9

Statistical Properties of Polarization in Fiber

The topic of this chapter is the statistics of polarization, polarization-mode dispersion, and polarization-dependent loss, all in relation to behavior in single-mode fibers. The origin of these statistical properties is the birefringence within the mode-field diameter of the fiber. If a perfectly isotropic fiber existed, the polarization state at the output would match that at the input for all time, frequency, temperature, and length. This, however, is not the case. As illustrated in Fig. 9.1, there are several causes of fiber birefringence, both intrinsic and extrinsic. While an isotropic fiber is stress-free and has a perfectly concentric core (a), perturbations such as core ovality (b), stress-induced index gradient (c), and micro-bubbles (d) introduce birefringence at any cross-section of the fiber. The existence and study of fiber birefringence was well-known by the 1970s [29, 53, 54]. The length-scale of the birefringence is called the birefringent-beat length: $L_B = \lambda_0/\Delta n$, where λ_o is the free-space wavelength and Δn is the birefringence. Typical birefringence values of early 1990s fiber are $\Delta n/n \sim 10^{-7}$, which corresponds to $L_B \simeq 20$ m at 1550 nm.

Birefringence, however, is only one contributing factor to the statistical behavior. A second factor is the fiber autocorrelation length L_C, which is the characteristic length over which the birefringent axis of the fiber changes. While the models used in the following are cast in the continuous limit, at the discrete level one can think of a fiber of length L segmented into N parts L_C long, where $N = L/L_C$. Within a segment the birefringence and its orientation is fixed, and at the junctions between adjacent segments the birefringent axis changes abruptly. In the continuous limit, L_C represents the characteristic length over which the randomly evolving birefringence loses memory of its preceding orientations. A typical value of L_C is 100 m.

The relationship between the fiber autocorrelation length L_C and the birefringent beat length L_B, and between L_C and the fiber length L, determines the regime in which the polarization ensembles behave. The principal regimes are illustrated in Fig. 9.2. In the first limit $L_C \gg L_B$, Fig. 9.2(a), the slow evolution of the birefringent axis along the fiber allows the optical field of the signal to follow. In the second limit $L_C \ll L_B$, Fig. 9.2(b), the birefringent

Jay N. Damask: Polarization Optics in Telecommunications, SSOS 101, 385–428 (2005)

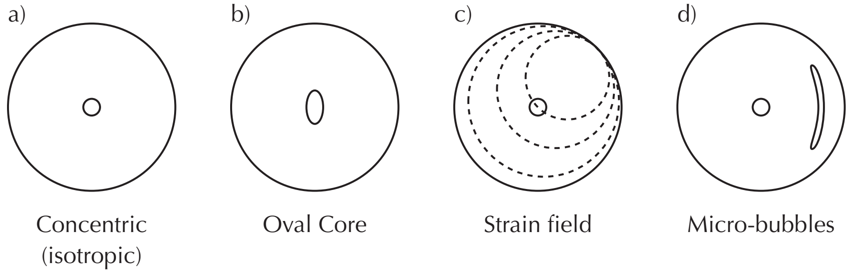

Fig. 9.1. Sources of birefringence within a cross-section of single-mode fiber. a) Perfect, isotropic fiber. b) Ovality of the core. c) Stress-induced index gradient. d) Physical defects like micro-bubbles or impurity concentrations.

axis changes too quickly for the optical field to follow, the result is a long-range average over the range of orientations. This effect is exploited to manufacture ultra-low PMD fiber: the fiber preform is spun during the drawing process at a rate designed to ensure $L_C \ll L_B$ [43]. For instance, Chen *et al.* [6] report a spin period of 1 m and a beat length of 10 m in their fiber. Higher resolution measurements are reported by Pietralunga *et al.* [46] and Galtarossa *et al.* [20]. To be sure, this is not a perfect cure as one must still include some length-scale for variation of the spin profile – this additional factor is illustrated in Fig. 9.2(c) – but the effective birefringence of the fiber is reduced by an order of magnitude.

The relationship between the autocorrelation length L_C and the total fiber length determines how the polarization-mode dispersion behaves. There are two extrema regimes, that of "low" (or "weak") mode coupling and that of "high" (or "strong") mode coupling (Fig. 9.2(d)). In the low mode-coupling regime, variation of birefringence orientation is low so the mean PMD increases linearly with length. In the high mode-coupling regime, the birefringence orientation is random beyond a correlation length so the mean PMD increases as the square-root of the length. As a practical matter, since square-root growth is slower than linear, one would like to reach this regime as quickly as possible. As shown in the following, the ratio L/L_C determines the regime; a low autocorrelation length L_C pushes a fiber toward high mode-coupling and root-length growth of the mean PMD.

An historic anecdote conveys the importance of the fiber autocorrelation length. Early fiber-transmission and characterization studies were done in the research lab where fiber is held on spools. When C. D. Poole went to measure for the first time a fiber spooled and then unspooled, he found that the PMD increased five-fold. This is an instance where the correlation length L_C is small on the spool, due to inhomogeneities of bending strain, and large unspooled. Indeed, de Lignie *et al.* report spooled and cabled measurements circa 1994 where they showed $L_C \sim 5$ m on the spool and $L_C \sim 500$ m cabled [9]. Their measurements from fiber to fiber show a wide range of values.

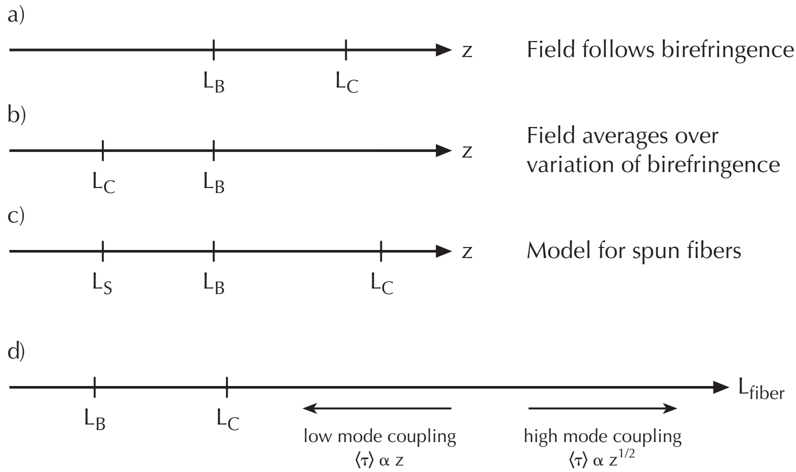

Fig. 9.2. Relationship between length scales within a single-mode fiber. a) Adiabatic regime $L_C \gg L_B$: the field follows the changing birefringence. b) Field-average regime $L_C \ll L_B$: the field cannot follow the changing birefringence vector and instead averages over the variation. c) Model for spun fibers where a third length scale L_S, the range over which the spin profile changes, is added. d) Low- and high-mode coupled PMD regimes: L_C in comparison with the total fiber length L.

For practical systems and design, there are three dimensions along which one would like to derive polarization-related statistics: propagation length, optical frequency, and time. In each case a statistical process must be defined to characterize the evolution on a microscopic level. The PMD evolution equation over length is well defined and the probability density converges in the limit of large ensembles. The ergodic nature of PMD lets "length" be replaced by "optical frequency" in the density functions and "long length" is replaced by "wide bandwidth." The PMD autocorrelation function connects the length and frequency regimes. There is, however, no definite process for the time evolution. Submarine cable changes at a slow rate while aerial fiber changes in the millisecond range. Moreover, there is likely no spatial homogeneity to the temporal changes – for instance, a train may cross a cable at a particular location – so one cannot expect a neat answer. When the temporal changes are spatial homogeneous, P. J. Leo *et al.* have developed a Rayleigh-distribution model of SOP change that is useful to define what "speed" of change means [32].

Statistics for polarization, PMD, and PDL are derived in the following using diffusion processes. The physics of a diffusion process is first captured by a stochastic differential equation (SDE) and then translated to its partial-differential equation (PDE) analogue. Exposition of these mathematical tools is beyond the scope of this text and the reader is referred to Arnold and Oksendal for SDEs [1, 44], and Risken for PDEs [55]. Finally, Davenport is an invaluable reference on applied probability is [8].

9.1 Polarization Evolution Model

An optical mode confined within a fiber propagates only along one dimension; denote this direction z. The longitudinal electric field will propagate according to the time-harmonic factor $\exp\left(-jzk_0\sqrt{\varepsilon_r}\right)$, where k_0 is the free-space wavenumber and ε_r is the relative permittivity of the fiber. The Helmholtz equation for the evolution of the electric field $\mathbf{E}$ in the plane perpendicular to z is therefore

$$\left(\frac{d^2}{dz^2} + k_0^2\vec{\varepsilon}_r\right)\mathbf{E} = 0$$

where $\vec{\varepsilon}_r$ is written in tensor form in anticipation of the following. The common permittivity $\bar{\varepsilon}_r$ may be separated from the differential part such that

$$\vec{\varepsilon}_r = \bar{\varepsilon}_r + \tfrac{1}{2}\Delta\vec{\varepsilon}_r \cdot \vec{\sigma} \tag{9.1.1}$$

When the propagation is lossless, the common permittivity is the square of the refractive index: $\bar{\varepsilon}_r = n^2$. The common phase is removed from $\mathbf{E}$ to isolate the polarization state evolution using the factorization

$$\mathbf{E} = \exp\left(-j\int_0^z k_0\, n(z')\, dz'\right)\ |s\rangle$$

where the integral simply accounts for the cumulative change of common index over the path; if the common index is fixed, the integral reduces to the more customary exponential phase factor. Substitution of this factorization and (9.1.1) into the wave equation, and dropping terms that are second-order in $|s\rangle$, makes

$$\left(\frac{d}{dz} + jk_0^2\frac{\Delta\vec{\varepsilon}_r \cdot \vec{\sigma}}{4n}\right)|s\rangle = 0$$

In the absence of polarization-dependent loss, $\Delta\vec{\varepsilon}_r$ is real and its magnitude to first-order in Δn is

$$\Delta\varepsilon_r = n_+^2 - n_-^2 = (n + \Delta n/2)^2 - (n - \Delta n/2)^2 \simeq 2n\Delta n$$

As a matter of notation, the magnitude of the birefringent vector $\vec{\beta}$ is defined as

$$\beta = |\vec{\beta}| = k_0 \Delta n = \frac{\omega}{c} \Delta n \tag{9.1.2}$$

where the Δ on $\Delta\beta$ has been dropped for convenience. Moreover, the birefringent beat length $L_B = \lambda_o/\Delta n$ is related to the birefringence β as

$$L_B = 2\pi/\beta \tag{9.1.3}$$

With these definitions in hand, the polarization state evolves in the fiber according to [25]

$$\left(\frac{d}{dz} + \frac{j}{2}\, \vec{\beta} \cdot \vec{\sigma} \right) |s\rangle = 0 \tag{9.1.4}$$

The birefringence vector $\vec{\beta}$ is the local birefringence at any position along the fiber. Equation (9.1.4) describes the response of the polarization state due to the local birefringence. Converting to Stokes space, the differential equation of motion is

$$\frac{d\hat{s}}{dz} = \vec{\beta} \times \hat{s} \tag{9.1.5}$$

As expected, the polarization state precesses about the local birefringent axis at a rate governed by the strength of the birefringence.

