relative to each other. The polarizations states of these two replicas coincide with the PSPs and the delay between them is the DGD of the fiber.

⁴ Notice that since the effect of the entire fiber link consists of a concatenation of many infinitesimal rotation operators, one can be certain that indeed the local effect can be described as a rotation around $\tau(\omega_0)$, regardless of what the polarization at ω_0 is.

⁵ More accurately, the principal states are the two polarizations states that are parallel and antiparallel to $\hat{\tau}$. It is customary to define the orientation of $\hat{\tau}$ such that the rotation of \mathbf{S} around it follows the right screw convention. In this case $\hat{\tau}$ coincides with the slow PSP.

With this we justified the first order PMD approximation that we initially introduced earlier. We have also seen that this approximation remains valid as long as the PMD vector can be considered constant within the bandwidth of the transmitted signal. This statement will be further quantified later. To account for the spectral dependence of the PMD vector it is customary to represent it in the form of a Taylor expansion

$$\boldsymbol{\tau}(\omega, z) = \boldsymbol{\tau}(\omega_0, z) + (\omega - \omega_0)\boldsymbol{\tau}'(\omega_0, z) + \frac{1}{2}(\omega - \omega_0)^2\boldsymbol{\tau}''(\omega_0, z) + \dots, \quad (1)$$

where the primes over $\boldsymbol{\tau}$ denote derivatives with respect to frequency. The term $\boldsymbol{\tau}(\omega_0, z)$ represents the first-order PMD approximation since it neglects the spectral variation of PMD altogether. The frequency derivatives of $\boldsymbol{\tau}$ represent the so-called high orders of PMD, where it has become customary to associate the n th frequency derivative of the PMD vector with the $n + 1$ PMD order [6]. Other, alternative, definitions of high orders of PMD have also been proposed recently [7, 8].

2. The PMD Equations

The above description can be mathematically formulated in the following way:

$$\boldsymbol{S}(\omega, z + dz) = \mathbf{R}(\boldsymbol{\beta}(\omega, z)dz)\boldsymbol{S}(\omega, z) \quad (2)$$

$$\boldsymbol{S}(\omega + d\omega, z) = \mathbf{R}(\boldsymbol{\tau}(\omega, z)d\omega)\boldsymbol{S}(\omega, z), \quad (3)$$

where $\mathbf{R}(\mathbf{V})$ is a rotation operator that rotates the vector that it acts upon by an angle $|\mathbf{V}|$ around the axis $\hat{\mathbf{V}}$. This operator can also be represented in an exponential operator form as $\mathbf{R}(\mathbf{V}) = \exp(\mathbf{V} \times)$, with $(\mathbf{V} \times)$ representing the operation of a cross product. When the rotation operators are expanded to first order with respect to their argument ($\mathbf{R}(\boldsymbol{\beta}dz) \simeq 1 + dz\boldsymbol{\beta} \times$ and $\mathbf{R}(\boldsymbol{\tau}d\omega) \simeq 1 + d\omega\boldsymbol{\tau} \times$) the more familiar form of expressions (2) and (3) follows [9, 10, 11]:

$$\frac{\partial \boldsymbol{S}}{\partial z} = \boldsymbol{\beta} \times \boldsymbol{S} \quad (4)$$

$$\frac{\partial \boldsymbol{S}}{\partial \omega} = \boldsymbol{\tau} \times \boldsymbol{S}. \quad (5)$$

Furthermore, by combining Eqs. (2) and (3) and using the algebra of rotations, it can also be shown (see appendix A and [11, 12]) that the evolution of $\boldsymbol{\tau}(\omega, z)$ is described by

$$\boldsymbol{\tau}(\omega, z + dz) = \mathbf{R}(\boldsymbol{\beta}(\omega, z)dz)\boldsymbol{\tau}(\omega, z) + \boldsymbol{\beta}'(\omega, z)dz, \quad (6)$$

where we recall that $\boldsymbol{\beta}'$ is the delay vector, which is equal to the frequency derivative of $\boldsymbol{\beta}(\omega, z)$. We should stress that while the vector $\boldsymbol{\tau}(\omega, z)$ in Eq. (6) reflects the PMD that is accumulated between the beginning of the link and point z , the vectors $\boldsymbol{\beta}(\omega, z)dz$ and $\boldsymbol{\beta}'(\omega, z)dz$ represent the birefringence and delay, respectively, of the section of fiber that starts at z and ends at $z + dz$. This clarification becomes important when the birefringence vector is characterized by abrupt changes along the link as in the case where the fiber is constructed from many sections of constant birefringence. Equation (6) can also be simplified if the rotation operator is expanded to first order in dz , but let us defer this simplification to a later stage, for reasons that will become clear shortly.