Wai and Menyuk propose two models of how the local birefringence varies along an unspun fiber [40, 60, 62]. In both models the fiber exhibits no chirality:

$$\text{No circular birefringence:} \quad \beta_3 = 0$$

This assertion has been experimentally verified for such fibers [21], while spun fibers show evidence of residual chirality [27]. The calculations that follow use the no-chirality assumption, while models for spun fibers can be found in [47]. Without a chiral factor, the birefringent matrix is

$$\vec{\beta} \cdot \vec{\sigma} = \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_2 & -\beta_1 \end{pmatrix} \tag{9.1.6}$$

In their first model, the birefringence magnitude is fixed and the angle θ on the Poincaré equator randomly varies. In their second model, the cartesian birefringent components (β_1, β_2) are independent random variables. Both models give the correct evolution of the mean-square DGD, but the latter model, while a bit more involved, generates aperiodic PMD spectra.

9.1.1 Random Birefringent Orientation

For this first model, the birefringent matrix is

$$\vec{\beta} \cdot \vec{\sigma} = \beta \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{9.1.7}$$

where θ is in Stokes space. The angle is modelled as a Brownian motion on $\mathbb{R}^1$ according to

$$\frac{d\theta}{dz} = g_\theta(z) \quad \longrightarrow \quad \langle g_\theta(z)\rangle = 0, \quad \langle g_\theta(z) g_\theta(z')\rangle = \sigma_\theta^2\, \delta(z - z') \tag{9.1.8}$$

where g_θ is a white-noise stochastic process. Physically, the angle θ is subject to impulsive "kicks" as z increases, which in turn drives a random walk in angle. The impulsive nature of the kicks means that there is no memory or correlation from one kick to the next; each is random in its own right.

The Brownian probability density of θ subject to this motion is

$$\rho_\theta(\theta, z) = \frac{1}{\sqrt{2\pi\sigma_\theta^2 z}} \exp\left(-\frac{\theta^2}{2\sigma_\theta^2 z}\right)$$

where, as characteristic with this process, the variance increases linearly with length: $\mathrm{var}(\theta) = \sigma_\theta^2 z$. The "strength" of the white-noise g_θ, σ_θ^2, is now apparent: the stronger the noise the shorter the fiber length is necessary to reach a nearly uniform angular distribution between $[-\pi/2, \pi/2]$.

As with any random walk, eventually there is complete loss of correlation between some earlier position and the present. In this case, the fiber autocorrelation length L_C is defined as the length over which the angle θ losses correlation. The autocorrelation of θ is calculated by the expectation value of $\cos\theta(z)$:

$$E\left[\cos\theta(z)\right] = \int_{\mathbb{R}} \cos\theta \rho_\theta(\theta) d\theta = \exp\left(-\frac{\sigma_\theta^2 z}{2}\right) \tag{9.1.9}$$

The autocorrelation length L_C is the length at which the autocorrelation falls to e^{-1}; therefore,

$$\sigma_\theta^2 = \frac{2}{L_C} \tag{9.1.10}$$

With this identification, the evolution of the probability density can be written in terms of L_C:

$$\rho_\theta(\theta, z) = \frac{1}{\sqrt{4\pi z/L_C}} \exp\left(-\frac{\theta^2}{4z/L_C}\right) \tag{9.1.11}$$

This distribution represents a diffusion of the angle θ with fiber length (Fig. 9.3). At an initial position the angle is known with certainty; propagating away from this point increases the uncertainty of the angle.

As the fiber length increases, the diffusion of θ approaches a uniform distribution. In fact, only the angle θ modulo π matters. Consider a starting angle of $\theta = 0$; at some distance z/L_C the density at $\theta = \pi/2$ will be within $1 - \varepsilon$ of uniform. Accounting for the wrap-around of the distribution modulo π, the length needed to fall within this error is $z/L_C = \pi^2/(8\varepsilon)$. For instance, it takes about $3L_C$ to realize a 5% deviation from a uniform distribution. Once the distribution is uniform there is no memory of the initial state.

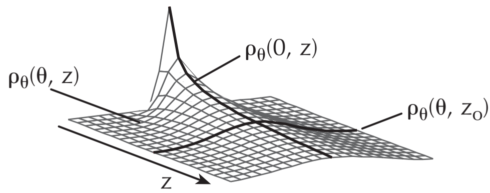

Fig. 9.3. Spatial evolution of the birefringent-angle θ probability density. At any position z_0 the density is a gaussian. On a length scale $z > L_C$ the density converges to a uniform distribution on $[-\pi/2, \pi/2]$.

9.1.2 Random Component Birefringence

For the second model, the entries of the birefringence matrix (9.1.6) are treated as independent Langevin processes:

$$\frac{d\beta_1}{dz} = -L_C^{-1}\beta_1 + g_1(z) \tag{9.1.12a}$$

$$\frac{d\beta_2}{dz} = -L_C^{-1}\beta_2 + g_2(z) \tag{9.1.12b}$$

The characteristics of the noise sources are

$$\langle g_1(z)\rangle = \langle g_2(z)\rangle = 0 \quad \langle g_1(z)g_2(z)\rangle = 0 \tag{9.1.13a}$$

$$\langle g_1(z)g_1(z')\rangle = \langle g_2(z)g_2(z')\rangle = \left\langle \beta^2 \right\rangle / L_C\ \delta(z-z') \tag{9.1.13b}$$

where $\left\langle \beta^2 \right\rangle$ is the rms value of the birefringence.

A Langevin equation describes a mean-reverting process driven by noise. Consider (9.1.12a) in the absence of $g_1(z)$: any initial condition decays exponentially over characteristic length L_C. Turning the noise source on disrupts this deterministic motion. The minus sign of the deterministic coefficient $-L_C^{-1}$ acts as a linear spring constant, pulling the solution back toward the mean with a strength proportional to the deviation. In the steady state, this spring force balances the noise source, resulting in a stationary gaussian distribution.

The solution to (9.1.12) is

$$\beta_i(z) = \beta_i(0)\,\mathrm{e}^{-z/L_C} + \int_0^z \mathrm{e}^{-(z-z')/L_C} g_i(z')dz'$$

The initial condition $\beta_i(0)$ decays exponentially on a scale given by the fiber autocorrelation length L_C. In the regime $z \gg L_C$ there is no memory of the initial state and a stationary distribution is reached. A two-dimensional sample-path of the birefringence is illustrated in Fig. 9.4. The steady-state density of β_i is readily determined by solution of the associated Fokker-Planck equation, and is

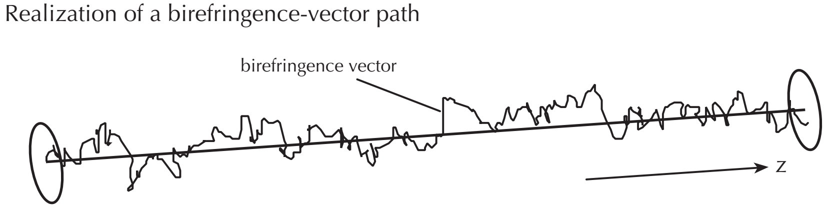

Fig. 9.4. Sample path of the birefringence vector in the steady-state. This path was calculated using a Karhunen-Loeve expansion of a Wiener process and a numerical integration of the Langevin equation (9.1.12). In the steady-state $\beta_{1,2}$ converge in distribution to i.i.d. stationary gaussian processes.

$$\rho_{\beta_i}(v) = \frac{1}{\sqrt{\pi \left\langle \beta^2 \right\rangle}} \exp\left(-\frac{v^2}{\left\langle \beta^2 \right\rangle}\right) \tag{9.1.14}$$

The variance of each component is $\mathrm{var}\,(\beta_i) = \left\langle \beta^2 \right\rangle /2$, which is independent of z. Moreover, as detailed in Appendix D, the radial and angular distributions of the local birefringence vector $\vec{\beta} = \hat{x}\beta_1 + \hat{y}\beta_2$ are Rayleigh and uniform distributions, respectively. Finally, the second moment of the Rayleigh distribution is $\left\langle \beta^2 \right\rangle$, which is what is expected on physical grounds. Therefore one writes

$$\mathrm{var}\,(\beta_i) = \frac{1}{2}\,\mathrm{var}\left(|\vec{\beta}|\right) \tag{9.1.15}$$

This and the preceding section have detailed physically reasonable forms of the local fiber birefringence vector $\vec{\beta}$ that drives the evolution of the polarization state and, consequently, the PMD evolution. A significant further study by Marcuse *et al.* details how these models are used to analyze pulse propagation, and particularly non-linear propagation, in fibers [36].

9.2 Polarization Diffusion in Single-Mode Fiber

The optical field evolves along a single-mode fiber subject to local birefringence perturbations. The characteristic length L_E is the length-scale over which the electric field losses memory and is called the polarization decorrelation length [60]. This is an additional length scale to those characteristic of the fiber, namely, the birefringent beat length L_B and the fiber autocorrelation length L_C.

One naturally expects the polarization decorrelation length to be related to the birefringent properties of the fiber. At one extreme, where $L_C \ll L_B$, the optical field averages over the rapidly varying birefringence, thereby changing slowly with respect to its initial state. At the other extreme where $L_B \ll L_C$, the field tends to follow the local birefringence and will, accordingly, change slowly with respect to the local state. Based on this physical picture, there are two choices for the polarization decorrelation length: a fixed definition $L_{E,\mathrm{fixed}}$

that measures the local field with respect to the initial birefringence, and a local definition $L_{E,\mathrm{local}}$ that measures the local birefringence.

To compute the polarization decorrelation length, the Stokes picture of polarization diffusion (9.1.5) is used. Recalling that the fiber model assumes no chirality, the component form of the precession equation reads

$$\frac{d}{dz}\begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} & & \beta_2 \\ & & -\beta_1 \\ -\beta_2 & \beta_1 & \end{pmatrix}\begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} \beta_2 S_3 \\ -\beta_1 S_3 \\ -\beta_2 S_1 + \beta_1 S_2 \end{pmatrix}$$

where capital S_k denotes a random variable. In order to simplify this expression prior to writing the diffusion generator, Wai and Menyuk introduce a rotation operator $R(z)$ to rotate the local birefringence to point along s_1 [62]. This operator is simple because $\beta_3 = 0$:

$$R(z) = \begin{pmatrix} \cos\theta(z) & -\sin\theta(z) & 0 \\ \sin\theta(z) & \cos\theta(z) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.2.1}$$

By defining the local Stokes coordinates such that $\tilde{s} = R(z)\hat{s}$, the precession equation (9.1.5) is transformed to

$$\frac{d}{dz}\,\tilde{s} = \left(R\vec{\beta} \times R^{-1}\right)\tilde{s} - RR_z^{-1}\tilde{s} \tag{9.2.2}$$

where R_z is the derivative of $R(z)$ with respect to z. This precession equation is called the local evolution equation, to distinguish it from (9.1.5) which describes fixed-reference evolution.

There are two derivations that can follow from (9.2.2) – the first using the fixed-birefringence fiber model, and the second using Langevin fiber model. The results of the two calculations are not qualitatively different; so the first, and simpler, model is detailed below.