The details of the evolution of the birefringence vector $\beta(\omega, z)$ along the fiber and its statistical properties have so far been completely ignored in our discussion. The study of the spatial variation of the birefringence is, in fact, a subject of continuing research [13]. Nevertheless, in most cases of practical interest the detailed statistics of the birefringence has little significance in predicting experimental results. At least, this appears to be the case with long optical links where the overall length of the link is significantly larger than the correlation length of the birefringence, which is the length-scale over which the birefringence vector loses its dependence on the initial conditions. Measurements of the correlation length [13] seem to indicate that it is within hundreds of meters, whereas the length of optical communications links is usually of the order of hundreds to thousands of kilometers - three orders of magnitude larger. Under such conditions it appears that the statistical details of $\beta(z)$ lose their relevance and only the mean square value of $\beta(z)$ determines the effect on signal propagation. This is in fact the limit in which $\beta(z)$ can be safely approximated as white Gaussian noise, meaning that $\langle \beta(z) \cdot \beta(z') \rangle \propto \delta(z - z')$ and $\langle \beta(z) \times \beta(z') \rangle = 0$, where the angular brackets denote ensemble averaging. In this situation the fiber-optic link can be legitimately described in terms of a large number of constant birefringence sections with statistically independent vectors of birefringence. Thus, we have now justified the picture that we have previously introduced merely for illustration purposes. Notice that the white noise approximation applies only to the dependence of the birefringence vector on position. Its dependence on frequency can be effectively approximated as

$$\beta(\omega, z) = \beta(\omega_0, z) + (\omega - \omega_0)\beta'(\omega_0, z). \quad (7)$$

The frequency independent part $\beta(\omega_0, z)$ can be safely ignored since it has no effect on the waveform distortions that PMD generates in optical communications systems. Formally, this term can be removed from all equations by a simple change of variables as can be found in [14]. We will also simplify the notation further by setting $\omega_0 = 0$, such that ω represents the offset from the central frequency of the signal. Finally, instead of referring to the birefringence vector β , or to the delay vector β' that are infinitely wild, it is more convenient to define $d\mathbf{W}(z) = \beta' dz$ such that $d\mathbf{W}(z)$ is a standard three dimensional Brownian motion whose most relevant property is

$$\langle dW_i(z)dW_j(z') \rangle = \frac{1}{3}\gamma_0^2\delta(z - z')\delta_{i,j}dzdz' + r(dz), \quad (8)$$

where dW_i and dW_j are the i and j components of $d\mathbf{W}$ (i, j are equal to 1, 2 or 3), γ_0 is a proportionality coefficient that will later be tied to the mean square DGD value and $r(dz)$ is a zero mean remainder whose dependence on dz satisfies $\lim_{dz \rightarrow 0} r(dz)/dz = 0$ in the mean square sense. Notice that using equation (8) one can show that the mean value of $\langle d\mathbf{W}(z) \cdot d\mathbf{W}(z) \rangle$ is $\gamma_0^2 dz$. Recall that as we explained following Eq. (6), $d\mathbf{W}(z)$ corresponds to the delay vector of the fiber section between z and $z + dz$, implying that $d\mathbf{W}(z)$ and $\tau(z)$ are statistically independent quantities.

We are now ready to transform Eq. (6) to a more usable form. This is done by expanding the rotation operator $R(\omega\beta'dz) = \exp(\omega d\mathbf{W} \times)$ to first order in dz . Notice, however that in view of (8), first order in dz implies second order in $d\mathbf{W}$, namely $\exp(\omega d\mathbf{W} \times) \simeq 1 + (\omega d\mathbf{W} \times) + 1/2(\omega d\mathbf{W} \times)(\omega d\mathbf{W} \times)$, which reduces after some algebraic manipulation to $1 + \omega d\mathbf{W} \times - \gamma_0^2(\omega^2/3)dz$. This allows us to rewrite Eq. (6) in the more convenient form [6]:

$$d\boldsymbol{\tau} = \omega d\mathbf{W} \times \boldsymbol{\tau} + d\mathbf{W} - \frac{\omega^2}{3} \gamma_0^2 \boldsymbol{\tau} dz. \quad (9)$$

It is important to stress that the rightmost term in (9) came up because of the fact that we approximate the birefringence of the fiber as a white noise process⁶. In the absence of that approximation this term disappears and Eq. (9) reduces to the familiar form introduced by Poole in 1991 [10]. The proportionality coefficient γ_0 can now be easily tied to the physical parameters of the system. Setting the frequency to be $\omega = 0$ we obtain $\boldsymbol{\tau} = \mathbf{W}$ as can be easily seen from Eq. (9). The mean square DGD is then $\langle \tau^2 \rangle = \gamma_0^2 z$, implying that γ_0^2 is simply the mean square DGD per unit of fiber length. Furthermore we find that $\boldsymbol{\tau}(\omega, z)$ is a Gaussian vector and therefore its magnitude $\tau = |\boldsymbol{\tau}|$ has a Maxwellian distribution [15] such that the probability density function of τ is given by

$$P(\tau) = \sqrt{\frac{54}{\pi}} \frac{\tau^2}{\langle \tau^2 \rangle^{3/2}} \exp\left(-\frac{3\tau^2}{2\langle \tau^2 \rangle}\right). \quad (10)$$

Our choice of $\omega = 0$ does not limit the generality of the result owing to the stationarity of PMD with respect to frequency. This can be learned from the same transformation that allowed the removal of the frequency independent term $\beta(\omega_0, z)$ of Eq. (7) [14].