For the first fiber model, the birefringent variation (9.1.7) and its noise source (9.1.8) is substituted into the local evolution equation. This gives

$$\frac{d}{dz}\begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \tilde{S}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\beta\tilde{S}_3 \\ \beta\tilde{S}_2 \end{pmatrix} + \begin{pmatrix} \tilde{S}_2 \\ -\tilde{S}_1 \\ 0 \end{pmatrix} g_\theta \tag{9.2.3}$$

This is a stochastic differential equation (SDE) in the Itô form. While it is quite beyond the scope of this text, the discerning reader should be aware that this equation is not quite correct. As Foschini shows [17], the Itô interpretation leads to an immediate departure of the polarization state from the unit sphere once (9.2.3) is integrated, while the associated Stratonovich interpretation does not. One has two choices on how to treat the analytics. Either the infinitesimal probability generator G accounts for the Stratonovich interpretation by adding a correction term or the SDE is translated from Stratonovich to Itô form and the Itô generator is used. Examples of both treatments are given below.

The infinitesimal probability generator governs the diffusion of the probability density associated with the stochastic differential equation

$$dX_{i,z} = b(z, X_{i,z})dz + \sigma(z, X_{i,z})dB_{i,z} \tag{9.2.4}$$

where b is the column-vector coefficient of the drift and σ is the column vector coefficient of the Brownian motion – either are functions of length and the random variable. Brownian motion is related to white noise by $dB_z = g_z dz$. The expectation of a sufficiently smooth functional ψ on coordinates X_i evolves according to

$$\frac{d\langle\psi\rangle}{dz} = \langle G\psi\rangle \tag{9.2.5}$$

This is Kolmogorov's backward equation (KBE). The generator G is the probability generator of the Itô diffusion (see Foschini [17], Menyuk [62], Oksendal [44] (esp. Theorem 7.3.3, and (6.1.3)), and Risken [55] for more details). The importance of the probability generator is that it transforms a stochastic differential equation into a partial differential equation (PDE). Many powerful analytic and numeric tools are available to solve PDEs, making problems cast in this form more tractable. The polarization and PMD diffusions that follow are prime examples of physical processes developed first in SDE form to capture the differential behavior of the process and then solved in PDE form to determine the global behavior subject to the boundary conditions.

The Itô-sense diffusion generator has two components: $G_I = G_d + G_s$. These components account for, respectively, the deterministic drift of the system and the stochastic fluctuation. The Itô generator for (9.2.4) is

$$G_I = \sum_i b_i(x_i)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_i\sum_j \left(\sigma\sigma^T\right)_{i,j}(x)\,\frac{\partial^2}{\partial x_i \partial x_j} \tag{9.2.6}$$

The Stratonovich-sense G_S generator makes a correction for the drift and, while more complicated to express in general, for the present case it may be written as

$$G_S = G_I - \frac{1}{2}\,\sigma^2 \sum_i x_i\,\frac{\partial}{\partial x_i} \tag{9.2.7}$$

The calculations in this section use the Stratonovich-sense generator.

While the derivations of the above equations are advanced, its application is straightforward. The generator for (9.2.3) is

$$\begin{aligned} G = &-\beta\tilde{S}_3\frac{\partial}{\partial\tilde{S}_2} + \beta\tilde{S}_2\frac{\partial}{\partial\tilde{S}_3} \\ &+ \frac{\sigma_\theta^2}{2}\left(\tilde{S}_2^2\frac{\partial^2}{\partial\tilde{S}_1^2} + \tilde{S}_1^2\frac{\partial^2}{\partial\tilde{S}_2^2} - 2\tilde{S}_1\tilde{S}_2\frac{\partial^2}{\partial\tilde{S}_1\partial\tilde{S}_2} - \tilde{S}_1\frac{\partial}{\partial\tilde{S}_1} - \tilde{S}_2\frac{\partial}{\partial\tilde{S}_2}\right) \end{aligned} \tag{9.2.8}$$

where the noise strength σ_θ^2 is related to the fiber autocorrelation length via (9.1.10).

The evolution of the moments of $\tilde{S}$ are calculated using this generator and the KBE. For instance, the functionals $\psi(\tilde{S}) = \tilde{S}$ and $\psi(\tilde{S}) = \tilde{S}^2$ give the evolution of the first- and second-moments of $\tilde{S}$. These results are used to associate the polarization decorrelation length L_E with the fiber parameters.

For the evolution of the mean values, the functional ψ is $\psi(\tilde{S}_i) = \tilde{S}_i$. Calculation of $G\psi$ generates the following system of equations:

$$\frac{d}{dz}\left\langle \tilde{S}_1 \right\rangle = -\frac{1}{L_C}\left\langle \tilde{S}_1 \right\rangle \tag{9.2.9a}$$

$$\frac{d}{dz}\left\langle \tilde{S}_2 \right\rangle = -\frac{2\pi}{L_B}\left\langle \tilde{S}_3 \right\rangle - \frac{1}{L_C}\left\langle \tilde{S}_2 \right\rangle \tag{9.2.9b}$$

$$\frac{d}{dz}\left\langle \tilde{S}_3 \right\rangle = \frac{2\pi}{L_B}\left\langle \tilde{S}_2 \right\rangle \tag{9.2.9c}$$

For a non-zero initial condition, $\left\langle \tilde{S}_1(z) \right\rangle$ monotonically decays to zero and the remaining mean values undergo a damped oscillation to zero. The long-range values of the polarimetric means are all zero; the polarization state with respect to the local birefringence ultimately becomes completely uncorrelated. In the particular case when the initial state of the system is $\tilde{S}_1 = 1$, that is, the launch polarization is aligned to the local birefringent axis, the mean polarimetric values of the remaining two coordinates are $\left\langle \tilde{S}_2(z) \right\rangle = \left\langle \tilde{S}_3(z) \right\rangle = 0$, and the mean along the initial axis decays as

$$\left\langle \tilde{S}_1(z) \right\rangle = \exp\left(-z/L_C\right) \tag{9.2.10}$$

This simple case is all that is necessary to associate the polarization decorrelation length $L_{E,\text{local}}$ with the fiber autocorrelation length. Since the characteristic length over which the mean polarization is preserved is by definition the polarization decorrelation length, in light of (9.2.10) one makes the association

$$L_{E,\text{local}} = L_C \tag{9.2.11}$$

In the local reference frame the two characteristic lengths are equal.

Wai and Menyuk detail the transformation to the fixed reference frame from the local frame [62]. While their work may be consulted for the details, the results are

$$\langle S_1(z) \rangle = \begin{cases} S_1(0)\exp\left(-z/L'_{E,\text{fixed}}\right) & L_C \ll L_B \\ S_1(0)\exp\left(-z/L''_{E,\text{fixed}}\right) & L_C \gg L_B \end{cases} \tag{9.2.12}$$

where

$$L'_{E,\text{fixed}} = \frac{L_C}{2\pi^2 \left(L_C/L_B\right)^2} \tag{9.2.13a}$$

$$L''_{E,\text{fixed}} = L_C/\, 2 \tag{9.2.13b}$$

These equations reenforce the physical understanding developed in the introduction of this chapter. With respect to the launched polarization state, when $L_C \ll L_B$ the polarization decorrelation length (9.2.13a) is much longer than the fiber autocorrelation length. This is because the birefringence changes too rapidly for the field to follow, which in turn makes the propagated field correlate with the launched field over a longer distance. Conversely, when $L_C \gg L_B$, the field follows the birefringence more faithfully, so the polarization state diffuses on a length scale more closely linked to the fiber autocorrelation length. In particular, notice that $L''_{E,\text{fixed}} = L_{E,\text{local}}/\, 2$, which makes sense because the field follows the local birefringence, so the local-frame characteristic length should indeed be longer than the fixed reference frame.

These associations between the fiber autocorrelation and polarization decorrelation lengths are useful in a practical sense. While L_C is central to the statistical description of the optical field, the polarization decorrelation length is the measurable quantity. Equation (9.2.12) provides a means in which to determine L_C through the measurement of $L_{E,\text{fixed}}$.

For the evolution of the polarimetric second-momemts, the functional ψ is set to $\psi(\tilde{S}_i) = \tilde{S}_i^2$ and the generator (9.2.8) remains the same. Calculation of $G\psi$ generates the following system of equations:

$$\begin{aligned}
\frac{d}{dz}\left\langle \tilde{S}_1^2 \right\rangle &= -\frac{2}{L_C}\left(\left\langle \tilde{S}_1^2 \right\rangle - \left\langle \tilde{S}_2^2 \right\rangle\right) \\
\frac{d}{dz}\left\langle \tilde{S}_2^2 \right\rangle &= \frac{2}{L_C}\left(\left\langle \tilde{S}_1^2 \right\rangle - \left\langle \tilde{S}_2^2 \right\rangle\right) - \frac{4\pi}{L_B}\left\langle \tilde{S}_2\tilde{S}_3 \right\rangle \\
\frac{d}{dz}\left\langle \tilde{S}_3^2 \right\rangle &= \frac{4\pi}{L_B}\left\langle \tilde{S}_2\tilde{S}_3 \right\rangle \\
\frac{d}{dz}\left\langle \tilde{S}_2\tilde{S}_3 \right\rangle &= \frac{2\pi}{L_B}\left(\left\langle \tilde{S}_2^2 \right\rangle - \left\langle \tilde{S}_3^2 \right\rangle\right) - \frac{2}{L_C}\left\langle \tilde{S}_2\tilde{S}_3 \right\rangle
\end{aligned} \tag{9.2.14}$$

As before there are local- and fixed-reference frame solutions, and the solutions differ for the two limits of L_C/L_B. The general form for the fixed-frame solution is

$$\left\langle S_{1,2}^2 \right\rangle \simeq \frac{1}{3}\left(1 \pm \frac{3}{2}\exp\left(-z/L_{E,1}\right) + \frac{1}{2}\exp\left(-z/L_{E,2}\right)\right) \tag{9.2.15a}$$

$$\left\langle S_3^2 \right\rangle \simeq \frac{1}{3}\left(1 - \exp\left(-z/L_{E,3}\right)\right) \tag{9.2.15b}$$

where in the first equation the "+" and "−" signs refer to $\tilde{S}_1$ and $\tilde{S}_2$, respectively. In the $L_C \ll L_B$ regime, the length scales are

$$L_{E,1} = 2L'_{E,\text{fixed}}, \quad \text{and} \quad L_{E,(2,3)} = 6L'_{E,\text{fixed}} \tag{9.2.16}$$

Conversely, in the $L_C \gg L_B$ regime, the length scales are

$$L_{E,1} = \tfrac{2}{7} L''_{E,\mathrm{fixed}}, \quad \text{and} \quad L_{E,(2,3)} = \tfrac{2}{3} L''_{E,\mathrm{fixed}} \tag{9.2.17}$$

In either length regime, the stationary variances of the diffusions converge to $\left\langle \tilde{S}_k^2 \right\rangle = 1/3$. The convergence rate for all three variances is roughly the same, but the small difference was studied by Wai and Menyuk in [61], where they showed that the absence of chirality in the fiber imparts a short-range anisotropy to the diffusions.