Finally, by normalizing the distance, the frequency and the modulus of the PMD vector in the following way [16]:

$$z_N = z/L, \quad (11)$$

$$\omega_N = \omega \sqrt{\langle \tau^2 \rangle}, \quad (12)$$

$$\boldsymbol{\tau}_N = \boldsymbol{\tau} / \sqrt{\langle \tau^2 \rangle}, \quad (13)$$

we arrive at a universal equation for the evolution of PMD

$$d\boldsymbol{\tau}_N = \omega_N d\mathbf{W}_N \times \boldsymbol{\tau}_N + d\mathbf{W}_N - \frac{\omega_N^2}{3} \boldsymbol{\tau}_N dz_N, \quad (14)$$

where $d\mathbf{W}_N(z)$ is a standard Brownian motion in three dimensions, satisfying Eq. (8) with $\gamma_0 = 1$. This equation is powerful because of the fact that it has no adjustable parameters. When it is solved (whether numerically or analytically), its solution can be scaled rigorously to any fiber, based solely on the mean square value of the total DGD of that fiber. In other words this normalized equation implies that *any* statistical property of PMD is determined completely by the mean DGD of the fiber. In the following we will only consider the normalized quantities in our analysis, but the subscript N will be omitted in order to simplify the notation. Notice also that Eq. (14) (as well as Eq. (9) for that matter) is a standard stochastic differential equation expressed in the Ito form [14, 17]. Using this formulation opens the door to the rich toolbox of stochastic calculus which we will use in the sections that follow.

⁶ In the terminology of stochastic calculus Eq. (9) is called an Ito stochastic differential equation [17] and the right-most term is called the Ito correction term that takes into account the peculiarities of white noise. One could use an alternative, Stratonovich formulation in which that term does not exist, but the terms $d\mathbf{W}(z)$ and $\boldsymbol{\tau}(z)$ are no longer independent. The Stratonovich formulation ends up being more complicated and less natural in implementations.

3. The Statistics of the Frequency Dependence of PMD and Its Related Quantities

The spectral behavior of PMD is of utmost importance in determining the effect of PMD in optical communications systems. To have a full understanding of the spectral dependence, one needs to obtain a complete description of PMD as a stochastic process in frequency, a goal which has so far been unaccomplishable. Instead, we will ask and answer somewhat less ambitious questions concerning the spectral dynamics of PMD. The first thing that we will want to know is how much the PMD vector changes on average in response to a given change in the optical frequency. The answer to this question will give us an idea concerning the bandwidth in which the first order PMD approximation remains valid. We will then ask more specific questions related to the dependence of the PSPs and the DGD on the optical frequency. The answer to the latter question will later be used for quantifying the accuracy of PMD measurements. Other, more advanced questions that we shall answer relate to the conditional moments of the PMD vector at one frequency when the PMD at some given other optical frequency is known.

An important quantity that answers the above questions to some extent is the frequency autocorrelation function [18, 19]. We may think of several such functions that are of interest. The first is the autocorrelation function of the PMD vector itself and it is defined as $\langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle$ where we have left the dependence on z implicit. Owing to the stationarity of PMD with respect to frequency, we expect this function to depend only on the frequency difference $\Delta\omega = \omega' - \omega$. This function gives us an idea on how much the optical frequency can be varied before the PMD vector loses memory of its initial value. The disadvantage of this function is that it treats the PMD vector as a whole and does not differentiate between its orientation (the PSPs) and its modulus (the DGD). The dependence of these individual quantities on the optical frequency can be characterized by relating to two additional autocorrelation functions; one for the DGD $\langle \tau(\omega) \tau(\omega') \rangle$ and one for the PSPs $\langle \hat{\boldsymbol{\tau}}(\omega) \cdot \hat{\boldsymbol{\tau}}(\omega') \rangle$ [16]. The width of all three of these quantities when expressed as a function of $\Delta\omega$ will provide us with an idea for the bandwidth in which the first order PMD approximation remains valid.

We will now briefly describe the extraction of these quantities. We start with the vector autocorrelation function $\langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle$ which can be obtained from a differential equation that is constructed as follows,

$$d\langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle = \langle \boldsymbol{\tau}(\omega) \cdot d\boldsymbol{\tau}(\omega') \rangle + \langle d\boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle + \langle d\boldsymbol{\tau}(\omega) \cdot d\boldsymbol{\tau}(\omega') \rangle, \quad (15)$$

where we may use Eq. (14) in order to express the differentials $d\boldsymbol{\tau}(\omega)$ and $d\boldsymbol{\tau}(\omega')$ which can be substituted into Eq. (15). If we recall that the average of terms that contain $d\mathbf{W}$ (in an odd power) is 0 and that $\langle d\mathbf{W} \cdot d\mathbf{W} \rangle = dz$ we obtain the equation

$$\frac{d\langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle}{dz} = 1 - \frac{1}{3} \Delta\omega^2 \langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle. \quad (16)$$

Equation (16) can be easily solved to yield the following autocorrelation function at the end of the link ($z = 1$)

$$\langle \boldsymbol{\tau}(\omega) \cdot \boldsymbol{\tau}(\omega') \rangle = \frac{3}{\Delta\omega^2} \left(1 - e^{-\Delta\omega^2/3} \right), \quad (17)$$

a result that has been originally obtained in [18] and [19]. We should note that the above derivation becomes particularly simple if one exploits the stationarity of PMD with respect to frequency and sets in Eq. (15) $\omega = 0$ and $\omega' = \Delta\omega$. The autocorrelation function (17) is plotted in Fig. 6. The autocorrelation bandwidth of the

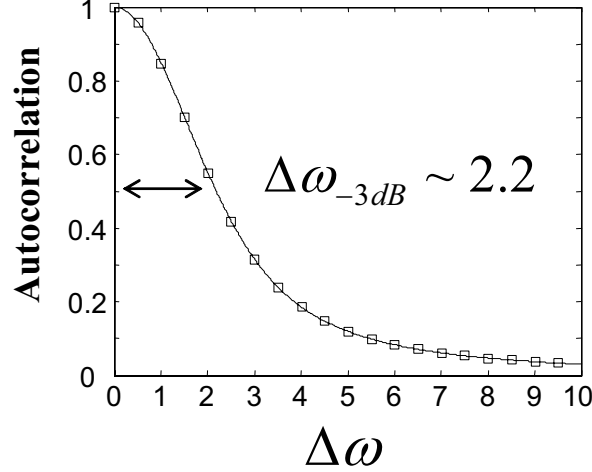


Fig. 6. The autocorrelation function of the PMD vector in normalized units.

PMD vector, which we define by the point where the autocorrelation function reduces to half of its maximal value is equal approximately $\Delta\omega_c = 2.2$. In real world units and in actual frequency (as opposed to angular frequency) this corresponds to $0.35/\sqrt{\langle\tau^2\rangle}$, or equivalently to $0.32/\langle\tau\rangle$ where $\langle\tau\rangle$ is the mean DGD and it is equal to $\langle\tau\rangle = \sqrt{8/(3\pi)\langle\tau^2\rangle}$ as is dictated by the relation between the first two moments of a Maxwellian random variable. The frequency autocorrelation function can also be used to derive the variances and covariances of the various PMD orders that are defined based on the Taylor expansion of $\tau(\omega)$ that we described above. By taking the corresponding derivatives of $\langle\tau(\omega) \cdot \tau(\omega')\rangle$ with respect to ω and ω' one obtains

$$\langle\tau^{(n-k)} \cdot \tau^{(n+k)}\rangle = \frac{(2n)!}{3^n(n+1)!} \langle\tau^2\rangle^{n+1}, \quad (18)$$

$$\langle\tau^{(n-k)} \cdot \tau^{(n+k+1)}\rangle = 0, \quad (19)$$

where $\tau^{(j)}$ denotes the j th derivative of τ with respect to frequency, or the $j-1$ order of PMD. Setting $k=0$, we find that the righthand side of Eq. (18) is the variance of the $n-1$ order of PMD. In a similar calculation to the one leading to Eq. (17) we can obtain a rigorous expression for the frequency autocorrelation of the square modulus of the PMD vector (or the square DGD) [16],

$$\langle\tau^2(\omega)\tau^2(\omega')\rangle = \langle\tau^2\rangle^2 + \frac{4\langle\tau^2\rangle}{\Delta\omega^2} - \frac{12}{\Delta\omega^4} \left[1 - e^{-\Delta\omega^2/3}\right], \quad (20)$$

which is plotted together with the results of computer simulations in Fig. 7. The width at half maximum of this function is given by $\Delta\omega_c = 2.8$. Unfortunately the same tech-

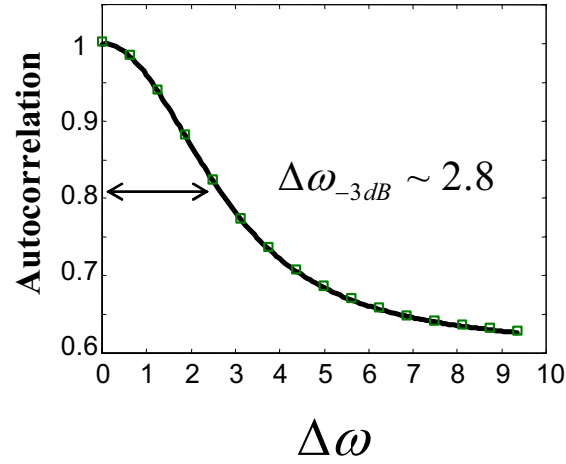


Fig. 7. Autocorrelation function of the square DGD [equation (20)] as well as results of a Monte Carlo simulation of a fiber with 1000 independent sections (squares).