To summarize, polarization decorrelation happens with its own characteristic length-scale L_E in a single-mode fiber. That length scale is related to the fiber birefringence parameters in either a local or fixed frame of reference. In the local frame, the polarization decorrelation and fiber autocorrelation lengths are equal. In the fixed frame, the relationship depends on the regime in which the fiber is characterized: for $L_C \ll L_B$ the diffusion occurs at a rate related to L_B; for $L_C \gg L_B$ the diffusion occurs at a rate related to L_C. The three polarimetric values all reach a mean of zero and a variance of 1/3 beyond the diffusion limit. The diffusions detailed in this section are for the fixed-birefringence fiber model, but the results do not qualitatively change for the Rayleigh-distributed birefringence model.

9.3 RMS Differential-Group Delay Evolution

The equation of motion for polarization evolution (9.1.5) was recast in the preceding section into a stochastic differential equation whose solutions showed the behavior of the statistical moments of the polarization state. In parallel with this procedure, the equation of motion for polarization-mode dispersion evolution is studied. Recall from (8.2.39) on page 339 that the differential equation of motion for the PMD vector $\vec{\tau}$ is

$$\frac{\partial \vec{\tau}}{\partial z} = \vec{\beta}_\omega + \vec{\beta} \times \vec{\tau} \tag{9.3.1}$$

where $\vec{\beta}$ is the local birefringence vector and $\vec{\beta}_\omega$ is its frequency derivative.

The solution to (9.3.1) for the mean-square magnitude of $\vec{\tau}$ in the diffusion limit is

$$\left\langle \tau^2(z) \right\rangle = 2 \left\langle \tau_c^2 \right\rangle \left(\mathrm{e}^{-z/L_C} + z/L_C - 1 \right) \tag{9.3.2}$$

where $\left\langle \tau_c^2 \right\rangle$ is the mean-square DGD for a segment L_C long. The mean-square solution is written at this point in the discussion because it is apparently independent of any reasonable derivation. Poole [48, 49] first derived this equation, shortly followed by Curti [7], Foschini [17], and Gisin [23, 24], and later by Wai and Menyuk [60]. Gisin [24] showed that (9.3.2) is the mean-square deviation of the probability density that solves the Telegrapher's equation (a

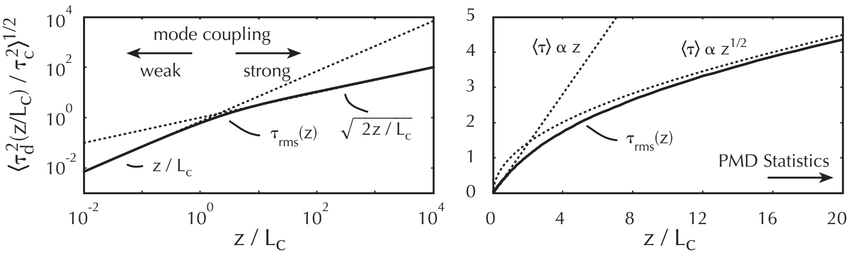

Fig. 9.5. The rms growth of $\langle \tau \rangle$ with length and its asymptotic limits. a) Log-log scale shows long-range behavior. For $z \ll L_C$ the rms growth is linear with length, while for $z \gg L_C$ the rms growth goes as root-length. The crossover is in the range $z \sim L_C$. b) Linear scale of the same. PMD fiber statistics are derived in the strong mode-coupling regime.

second-order parabolic PDE). While the details depend on which report is consulted, all of these researchers apply a gaussian-correlated noise to the fiber birefringence.

The root-mean-square growth of the DGD, defined by $\langle \tau \rangle_{\text{rms}} = \sqrt{\langle \tau^2(z) \rangle}$, shows two asymptotic limits:

$$\langle \tau \rangle_{\text{rms}}(z) = \begin{cases} (z/L_C)\, \langle \tau_c \rangle_{\text{rms}} & \text{weak coupling: } z \ll L_C \\ \sqrt{2z/L_C}\, \langle \tau_c \rangle_{\text{rms}} & \text{strong coupling: } z \gg L_C \end{cases} \tag{9.3.3}$$

where $\langle \tau_c \rangle_{\text{rms}} = \sqrt{\langle \tau_c^2 \rangle}$. The cross-over between these two limits is $z \sim L_C$. The two limits of the rms evolution function are shown in Fig. 9.5.

At one limit, the so-called "weak" mode-coupling limit, $\langle \tau \rangle_{\text{rms}}$ grows linearly with length z. In this regime the local birefringence is nearly aligned along the fiber, making the cumulative birefringence additive. At the other limit, the "strong" mode-coupling limit, $\langle \tau \rangle_{\text{rms}}$ grows as root-length: $z^{1/2}$. In this regime the fiber length is well beyond the fiber autocorrelation length, so the cumulative birefringence grows statistically. Given that the birefringence of fiber segments L_C long are uncorrelated, the birefringence variances across segments add, resulting in a standard deviation that grows as root-length. This behavior was already seen in the probability density (9.1.11) for the angular distribution of birefringence.

The analytic calculation of the PMD probability densities are made in the strong-coupling regime. The weak-coupling regime poses few novel problems because the statistics will closely match that of the local birefringence. The intermediate region that connects the weak- and strong-coupling regimes can be treated one of two ways. For the case of fiber the PMD statistics evolve through the intermediate region toward their stationary end-points. Tan *et al.* have studied these transient statics and show that the Maxwellian distribution is achieved by approximately $z \sim 30L_C$, but that the distribution tails take longer to mature [58, 63]. Another case is for PMD emulators, where a fixed number (3–20) of sections is used. These statistics have also been investigated and shown to exhibit deviation from the stationary forms [30, 34].

Equation (9.3.2) is derived here following the diffusion formalism. Use of the fixed-birefringence model of §9.1.1 makes for a simpler calculation; the result is easily extended to Rayleigh-distributed birefringence. As with the polarization calculations, the PMD diffusion equation is simpler when converted to a local reference frame. Define $\tilde{\tau} = R(z)\vec{\tau}$, where $R(z)$ is as in (9.2.1). The resulting stochastic differential equation for (9.3.1) in component form is (cf. (9.2.3))

$$\frac{d}{dz}\begin{pmatrix}\tilde{\tau}_1\\ \tilde{\tau}_2\\ \tilde{\tau}_3\end{pmatrix} = \begin{pmatrix}\beta_\omega\\ -\beta\tilde{\tau}_3\\ \beta\tilde{\tau}_2\end{pmatrix} + \begin{pmatrix}\tilde{\tau}_2\\ -\tilde{\tau}_1\\ 0\end{pmatrix} g_\theta \tag{9.3.4}$$

The infinitesimal probability generator is

$$\begin{aligned} G = \beta_\omega \frac{\partial}{\partial\tilde{\tau}_1} - \beta\tilde{\tau}_3\frac{\partial}{\partial\tilde{\tau}_2} + \beta\tilde{\tau}_2\frac{\partial}{\partial\tilde{\tau}_3} \\ + \frac{\sigma_\theta^2}{2}\left(\tilde{\tau}_2^2\frac{\partial^2}{\partial\tilde{\tau}_1^2} + \tilde{\tau}_1^2\frac{\partial^2}{\partial\tilde{\tau}_2^2} - 2\tilde{\tau}_1\tilde{\tau}_2\frac{\partial^2}{\partial\tilde{\tau}_1\partial\tilde{\tau}_2} - \tilde{\tau}_1\frac{\partial}{\partial\tilde{\tau}_1} - \tilde{\tau}_2\frac{\partial}{\partial\tilde{\tau}_2}\right) \end{aligned} \tag{9.3.5}$$

Finally, the functional of interest is $\psi = \tilde{\tau}^2 = \tilde{\tau}_1^2 + \tilde{\tau}_2^2 + \tilde{\tau}_3^2$. A requisite auxiliary functional $\psi' = \tilde{\tau}_1$ is also needed. Calculating $G\psi$ and $G\psi'$, and combining the results, yields the equation of motion

$$\left(\frac{d^2}{dz^2} + \frac{1}{L_C}\frac{d}{dz}\right)\left\langle\tilde{\tau}^2\right\rangle = 2\beta_\omega^2 \tag{9.3.6}$$

The solution to this equation is (9.3.1). A few details need clarification. The length of the PMD vector is invariant under rotation, so $|\tilde{\tau}| = |R(z)\vec{\tau}|$. The product $\beta_\omega L_C$ is the characteristic DGD per segment L_C long. To see this, simply expand the terms: $\tau_c = \beta_\omega L_C = \Delta n_{\mathrm{g}} L_C / c$. Finally, the replacement of τ_c^2 in the solution of (9.3.6) with $\left\langle\tau_c^2\right\rangle$ as reported in (9.3.2) comes with the Rayleigh-birefringence derivation detailed by Wai and Menyuk [62].

9.4 PMD Statistics

The PMD statistics that yield to analytic study are those related to the first- and second-order PMD vectors $(\vec{\tau}, \vec{\tau}_\omega)$ and the autocorrelation of the PMD vector $\vec{\tau}$ with frequency. The statistics that have been analytically solved are for $\vec{\tau}$ and its components τ_i; $\vec{\tau}_\omega$ and its components $\tau_{\omega,i}$; and the perpendicular and parallel components of the second-order vector: depolarization $|\vec{\tau}_{\omega,\perp}|$ and polarization-dependent chromatic dispersion $|\vec{\tau}|_\omega$, respectively. Additionally, the relation between depolarization and PDCD conditional on the DGD has been determined. The joint density of the magnitudes (τ, τ_ω), however, is not analytic but has been solved using importance sampling (IS) and, separately, special numerical techniques.

A remarkable property of PMD statistics is that they scale with a single scaling factor: the mean fiber DGD $\bar{\tau}$. The mean fiber DGD is itself related to the fiber length, fiber autocorrelation length, and birefringence variance. That association is clarified below, but once made the mean fiber DGD becomes its own "unit." The mean fiber DGD is so ubiquitous in the statistics and measurement of PMD that a corruption of terms has entered the literature where "PMD" is *defined* as the mean fiber DGD. There should be no confusion, however, between PMD as a vector and mean DGD as a statistical unit.

The expression for the rms DGD evolution in the strong coupling limit connects the microscopic scale of the birefringence variation with the macroscopic properties of PMD statistics. This expression is therefore the gateway between micro- and macroscopic views of the same process. While this expression was derived above, the derivation of the PMD probability densities requires analytic tools that are well beyond the scope of this text. Instead, the results, principally of Foschini, will be quoted and the ambitious reader is referred to the cited papers.

The mean fiber DGD is connected to the stochastic model of PMD evolution in the following way. In the $z \gg L_C$ limit, the mean-square DGD grows as

$$\left\langle \tau^2(z) \right\rangle = 2 \left\langle \tau_c^2 \right\rangle z / L_C \tag{9.4.1}$$

As discussed in the following, the cartesian components of the PMD vector are i.i.d. gaussian random variables. The probability density of the length of the PMD vector is therefore Maxwellian. The first and second moments of this density are related (see Table D.1 on page 507), so (9.4.1) may be rewritten for the mean DGD:

$$\bar{\tau} \equiv \langle \tau(z) \rangle = \sqrt{\frac{8}{3\pi} \frac{2 \left\langle \tau_c^2 \right\rangle z}{L_C}} \tag{9.4.2}$$

There is some variation in the literature on the interpretation of $2\left\langle \tau_c^2 \right\rangle / L_C$. Curti, for instance, writes $\bar{\tau} = \sqrt{8z/\pi L_C}\, d\tau$ [7]. Since (9.3.2) was derived using a Rayleigh statistic for the birefringence, the mean segment DGD is related to its second moment by $\left\langle \tau_c^2 \right\rangle = 4 \left\langle \tau_c \right\rangle^2 / \pi$. Association with Curti gives $d\tau = \sqrt{8/3\pi}\, \langle \tau_c \rangle$, which is an 8.5% difference. Separately, Poole and Favin [52] write $\bar{\tau} = \sqrt{8N/3\pi} \Delta\tau_p$, where $\Delta\tau_p$ is the fixed birefringence of a retardation plate and N is the number of plates. The connection to (9.4.2) requires $N = 2z/L_C$. This interpretation is used below to develop a discrete waveplate model. Other researchers reproduce (9.4.2) in the continuous limit [17, 23, 51, 62].