nique does not allow the rigorous analytical derivation of the autocorrelation functions of the DGD and of the PSPs. These quantities can however be numerically extracted from Eq. (14) and since, as we have already explained, this equation is universal, the results can be scaled to any fiber based on its mean DGD value. The results for the autocorrelation functions of the DGD and the PSPs appear in Figure 8 together with a comparison with the autocorrelation functions obtained from experimental data. The correlation bandwidths are $\Delta\omega_c = 2.2$ for the PSPs and $\Delta\omega_c = 2.8$ for the DGD. The two quantities decorrelate at a fairly similar rate, contrary perhaps to some of the common intuition. Nevertheless, since the autocorrelation function of the DGD only changes within 15% of its maximal value, its contribution to the decorrelation of the entire PMD vector is secondary. As a result the autocorrelation function of the PSPs is very similar to the autocorrelation function of the PMD vector (17) as can be observed in Fig. 9. Notice that the autocorrelation function of the PSPs is slightly narrower than the vector autocorrelation function. This is opposite to what would be the case if the DGD and the PSPs were statistically independent, where $\langle \tau(\omega) \cdot \tau(\omega') \rangle$ would be equal to $\langle \tau(\omega) \tau(\omega') \rangle \langle \hat{\tau}(\omega) \cdot \hat{\tau}(\omega') \rangle$. It is interesting that an excellent analytical approximation for the autocorrelation function of the DGD can also be obtained by comparison between the autocorrelation functions $\langle \tau(\omega) \tau(\omega') \rangle$ and $\langle \tau^2(\omega) \tau^2(\omega') \rangle$. In Fig. 10 we plot these two functions after subtracting their values at $\Delta\omega \rightarrow \infty$ and normalizing their peaks to unity. Notice the great resemblance between the two functions that is evident both on a linear and a logarithmic scale. This suggests that the autocorrelation of the DGD can be very accurately approximated by the expression

$$\langle \tau(\omega) \tau(\omega') \rangle \simeq \frac{3\pi - 8}{\pi} \left[\frac{2}{\Delta\omega^2} - \frac{6}{\Delta\omega^4} \left(1 - e^{-\Delta\omega^2/3} \right) \right] + \frac{8}{3\pi} \langle \tau^2 \rangle, \quad (21)$$

which is obtained from Eq. (20).

Finally, it is also interesting to mention without derivation a number of more advances quantities that reflect the spectral behavior of PMD as a stochastic process in

frequency. Two such quantities are the mean and the mean square value of PMD at one frequency when the value of the PMD vector at some other, offset optical frequency is known. These results are presented below in our normalized units [20]

$$\langle \tau(\omega) | \tau(\omega') = \tau_0 \rangle = \frac{3\tau_0}{\Delta\omega^2} \left[1 - \exp\left(-\frac{\Delta\omega^2}{3}\right) \right], \quad (22)$$

$$\langle \tau^2(\omega) | \tau(\omega') = \tau_0 \rangle = 1 + (\tau_0^2 - 1) \frac{6}{\Delta\omega^4} \left[\Delta\omega^2 - 3 + 3 \exp\left(-\frac{\Delta\omega^2}{3}\right) \right]. \quad (23)$$

Notice that expression (22) has exactly the same functional dependence on optical frequency as does the autocorrelation function of the PMD vector (17). The evolution of the PMD from its value at ω' (which is τ_0) to its unconditional value (which is 0) at large ω , is therefore characterized only by the autocorrelation bandwidth. A property that is common to both conditional moments of PMD is that the functional dependence on the frequency offset is independent of τ_0 . This feature is obvious in the case of the first moment (22) and it can be easily observed in the case of Eq. (23) by expressing the quantity $(\langle \tau^2(\omega) | \tau_0 \rangle - \langle \tau^2 \rangle) / (\tau_0^2 - \langle \tau^2 \rangle)$ and showing that it does not depend on τ_0 . To visualize the consequences of this property, consider for example a WDM optical communications system in which one of the channels suffers from an outage due to PMD. One may think that the rate at which the PMD vector de-correlates may be affected by the fact that the DGD at the optical frequency of the impaired channel is very large. The conditional moments (22) and (23) clearly refute this stipulation. Perhaps somewhat contrary to what may be considered intuitive, expressions (22) and (23) indicate that the rate at which the PMD vector loses memory of its value does not depend on whether its modulus, the DGD, is significantly different from its average. It is perhaps relevant to discuss the value that the results (22) and (23) add to the PMD vector autocorrelation function (17) whose form is nearly identical to that of equation (22). The autocorrelation function of the DGD describes the way in which the PMD

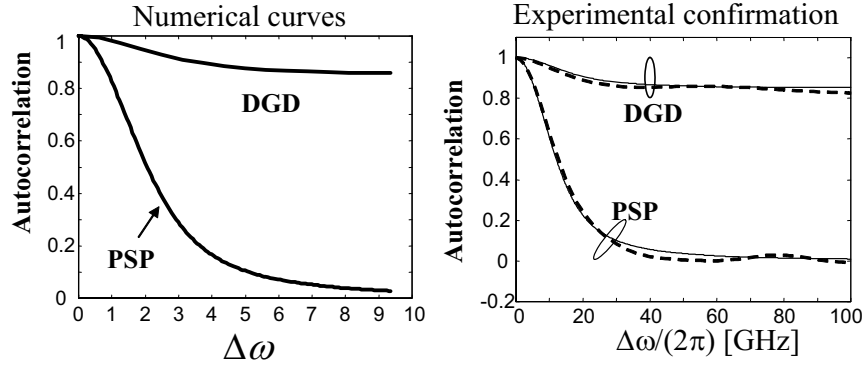


Fig. 8. On the left are the numerically extracted autocorrelation functions for the DGD and the PSPs in normalized units. On the right figure the numerical curves are compared with the results in a simulated fiber with mean DGD of 27ps (after [16]). The oscillations in the experimental curves result from a limited measurement bandwidth and they can be reproduced in simulations [19]