Armed with the definition of the mean fiber DGD $\bar{\tau}$, the statistics of PMD are presented below. The principal contributors to this field are Curti [7], who first derived the Maxwellian DGD distribution; Foschini [14–17], who derived the remaining PMD distributions; Karlsson [31], and Shtaif and

Mecozzi [56, 57], who derived the PMD autocorrelation functions; Ibragimov and Shtengel [28], who derived a conditional expression; and Fogal, Biondini, and Kath, who developed the IS methods [2, 11, 12].

9.4.1 Probability Densities

The expressions for the probability densities of the first- and second-order PMD vector are tabulated in Table 9.1. The vectors $\vec{\tau}$ and $\vec{\tau}_\omega$ are statistically dependent on one another. Plots of these densities are shown in Fig. 9.6 on linear scale, to emphasize the distribution about the mean, and semi-log scale, to emphasize the fall-off of the distribution tails.

The probability density for the DGD τ, which is the magnitude of the PMD vector $\tau = |\vec{\tau}|$, is Maxwellian. The origin of the Maxwellian distribution comes from the distribution of the radius of a sphere, where the cartesian coordinates of the sphere are i.i.d. gaussian random variables (see Appendix D). An important relation between the mean and mean-square for a Maxwellian distribution is

$$\left\langle \bar{\tau}^2 \right\rangle = \left(\frac{8}{3\pi}\right) \left\langle \bar{\tau} \right\rangle^2 \tag{9.4.3}$$

The $8/3\pi$ factor appears frequently in the discussion of PMD statistics. The Maxwellian and gaussian distributions scale linearly in $\bar{\tau}$. That is, the measured DGD values from any fiber can be normalized by the mean fiber DGD $(\tau/\bar{\tau})$ to produce a unit-scaled statistic.

The probability density for the magnitude SOPMD $\tau_\omega = |\vec{\tau}_\omega|$ is sech-tanh in form. The origin of the sech-tanh distribution comes from the distribution of the radius of a sphere, where the cartesian coordinates of the sphere are i.i.d. hyperbolic secant (sech) random variables. In contrast to the Maxwellian and gaussian distributions, the sech-tanh and sech distributions scale quadratically with $\bar{\tau}^2$: measurements can be normalized to a unit statistic by $\tau_\omega/\bar{\tau}^2$.

The second moments of the τ and τ_ω magnitudes are related by

$$\left\langle \tau_\omega^2 \right\rangle = \frac{1}{3} \left\langle \bar{\tau}^2 \right\rangle^2 \tag{9.4.4}$$

Moreover, a comparison of the tails of the Maxwellian and sech-tanh distributions (Fig. 9.6), shows that eventually the Maxwellian falls off faster. This is because the underlying gaussian distribution falls off quadratically (on a log scale) while that of the sech fall off linearly.

The cartesian components of the SOPMD vector are i.i.d. random variables, so their variance is one-third that of the vector length. Yet, the cartesian components are not directly or easily related to the distortion PMD imparts on a signal. However, the projections of the SOPMD vector onto the first-order vector are directly related, so one asks about the conditional dependence of these projections.

Statistic
DGD$^{(a--d}$
SOPMD$^{(d)}$
$\vec{\tau}$ compone
$\vec{\tau}_\omega$ compon
PDCD$^{(e,f)}$
Depolariza
Depolariza vector$^{(g)}$

$^{(*)}$ $\left\langle \bar{\tau}^2 \right\rangle =$

$^{(a)}$ Curti *e*

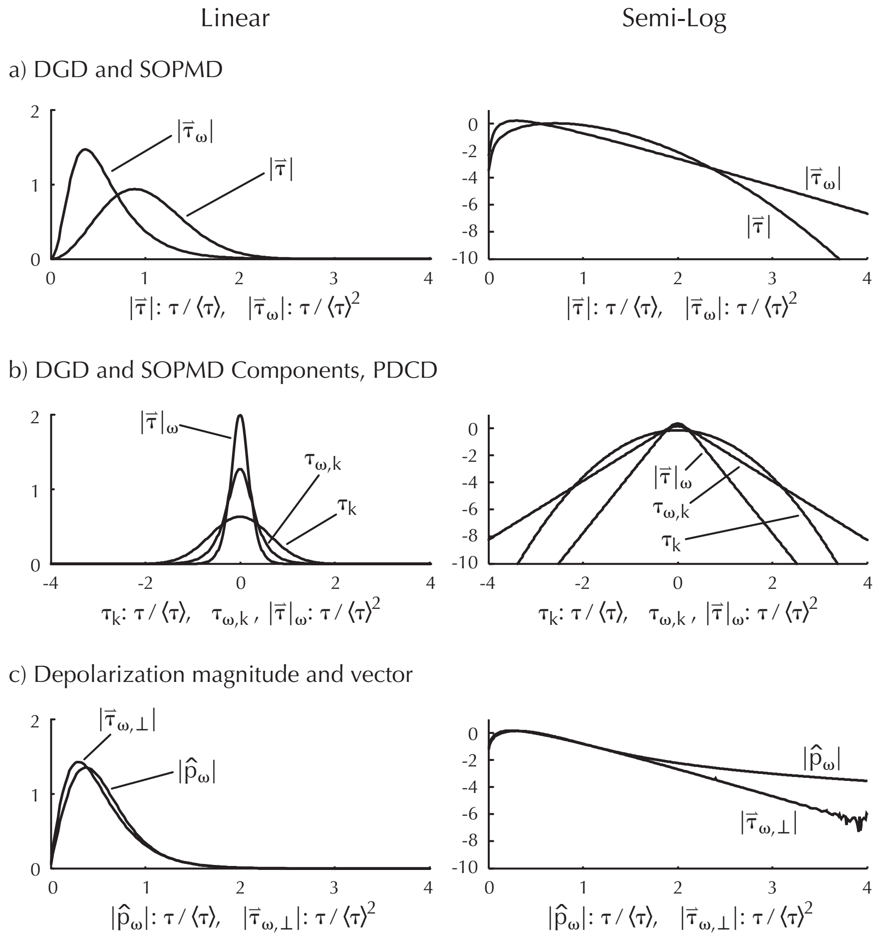

Fig. 9.6. Probability densities of first- and second-order PMD statistics, linear and semi-log scales. The log scale is in $\log_{10}$.

The second-order PMD vector is projected onto the direction of the first-order vector $(\vec{\tau}_\omega \cdot \hat{p})$ to produce parallel and perpendicular components. This action conditions the SOPMD vector to $\hat{p}$. Expanding $\vec{\tau}_\omega$ from $\vec{\tau} = \tau\hat{p}$ makes

$$\vec{\tau}_\omega = \tau_\omega \hat{p} + \tau \hat{p}_\omega \tag{9.4.5a}$$

$$= \vec{\tau}_{\omega,\parallel} + \vec{\tau}_{\omega,\perp} \tag{9.4.5b}$$

The parallel component is the polarization-dependent chromatic dispersion. This component changes the chromatic dispersion of the fiber from D to $D_{\text{eff}} = D \pm \tau_{\omega,\parallel}$ and, accordingly, induces pulse compression or expansion [15, 50]. Nelson gives the wavelength dependence of this component [42]. The PDCD magnitude has two synonymous notations: $|\vec{\tau}|_\omega = \tau_{\omega,\parallel}$. The perpen-

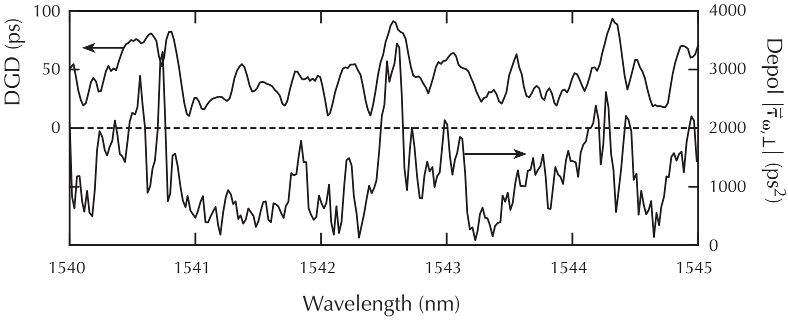

Fig. 9.7. Measurements of DGD τ and depolarization component $|\vec{\tau}_{\omega,\perp}|$ of SOPMD from high-PMD fiber spool. Courtesy G. Shtengel [28].

dicular component is the depolarization, which is the tendency for the PMD vector to change direction. The effects of this component were treated in §8.2.

There is a strong tendency for $\vec{\tau}_\omega$ to point away from $\vec{\tau}$. The mean-square value of the PDCD and depolarization components in relation to that of the second-order vector are

$$\left\langle \tau^2_{\omega,\parallel} \right\rangle = \frac{1}{9}\left\langle \tau^2_\omega \right\rangle, \quad \text{and} \quad \left\langle \tau^2_{\omega,\perp} \right\rangle = \frac{8}{9}\left\langle \tau^2_\omega \right\rangle \tag{9.4.6}$$

(Note that the rms of both components scale as $\bar{\tau}^2$, as does the full SOPMD vector.) Even in comparison to a cartesian component, the PDCD component is diminished:

$$\left\langle \tau^2_{\omega,\parallel} \right\rangle = \frac{1}{3}\left\langle \tau^2_{\omega,i} \right\rangle \tag{9.4.7}$$

Depolarization is clearly the dominant component of SOPMD and is therefore the dominant impairment on an optical signal.

One may ask how do the SOPMD-projected components vary with a particular sample value of DGD for a fixed mean DGD. This question comes about when testing PMD compensators: when the DGD value is high, what form of SOPMD in a fiber can one expect? The answer is that the higher the DGD, the higher the expected depolarization. Ibragimov and Shtengel [28] show that the conditional expectations scale as

$$\left\langle \tau^2_{\omega,\parallel} | \tau \right\rangle = \frac{1}{9}\left\langle \tau^2_\omega \right\rangle \tag{9.4.8a}$$

$$\left\langle \tau^2_{\omega,\perp} | \tau \right\rangle = \frac{2}{9}\left(\tau^2 \sqrt{3\left\langle \tau^2_\omega \right\rangle} + \left\langle \tau^2_\omega \right\rangle \right) \tag{9.4.8b}$$

When the sample value τ^2 equals its mean-square value (9.4.4), the conditional expression (9.4.8b) reduces to (9.4.6). These expressions show that the rms PDCD magnitude is determined solely by the mean fiber DGD, while the rms depolarization magnitude scales with sample DGD. These relations are borne out in experiment. Figure 9.7 illustrates DGD and magnitude-depolarization

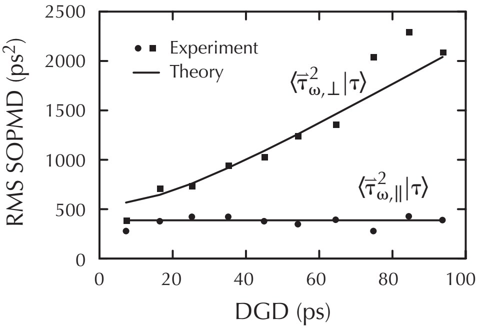

Fig. 9.8. Measurements of mean-square PDCD and depolarization as a function of DGD. Data taken on a single fiber with fixed mean DGD. Courtesy G. Shtengel [28].

measurements taken on a high-PMD fiber; the correlation is evident. Figure 9.8 shows an analysis of the data superimposed with the conditional distributions.

Finally, the joint probability distribution function (JPDF) of (τ, τ_ω) is a distribution central to the characterization and validation of receiver performance against PMD. The JPDF is not analytic, but can be calculated via brute force, special numerical techniques, or using importance sampling methods. A brute-force example of the JPDF is shown in Fig. 9.9. The lines indicate contours of constant probability. This calculation by P. J. Leo took a week running on a distributed computer network; the code formalism used the Stokes-based PMD concatenation rules for $\vec{\tau}$ and $\vec{\tau}_\omega$, thereby saving an order-of-magnitude in time compared to the equivalent Jones-matrix calculation. For comparison, a JPDF from measured field data is shown in Fig. 9.10.

The importance-sampling methods developed by Fogal, Biondini, and Kath [2, 3, 11, 12] also calculate the PMD vector in Stokes space, but with the innovation that a bias is added to the scattering of the polarization state from section to section. The bias can be made to accumulate preferentially large DGD values, large SOPMD values, or both. Importance sampling provides the formalism to rescale the resulting statistics by the probability of the underlying bias. Contours that extend to 10^{-20} have been demonstrated. Moreover, IS methods have been adopted to the expedient estimation of system outage probabilities due to PMD [33].

Finally, Forestieri has developed special numerical techniques to evaluate the PDCD component density and, subsequently, the joint first- and second-order density [13]. The author reports an efficient algorithm for fast evaluation of the JPDF, and shows that in comparison to the IS method the tails of the distribution fall more slowly.

The JPDF of (τ, τ_ω) is universal because of the scaling rules for DGD and SOPMD. In particular, the DGD and SOPMD axes scale as $\bar{\tau}$ and $\bar{\tau}^2$, respectively. Therefore, once the JPDF is calculated for some $\bar{\tau}$, it can be

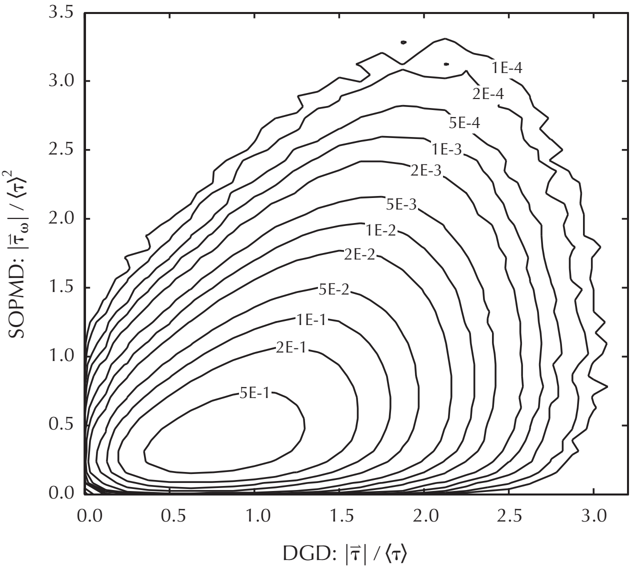

Fig. 9.9. Joint first- and second-order PMD density function. This is a universal distribution, scaled on the abscissa by $\langle\tau\rangle$ and on the ordinate by $\langle\tau\rangle^2$. Calculated from 10^9 realizations of a 2000-section fiber, using Stokes-based concatenation rules. Courtesy P. J. Leo.

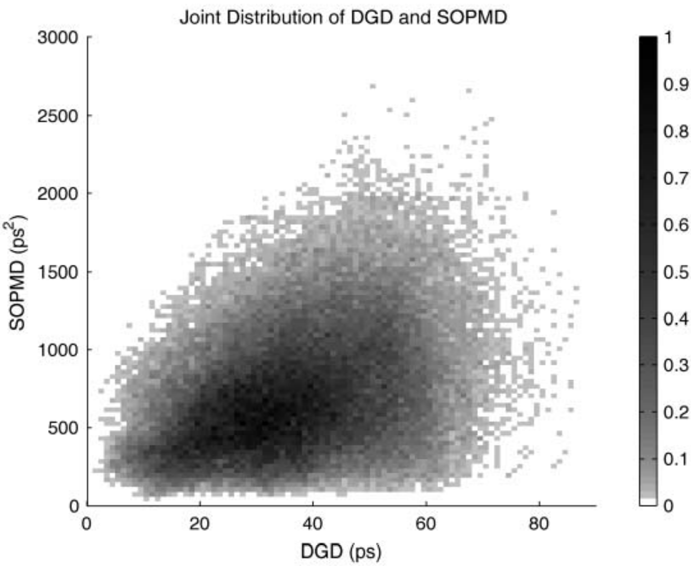

Fig. 9.10. Measured joint first- and second-order PMD magnitudes from fiber described in 9.11.

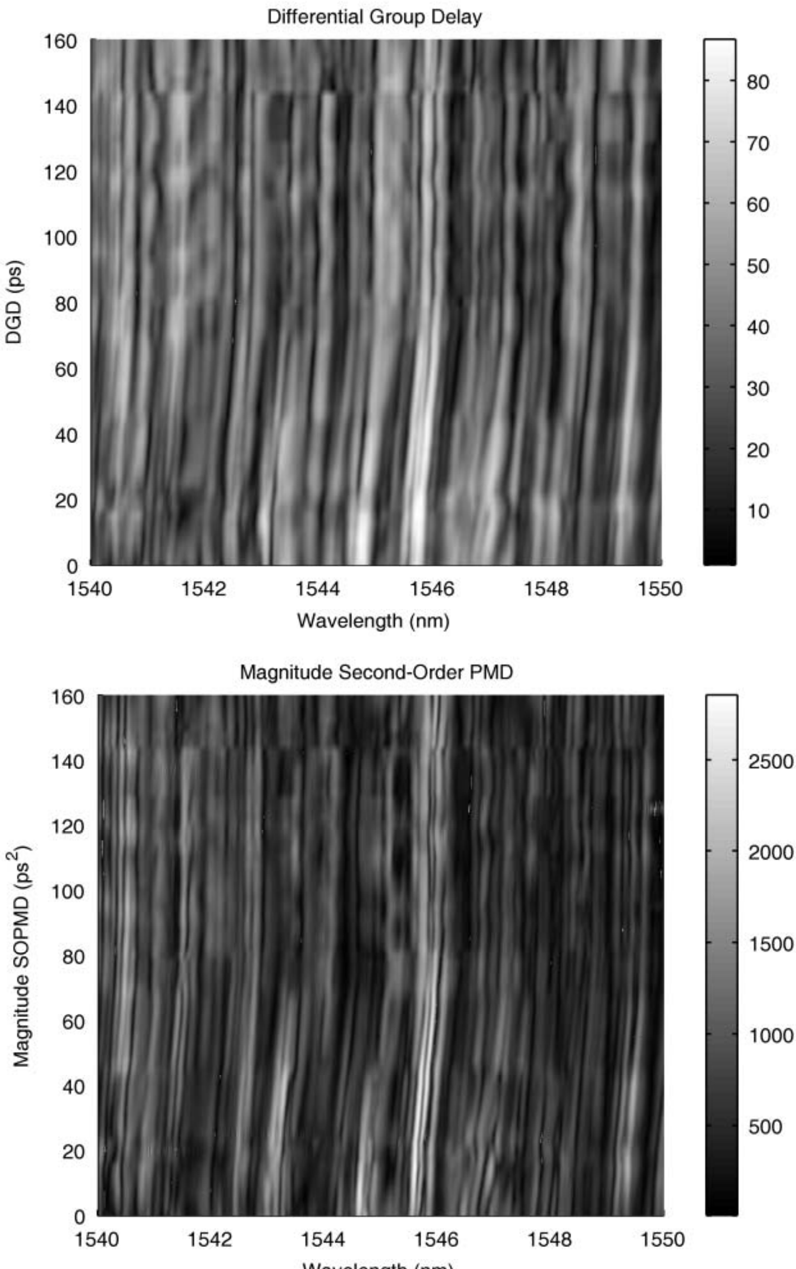

Fig. 9.11. Field measurements of differential-group delay and magnitude second-order PMD over 10 nm and 160 hrs. Notice the general adiabatic evolution is disrupted at about 145 hr. Courtesy D. Peterson, MCI [45].

arbitrarily rescaled to associate with any fiber. Such scaling, especially with the range available using IS, can be exploited to make solid predictions of the outage probability due to PMD in lightwave systems.

The nature of the JPDF also reveals the fallacy of the "PMD is DGD" concept in testing receivers and PMD compensators, whether optical or electrical. At almost any level of DGD there is nearly zero probability that SOPMD is zero, or even small. Yet most receiver testing at the time of this writing is done by introducing only DGD. Such results can significantly underestimate the receiver operation in real-world environments (but they glorify product performance to the parties that fund the work).

An example of rather adiabatic temporal behavior of first- and second-order PMD is shown in Fig. 9.11. The data was taken over 160 hrs in half-hour increments on installed fiber from a particularly old link [45]. At any time the wavelength variation is high, but the temporal evolution for this buried cable is slow. An exception exists at about 145 hrs where the fiber was disturbed (unintentionally). The disturbance must have been small because the spectra before and after the event are well correlated. A large disturbance will erase all memory of the past and set the fiber in a new state. This data also shows that for channels that fall on high first- and/or second-order states, the high PMD levels may persist for a long time before there is a change. A detailed study of the temperature dependence of PMD in installed cables is reported by Brodsky *et al.* [5].

9.4.2 Autocorrelation Functions

The autocorrelation function of the PMD vector is an essential descriptor for the characterization of PMD behavior. There are four properties that the autocorrelation function reveal:

- PMD is a wide-sense stationary process in frequency. The autocorrelation function is the only demonstrated method to connect ensemble averages with frequency averages.
- The PMD-vector correlation bandwidth depends only on the mean DGD $\bar{\tau}$.
- The variance of an estimated value of $\bar{\tau}$ derived from measurement depends inversely on the total measurement bandwidth.
- All moments of the PMD vector depend only on $\bar{\tau}$ and the moment order.