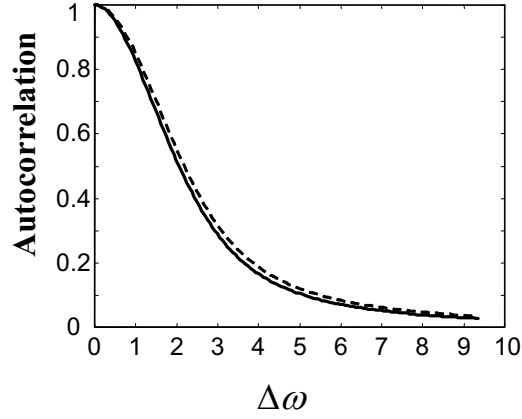


Fig. 9. Autocorrelation functions of the PSPs (solid) and of the PMD vector (17) (dashed).

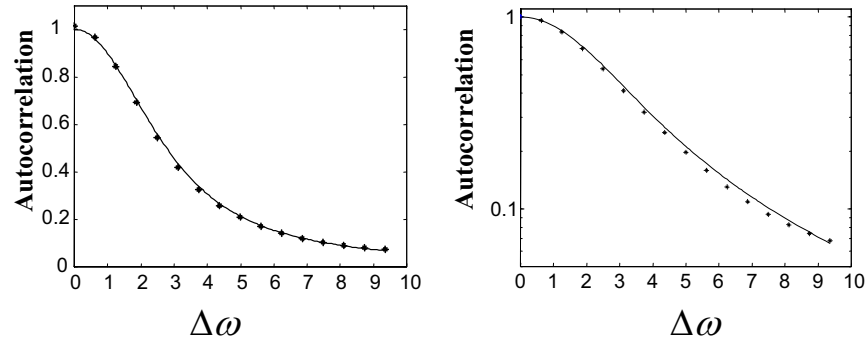


Fig. 10. The normalized autocorrelation functions of the DGD (dots) and of the square DGD (solid curve). The left-hand figure is plotted on linear scale and the right-hand figure is plotted on a logarithmic scale.

vector loses its memory on average, independent of whether its value at some particular frequency is known. It can in fact be obtained from the conditional first moment (22) through the relation $\langle \tau_0 \cdot \tau_\omega \rangle = \langle \tau_0 \cdot \langle \tau_\omega | \tau_0 \rangle \rangle$, whereas the opposite is of course not true.

4. Assessment of the Accuracy of PMD Measurements

We have already emphasized that the only parameter determining the statistical properties of PMD in long optical fibers is the mean DGD. In this section we discuss the accuracy of its measurements [16],[21]. The experimental assessment of the mean DGD is based on the spectral averaging of the DGD that is measured in a broad range of frequencies. In the limit of very large bandwidth, this procedure yields a value that

converges to the ensemble average of the DGD⁷, which we have previously denoted by $\langle \tau \rangle$. Yet, when the measured bandwidth is limited, the measured frequency average is a random variable whose mean is the desired quantity $\langle \tau \rangle$ and whose standard deviation reflects the accuracy of the measurement. Using the autocorrelation function for the DGD, which we presented in the previous section, we may now obtain an explicit expression for the accuracy of the measurement. Formally, the measured frequency average is given by

$$\bar{\tau}(B) = \frac{1}{B} \int_{-B/2}^{B/2} d\omega \tau(\omega). \quad (24)$$

The mean of $\bar{\tau}$ is obviously equal to $\langle \tau \rangle$ and its variance is given by

$$\begin{aligned} \text{Var}(\bar{\tau}(B)) &= \langle \bar{\tau}^2(B) \rangle - \langle \tau \rangle^2 \\ &= \frac{1}{B^2} \int_{-B/2}^{B/2} d\omega \int_{-B/2}^{B/2} d\omega' \langle \tau(\omega) \tau(\omega') \rangle - \langle \tau \rangle^2. \end{aligned} \quad (25)$$

Then, using Eq. (21) for $\langle \tau(\omega) \tau(\omega') \rangle$ one can obtain a lengthy, but analytical solution that is plotted in Fig. 11. Its width at half maximum is 10.4, approximately 4 times larger than the correlation bandwidth limit of the DGD. If we concentrate on the important limit of such measurements, where $B \gg 1$ then expression (25) approaches $\text{Var}(\bar{\tau}(B)) \simeq 1.23/B$, or $\text{Var}(\bar{\tau}(B)) \simeq 1.34\langle \tau \rangle/B$ in real units. This asymptotic solution is shown by the dashed curve in Figure 11 and it becomes an excellent approximation to the complete solution when $B > 30$.