The autocorrelation function tells how large a frequency separation is necessary for two PMD vectors $\vec{\tau}(\omega)$ and $\vec{\tau}(\omega')$ to become statistically independent. Physically, PMD is generated by the fiber birefringence $\beta(z,\omega)$. The longitudinal evolution of the birefringence $\beta(z)$ is modelled as white noise or a concatenation of waveplates, while the frequency dependence $\beta(\omega)$ is linear: $\beta(\omega) = w\Delta n/c$. Recall from the PMD concatenation rules §8.2.4 that component PMD vectors precess about one another with frequency. For small frequency variation the output PMD vector retains its length and pointing

Table 9.2. Autocorrelation Functions for PMD Vector $\vec{\tau}$ and DGD Squared τ^2

Parameter	**Form**	**Expression**
PMD vector$^{(a,b)}$	$\langle \vec{\tau}_{\omega'} \cdot \vec{\tau}_{\omega} \rangle$	$R_{\vec{\tau}}(\mathcal{W}) = \mathrm{sinhc}(\mathcal{W}/6)\exp(-\mathcal{W}/6)$
DGD squared$^{(c)}$	$\left\langle \tau^2_{\omega'} \tau^2_{\omega} \right\rangle$	$R_{\tau^2}(\mathcal{W}) = \frac{3}{5}\left(1 + \frac{2}{3}\,\frac{(1 - R_{\vec{\tau}}(\mathcal{W}))}{\mathcal{W}/6}\right)$
Estimator uncertainty$^{(a,b,d,e)}$		$\bar{\tau}_{\mathrm{est}}(B_f) \simeq \bar{\tau}\left(1 \pm \frac{0.9}{\sqrt{B_f\,\bar{\tau}}}\right)$
Bias error of statistic moments$^{(f)}$		$\bar{\tau}(B_f) = \sqrt{\frac{8}{3\pi}\langle \bar{\tau}^2 \rangle} \pm \frac{8}{9\sqrt{2}}\frac{1}{2\pi B_f}$

$$\mathcal{W} = \Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle, \quad \Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle = \frac{3\pi^3}{2}\left(\Delta f\,\bar{\tau}\right)^2, \quad R_\tau = \sqrt{R_{\tau^2}}, \quad R_{\hat{\tau}} = R_{\vec{\tau}}/R_\tau$$

$$\Delta\omega_{\mathrm{PMD-ACF}}\sqrt{\langle \bar{\tau}^2 \rangle} \simeq 4.4, \quad \Delta f_{\mathrm{PMD-ACF}}\,\bar{\tau} \simeq \frac{2}{\pi}$$

Note: Terms such as τ_ω denote $\tau_\omega = \tau(\omega)$, not a frequency derivative.

$^{(a,c)}$ Shtaif and Mecozzi [39, 56, 57], $^{(b)}$ Karlsson and Brentel [31],

$^{(d)}$ Gisin *et al.* [22], $^{(e)}$ The Poole and Favin coefficient was 0.66 rather than 0.9 [52],

$^{(f)}$ Boroditsky *et al.* [4].

direction. As the frequency is further detuned, the PMD vector takes on a different length and pointing direction. The PMD autocorrelation function tells how large, on average, a frequency separation is necessary for two such PMD vectors to become statistically uncorrelated.

The normalized autocorrelation functions for the PMD vector $R_{\vec{\tau}}(\mathcal{W})$ and the DGD squared $R_{\tau^2}(\mathcal{W})$ are given in Table 9.2[1]. Both functions are even functions of the argument, continuous, dependent only on frequency difference $\Delta\omega$ and not absolute frequency, and have a maximum at $\Delta\omega = 0$. These properties are characteristic of a wide-sense stationary process in frequency.

The PMD vector autocorrelation function originates with the dot-product of the PMD vector at two different frequencies, averaged over frequency:

$$\langle \vec{\tau}(\omega') \cdot \vec{\tau}(\omega) \rangle = \left\langle \bar{\tau}^2 \right\rangle R_{\vec{\tau}}(\mathcal{W}) \tag{9.4.9}$$

Figure 9.12(a) plots $R_{\vec{\tau}}(\mathcal{W})$ as a function of cyclic frequency $\Delta f\,\bar{\tau}$. The product $\Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle$ (or $\Delta f\,\bar{\tau}$) alone determines the form of the ACF. In an extensive study Lin and Agrawal have connected this and higher-order ACFs to pulse broadening and distortion [35].

[1] The function sinhc is a sinh function divided by its argument. In general a trigonometric function followed by the letter "c" is to be divided by its argument. At the origin, sinhc(0) = 1.

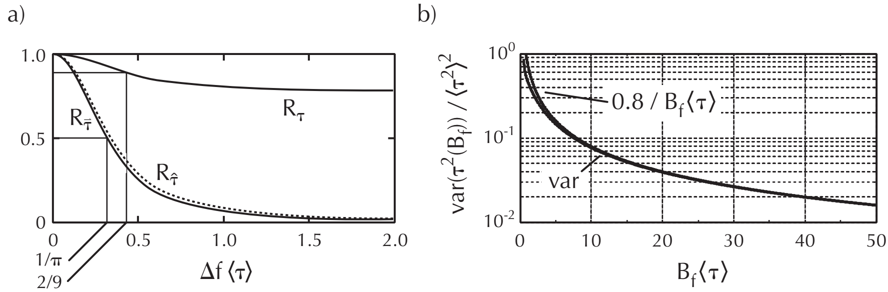

Fig. 9.12. Autocorrelation function and log variance. a) Autocorrelation function $R_{\vec{\tau}}$ of the PMD vector as function of $\Delta f\,\bar{\tau}$. FWHM is $\Delta f\,\bar{\tau} \simeq 2/\pi$. The PMD-vector ACF is dominated by depolarization, $R_{\hat{\tau}}$ (dashed curve). The root-mean-square DGD ACF has a secondary effect. b) Normalized $\log_{10}$ variance of estimated value $\langle \tau^2(B_f) \rangle$ as function of measurement bandwidth (in Hertz) B_f. Variance is well approximated by $16\sqrt{2}/\left(9\pi B_f \bar{\tau}\right)$ for $B_f \bar{\tau} > 5$.

Especially important is the autocorrelation function bandwidth, defined as the full-width half-maximum (FWHM) of the function. In cyclic frequency the FWHM for the PMD vector is

$$\Delta f_{\mathrm{PMD-ACF}}\,\bar{\tau} \simeq \frac{2}{\pi} \tag{9.4.10}$$

This is a useful relation to know. For instance, with $\bar{\tau} = 30$ ps, the correlation bandwidth is $\Delta f \simeq 21$ GHz, or $\Delta\lambda \simeq 0.16$ nm. Measurements of high-PMD fiber should at least have a 20 GHz resolution to capture the variation, but any higher resolution creates correlated measurements that oversample the PMD. Statistical estimates should be made only after the sample points of a high-resolution measurement are decimated (or are treated with other signal-processing methods) down to the autocorrelation bandwidth.

Another example is the failure of estimating mean DGD on installed fiber when the measurements are taken through a narrowband optical filter such as one port of a multiplexer. A typical channel bandwidth for such a filter is 40 GHz. The mean fiber DGD must be 32 ps or greater to have more than one statistically independent measurement within the filter passband.

As $R_{\vec{\tau}}$ is the autocorrelation of the PMD vector, the question is which component of that vector, the DGD or the pointing direction, dominates the decorrelation. In the preceding section it was determined that depolarization dominates PDCD, which is representative of the strong tendency for the PMD vector to change direction. The autocorrelation function also reflects this behavior.

The autocorrelation function of the mean-square DGD was derived by Shtaif and Mecozzi [57] to answer just this question. As the DGD-squared is the dot-product of the PMD vector with itself $\tau^2(\omega) = \vec{\tau}(\omega) \cdot \vec{\tau}(\omega)$, the correlation between $\tau^2(\omega')$ and $\tau^2(\omega)$ is

$$\left\langle \tau^2(\omega')\tau^2(\omega)\right\rangle = \frac{5}{3}\left\langle \bar{\tau}^2\right\rangle^2 R_{\tau^2}(\mathcal{W}) \tag{9.4.11}$$

where R_{τ^2} is listed in Table 9.2. The autocorrelation functions R_τ for the DGD and $R_{\hat{\tau}}$ for the pointing-direction of the PMD are then extracted as shown in the table. These two functions are also plotted in Fig. 9.12(a). The correlation bandwidths are all about the same (with $\Delta f_{\mathrm{ACF-DGD}}\,\bar{\tau} \simeq 4/9$) but the DGD ACF falls from $R_\tau(0) = 1$ to $R_\tau(\infty) = \sqrt{3/5}$, or about 22%. The PMD-vector ACF is dominated by the unit-vector ACF $R_{\hat{\tau}}$, consistent with the tendency of the PMD vector to change direction, or depolarize. The autocorrelation function for the PSP has been recently reported by Bao *et al.* [26].

The PMD-vector autocorrelation function can be modified to determine the weights of all moments of the PMD vector relative to the mean-square DGD [56]. Unlike the preceding ACFs, the moment relations apply to a single frequency. The two moment relations are

$$\left\langle \vec{\tau}^{(n-k)} \cdot \vec{\tau}^{(n+k+1)} \right\rangle = 0 \tag{9.4.12a}$$

$$\left\langle \vec{\tau}^{(n-k)} \cdot \vec{\tau}^{(n+k)} \right\rangle = (-1)^k\, \frac{(2n)!}{3^n(n+1)!}\, \left\langle \tau^2\right\rangle^{n+1} \tag{9.4.12b}$$

where $\vec{\tau}^{(n)}$ refers to the $(n{+}1)$ order of $\vec{\tau}$. While even/odd moments vanish, like moments ($k = 0$) grow quickly with n. This is another reflection of the disorder and complicated structure of the DGD spectrum. Recall from the Fourier analysis of the DGD spectrum that an increase in the number of elementary PMD segments in a concatenation increases the number of Fourier components in the spectrum. Since $\left\langle \tau^2\right\rangle \propto z$, higher moments grow increasingly quickly as the fiber length increases, consistent with the Fourier picture.

The factorial-function coefficient to $\left\langle \tau^2\right\rangle$ in (9.4.12b) grows very quickly with n. The origin of this coefficient is the white noise that underlies the model. The moments of a Brownian motion are Hermite polynomials evaluated at the origin. The resulting Hermite coefficients grow as $(2n)!/n!$; this growth is reflected in the moments of the PMD vector since it is a derived process from Brownian motion of the birefringent vector. The relative growth of the coefficient for successive moments is

$$\frac{\left\langle \vec{\tau}^{(n)} \cdot \vec{\tau}^{(n)} \right\rangle}{\left\langle \vec{\tau}^{(n-1)} \cdot \vec{\tau}^{(n-1)} \right\rangle} = \frac{2n(2n-1)}{3(n+1)} \simeq \frac{4}{3}\, n$$

For large n the growth is linear in n, again consistent with Hermite polynomial behavior.

Autocorrelation Function Derivations

The PMD-related ACFs can be derived exclusively using a stochastic-calculus treatment of the birefringence and PMD vector [37]. The PMD evolution equation (9.3.1) on page 397 may be rewritten in SDE form as

$$d\vec{\tau} = d\vec{B}_z + \omega\, d\vec{B}_z \times \vec{\tau} \tag{9.4.13}$$

where

$$d\vec{B}_z = \vec{\beta}_\omega dz, \;\; d\vec{B}_z \cdot d\vec{B}_z = \gamma^2 dz, \;\; \text{and} \;\; \langle dB_{z,j}\, dB_{z,k}\rangle = 0 \quad j \neq k \tag{9.4.14}$$

The differential $d\vec{B}_z$ is a three-entry column vector of i.i.d. Brownian motions that are the driving force in the evolution of $d\vec{\tau}$. The Brownian motion represents the local birefringence vector. The subscript z denotes that this motion is longitudinal in z and defines the domain over which averages will be taken. Brownian motion grows as $\sqrt{z}$, so according to the rules of stochastic calculus terms up to order z are included, thus $d\vec{B}_z \cdot d\vec{B}_z = \gamma^2 dz$, where γ^2 is the strength of the motion. (White noise is the formal derivative of Brownian motion: $dB_z = g_z dz$.)