An alternative method for extracting the mean DGD is to measure the frequency average of the square of the DGD and then use the relation $\langle \tau \rangle = \sqrt{8\langle \tau^2 \rangle / (3\pi)}$. The measured quantity in this case is

$$\bar{\tau}^2(B) = \frac{1}{B} \int_{-B/2}^{B/2} d\omega \tau^2(\omega) \quad (26)$$

and its variance can be expressed in a way similar to equation (25), except that the autocorrelation function of the square DGD (20) needs to be placed inside the integral. The asymptotic solution in this case is $\sqrt{0.5\pi(8/3)^3}/B$, or in real units $(\pi^2/\sqrt{2})\langle \tau \rangle^3/B$. If we use this value to obtain the variance of the mean DGD that we extract with this method we obtain $\text{Var}[\sqrt{8\bar{\tau}^2(B)/(3\pi)}] \simeq 1.26\langle \tau \rangle/B$, slightly better than what we obtain from the frequency averaging of τ . A point to notice with the second method for measuring the mean DGD is that the obtained result is slightly biased on average with respect to the mean DGD value [22], meaning that the mean of the measured average is not equal to the mean DGD. Nevertheless, the difference between the two values in the relevant range of parameters ($B \gg 1$) is approximately $\langle \tau \rangle - \langle \bar{\tau}(B) \rangle \simeq 0.6/B$ implying an insignificant estimation error. Figure 11 shows that the measured bandwidth should always be large enough (relative to $\langle \tau \rangle^{-1}$) in order to ensure the accuracy of the result. When the mean DGD of the link is small, as is the case in many modern communications systems operating on new fiber, an accurate measurement of PMD often requires an unrealistically broad measurement bandwidth and the development of alternative methods that optimize measurement accuracy becomes highly relevant [22].

⁷ This is because $\langle \tau(\omega) \tau(\omega') \rangle \rightarrow \langle \tau \rangle^2$ when $\Delta\omega \rightarrow \infty$, which ensures that $\tau(\omega)$ is mean ergodic with respect to frequency.

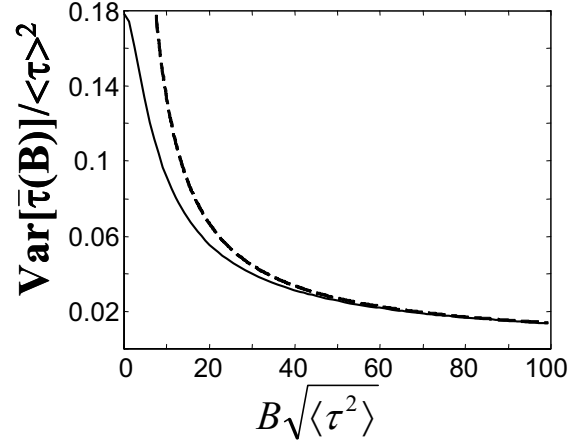


Fig. 11. Variance of the measured mean DGD when it is frequency averaged over an angular frequency bandwidth B and divided by the square of the mean DGD. The dashed curve shows the asymptotic dependence $\text{Var}(\bar{\tau}(B)) \simeq 1.34\langle\tau\rangle/B$.

5. Final Note

It is well recognized that PMD is one of the most significant limiting mechanisms in optical communications and as such it has been the subject of very extensive research. In this article we assumed the most broadly used model for PMD, where the fiber is described as a concatenation of a large number of statistically independent constant birefringence sections. Over the years this model has proven itself as extremely robust, and predictions that have been made on its basis were found to be in excellent agreement with experimental results. Nevertheless, as the field of optical communications constantly evolves, it may become necessary to consider situations that require a more refined modelling of the polarization properties of optical fibers. For example, this may be the case when the fiber link is known to have a small number of high PMD contributors that are placed within an otherwise very low PMD fiber. Studies of the evolution of PMD in time may also require a more careful description of the fiber. So far, however, it has been extremely difficult to refine the modelling of optical fibers in a way that is both consistent with the experimental reality and manageable in the mathematical sense. The search for such models will remain an active topic in the future research of PMD.

A. The Derivation of Eq. (6)

Equation (2) can be used to express $\mathcal{S}_{\omega+d\omega, z+dz}$ in the following way

$$\mathcal{S}_{\omega+d\omega, z+dz} = \mathcal{R}[\beta_{\omega+d\omega, z} dz] \mathcal{S}_{\omega+d\omega, z}, \quad (27)$$

where in order to shorten the notation we use indices to denote the dependence of the various parameters on z and ω [for example $\mathbf{S}_{z,\omega} = \mathbf{S}(z, \omega)$ and so on]. Using Eq. (3) we get

$$\mathbf{S}_{\omega+d\omega, z+dz} = \mathbf{R}[\boldsymbol{\beta}_{\omega+d\omega, z} dz] \mathbf{R}[\boldsymbol{\tau}_{\omega, z} d\omega] \mathbf{R}^{-1}[\boldsymbol{\beta}_{\omega, z} dz] \mathbf{S}_{\omega, z+dz}. \quad (28)$$

Expanding $\boldsymbol{\beta}_{\omega+d\omega, z} \simeq \boldsymbol{\beta}_{\omega, z} + \boldsymbol{\beta}'_{\omega, z} d\omega$ and concentrating on the cases where $\boldsymbol{\beta}'(\omega, z)$ is parallel to $\boldsymbol{\beta}(\omega, z)$, we may rewrite the above equation in the form

$$\mathbf{S}_{\omega+d\omega, z+dz} = \mathbf{R}[\boldsymbol{\beta}_{\omega+d\omega, z} dz] \mathbf{R}[\boldsymbol{\tau}_{\omega, z} d\omega] \mathbf{R}^{-1}[\boldsymbol{\beta}_{\omega+d\omega, z} dz] \mathbf{R}[\boldsymbol{\beta}'_z d\omega dz] \mathbf{S}_{\omega, z+dz}, \quad (29)$$

which is equivalent to

$$\mathbf{S}_{\omega+d\omega, z+dz} = \mathbf{R}[\mathbf{R}[\boldsymbol{\beta}_{\omega+d\omega, z} dz] \boldsymbol{\tau}_{\omega, z} d\omega] \mathbf{R}[\boldsymbol{\beta}'_z d\omega dz] \mathbf{S}_{\omega, z+dz}, \quad (30)$$

where we have used the relation $\mathbf{R}(\mathbf{A})\mathbf{R}(\mathbf{B})\mathbf{R}^{-1}(\mathbf{A}) = \mathbf{R}(\mathbf{R}(\mathbf{A})\mathbf{B})$, that corresponds to the rotation of the entire space by the operator $\mathbf{R}(\mathbf{A})$. To first order in $d\omega$, Eq. (30) is equivalent to

$$\mathbf{S}_{\omega+d\omega, z+dz} = \mathbf{R}[\mathbf{R}[\boldsymbol{\beta}_{\omega+d\omega, z} dz] \boldsymbol{\tau}_{\omega, z} d\omega + \boldsymbol{\beta}'_z d\omega dz] \mathbf{S}_{\omega, z+dz}. \quad (31)$$

On the other hand, we can relate $\mathbf{S}_{\omega+d\omega, z+dz}$ with $\mathbf{S}_{\omega, z+dz}$ with the help of Eq. (3), obtaining

$$\mathbf{S}_{\omega+d\omega, z+dz} = \mathbf{R}[\boldsymbol{\tau}_{\omega, z+dz} d\omega] \mathbf{S}_{\omega, z+dz}. \quad (32)$$

Comparison between Eqs. (31) and (32) yields Eq. (6).

References

1. H. Kogelnik, R. M. Jopson and L.E. Nelson, Polarization mode dispersion. In: *Optical Fiber Telecommunications IVb: Systems and Impairments*, edited by I.P. Kaminow and T. Li (Academic Press, San Diego, 2002), Chap. 15.
2. M. Born, E. Wolf, *Principles of Optics*, 6th Ed. (Pergamon, New York, 1986)
3. R.M.A. Azzam, N. M. Bashara, *Ellipsometry and Polarized Light* (NHPL, Elsevier, New York, 1977)
4. D.S. Kliger, J. W. Lewis, C.E. Randall, *Polarized Light in Optics and Spectroscopy* (Academic Press, New York, 1990)
5. C. D. Poole, R. E. Wagner, Electron. Lett., **22**, 1029 (1986)
6. G. J. Foschini, C. D. Poole, J. Lightwave Technol., **9**, 1439 (1991)
7. A. Eyal, W. K. Marshall, A. Yariv, M. Tur, Electron. Lett., **19**, 1658 (1999)
8. H. Kogelnik, J. Opt. Fiber Commun. Rep. **1**, 107–122 (2004) DOI: 10.1007/s10297-004-0005-1
9. R. Ulrich, A. Simon, Appl. Opt., **18**, 2241 (1979)
10. C. D. Poole, J. H. Winters, J. A. Nagel, Opt. Lett., **16**, 372 (1991)
11. J. P. Gordon, H. Kogelnik, Proc. Natl. Acad. Sci. USA, **97**, 4541 (2000)
12. N. J. Frigo, IEEE J. Quantum Electron., **22**, 2131 (1986)

13. A. Galtarossa, L. Palmieri, M. Schiano, T. Tambosso, *Opt. Lett.*, **26**, 962 (2001)
14. J. P. Gordon, *J. Opt. Fiber Commun. Rep.*, **1**, 210–217 (2004) DOI: 10.1007/s10297-004-0003-3
15. F. Curti, B. Daino, G. De Marchis, F. Matera, *J. Lightwave Technol.*, **9**, 1162 (1990)
16. M. Shtaif, A. Mecozzi, *Opt. Lett.*, **25**, 707 (2000)
17. C. W. Gardiner, *Handbook of Stochastic Methods*, Springer Series in Synergetics 13 (Springer, Berlin/Heidelberg/New York, 1985)
18. M. Karlsson, J. Brentel, *Opt. Lett.*, **24**, 939 (1999)
19. M. Shtaif, A. Mecozzi, J. Nagel, *IEEE Photon. Technol. Lett.*, **12**, 53 (2000)
20. A. Mecozzi, M. Shtaif, *IEEE Photon. Technol. Lett.*, **15**, December 2003, available at <http://ieeexplore.ieee.org>.
21. N. Gisin, B. Gisin, J. P. Won der Weid, R. Passy, *IEEE Photon. Technol. Lett.*, **8**, 1671 (1996)
22. M. Boroditsky, M. Brodsky, P. Magill, N.J. Frigo, M. Shtaif, "Second order PMD statistics analyses improves the accuracy of mean DGD estimates," to be published in *IEEE Photon. Technol. Lett.* (2004)

Polarization Mode Dispersion

Galtarossa, A.; Menyuk, C.R. (Eds.)

2005, X, 296 p. 109 illus., Hardcover

ISBN: 978-0-387-23193-8