The SDE (9.4.13) is in the Stratonovich form and must be translated into the Itô form. That translation generates an additional term, which when included makes

$$d\vec{\tau} = d\vec{B}_z + \omega\, d\vec{B}_z \times \vec{\tau} - \frac{\gamma^2\omega^2}{3}\,\vec{\tau} dz \tag{9.4.15}$$

This correction term shifts the drift of $\vec{\tau}$ but not the random behavior.

The first calculation is the mean-square of the DGD, or τ^2. Treating $\vec{\tau}$ as a stochastic variable, the differential of τ^2 using Itô's chain rule is

$$d(\tau^2) = 2\vec{\tau} \cdot d\vec{\tau} + d\vec{\tau} \cdot d\vec{\tau} \tag{9.4.16}$$

The clarify the following calculations, the Stratonovich form of $d\vec{\tau}$ (9.4.13) is first substituted into (9.4.16). Keeping only terms that survive an average, this partial solution gives

$$\begin{aligned} d(\tau^2) &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \omega^2 (d\vec{B}_z \times \vec{\tau})^2 \\ &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \omega^2 \left(\gamma^2\tau^2 dz - (d\vec{B}_z \cdot \vec{\tau})^2\right) \\ &= 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz + \frac{2}{3}\,\omega^2\gamma^2\tau^2 dz \end{aligned}$$

where

$$(d\vec{B}_z \cdot \vec{\tau})^2 = \sum_{k=1}^{3} dB_{z,k} \cdot dB_{z,k}\tau_k^2 + \text{cross-terms}$$

and $dB_{z,k} \cdot dB_{z,k} = \gamma^2 dz/3$. Now, adding the Itô drift correction from (9.4.16) back into $d\vec{\tau}$ and keeping terms only up to order dz gives

$$d(\tau^2) = 2\vec{\tau} \cdot d\vec{B}_z + \gamma^2 dz \tag{9.4.17}$$

Averaging this expression over z eliminates terms of order $d\vec{B}_z$. Subsequent integration over length results in the mean-square evolution of the DGD

$$\left\langle \tau^2(z) \right\rangle = \gamma^2 z \tag{9.4.18}$$

Identification with (9.3.2) on page 397 gives $\gamma^2 = 2 \left\langle \tau_c^2 \right\rangle / L_C$.

The ACF for the PMD vector can now be calculated. Denoting $\vec{\tau}_\omega = \vec{\tau}(\omega)$ (rather than the frequency derivative of $\vec{\tau}$), the Itô differential of the dot product $\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}$ is

$$d\left(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}\right) = \left(d\vec{\tau}_\omega\right) \cdot \vec{\tau}_{\omega'} + \vec{\tau}_\omega \cdot \left(d\vec{\tau}_{\omega'}\right) + d\vec{\tau}_\omega \cdot d\vec{\tau}_{\omega'}$$

Substitution of (9.4.15) and keeping terms only up to order dz gives

$$d\left(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}\right) = d\vec{B}_z \cdot \left(\vec{\tau}_\omega + \vec{\tau}_{\omega'}\right) + \Delta\omega\, \vec{\tau}_{\omega'} \cdot \left(d\vec{B}_z \times \vec{\tau}_\omega\right) + \gamma^2 dz$$
$$-\frac{1}{3}\, \gamma^2(\omega^2 + \omega'^2)(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'})dz + \omega\omega' \left(\gamma^2(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'})dz - \left(d\vec{B}_z \cdot \vec{\tau}_\omega\right)\left(d\vec{B}_z \cdot \vec{\tau}_{\omega'}\right)\right)$$

Subsequent averaging over z eliminates many terms. The average over the product of inner products in particular gives

$$\left\langle \left(d\vec{B}_z \cdot \vec{\tau}_\omega\right)\left(d\vec{B}_z \cdot \vec{\tau}_{\omega'}\right) \right\rangle = \frac{1}{3}\, \gamma^2 \left(\vec{\tau}_\omega \cdot \vec{\tau}_{\omega'}\right) dz \tag{9.4.19}$$

Completing the average over all terms gives the differential form of the autocorrelation

$$d\left\langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \right\rangle = \left(1 - \frac{1}{3}\, \Delta\omega^2 \left\langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \right\rangle\right) \gamma^2 dz$$

Integration gives

$$\left\langle \vec{\tau}_\omega \cdot \vec{\tau}_{\omega'} \right\rangle = \frac{3}{\Delta\omega^2}\left(1 - \exp\left(-\frac{\Delta\omega^2\gamma^2 z}{3}\right)\right) \tag{9.4.20}$$

Replacement of $\left\langle \bar{\tau}^2 \right\rangle = \gamma^2 z$ and normalization by $\left\langle \bar{\tau}^2 \right\rangle$ results in the normalized autocorrelation function listed in Table 9.2:

$$R_{\vec{\tau}}\left(\Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle\right) = \operatorname{sinhc}\left(\frac{\Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle}{6}\right) \exp\left(-\frac{\Delta\omega^2 \left\langle \bar{\tau}^2 \right\rangle}{6}\right) \tag{9.4.21}$$

The limits of the ACF are $R_{\vec{\tau}}(0) = 1$ and $R_{\vec{\tau}}(\infty) = 0$. It is remarkable that the only terms that enter the ACF are the frequency difference $\Delta\omega$ and the mean-square DGD $\left\langle \bar{\tau}^2 \right\rangle$. The mean-square DGD in turn is directly proportional to the square of the mean fiber DGD. Once again the mean fiber DGD is the “unit” by which a PMD-related statistical quantity is governed.

Lastly, the autocorrelation of the DGD squared is calculated. The kernel of the calculation is $\tau_\omega^2 \tau_{\omega'}^2$ where $\tau_\omega^2 = \vec{\tau}_\omega \cdot \vec{\tau}_\omega$. Treating τ_ω^2 as a stochastic variable, the differential is

$$d\left(\tau_\omega^2 \tau_{\omega'}^2\right) = \left(d\tau_\omega^2\right) \tau_{\omega'}^2 + \tau_\omega^2 \left(d\tau_{\omega'}^2\right) + \left(d\tau_\omega^2\right)\left(d\tau_{\omega'}^2\right) \tag{9.4.22}$$

Substitution of (9.4.17) into (9.4.22) and keeping terms only up to order dz gives

$$d\left(\tau_\omega^2\tau_{\omega'}^2\right) = 2\tau_{\omega'}^2\left(d\vec{B}_z\cdot\vec{\tau}_\omega\right) + 2\tau_\omega^2\left(d\vec{B}_z\cdot\vec{\tau}_{\omega'}\right) + \left(\tau_\omega^2+\tau_{\omega'}^2\right)\gamma^2 dz + 4\left(d\vec{B}_z\cdot\vec{\tau}_\omega\right)\left(d\vec{B}_z\cdot\vec{\tau}_{\omega'}\right)$$

Averaging over z and using (9.4.19) leaves

$$d\left\langle\tau_\omega^2\tau_{\omega'}^2\right\rangle = 2\gamma^4 z dz + \frac{4\gamma^2}{3}\left\langle\vec{\tau}_\omega\cdot\vec{\tau}_{\omega'}\right\rangle dz \tag{9.4.23}$$

where an Itô isometry removes the random component in $\left\langle\tau_\omega^2\right\rangle$:

$$\begin{aligned}\left\langle\tau_\omega^2\right\rangle &= \left\langle\int d(\tau_\omega^2)\right\rangle = \left\langle\int 2\vec{\tau}_\omega\cdot d\vec{B}_z + \gamma^2 dz\right\rangle\\ &= \int\gamma^2 dz\end{aligned}$$

This average is no longer a function of ω. Substitution of the PMD-vector ACF into (9.4.23) makes for a straightforward integration, resulting in

$$\left\langle\tau_\omega^2\tau_{\omega'}^2\right\rangle = \left\langle\bar{\tau}^2\right\rangle^2 + \frac{4\left\langle\bar{\tau}^2\right\rangle}{\Delta\omega^2} - \frac{12}{\Delta\omega^4}\left(1-\exp\left(-\frac{\Delta\omega^2\left\langle\bar{\tau}^2\right\rangle}{3}\right)\right) \tag{9.4.24}$$

Subsequent normalization as shown in (9.4.11) and rearrangement of terms produces the normalized ACF for the DGD squared:

$$R_{\tau^2}\left(\Delta\omega^2\left\langle\bar{\tau}^2\right\rangle\right) = \frac{3}{5}\left(1+\frac{2}{3}\frac{\left(1-R_{\vec{\tau}}\left(\Delta\omega^2\left\langle\bar{\tau}^2\right\rangle\right)\right)}{\Delta\omega^2\left\langle\bar{\tau}^2\right\rangle/6}\right) \tag{9.4.25}$$

As with the PMD-vector ACF, the DGD-squared ACF depends only on the frequency difference and mean fiber DGD. The limits of this autocorrelation are $R_{\tau^2}(0) = 1$ and $R_{\tau^2}(\infty) = 3/5$.

9.4.3 Mean-DGD Measurement Uncertainty

The PMD ACF gives the minimum bandwidth over which two neighboring PMD vectors are statistically independent. The PMD ACF can also be used to determine the uncertainty of an estimator of the mean DGD of a fiber. This important application has been studied by Gisin *et al.* [22], Karlsson and Brentel [31], Shtaif and Mecozzi [57], and Boroditsky *et al.* [4].

There is a difference in framework between the first three reports and the most recent. In particular, the relation between the mean-square DGD and average DGD is $\bar{\tau}^2 = 8/3\pi\left\langle\bar{\tau}^2\right\rangle$ is considered exact in the former reports while

Boroditsky *et al.* explain that equality holds only over infinite bandwidth (or ensemble averages). In fact, in the limit of zero bandwidth there is an 8% systematic error between mean-square and average DGD. In the broadband regime $B\left\langle\bar{\tau}^2\right\rangle > 30$ (discussed below), the error between the DGD moments is

$$\bar{\tau} = \sqrt{\frac{8}{3\pi}\left\langle\bar{\tau}^2\right\rangle} \pm \frac{8}{9\sqrt{2}}\frac{1}{B} \tag{9.4.26}$$

where B is the full measurement bandwidth in radians. Measurements of low-PMD fibers are susceptible to this error.

Putting aside this systematic error for the moment, there are two ways to estimate the mean DGD from a measurement: average the DGD values across frequency, or average the DGD-squared values across frequency and take the square-root. The former is a straight average, while the latter is gives the rms value. The studies show that the rms average gives a slightly better estimate. The variance of the rms estimate is detailed here, and Shtaif and Mecozzi give a brief comparison.

Consider the estimate of the mean-square DGD over a radian bandwidth $B = \omega_2 - \omega_1$:

$$\bar{\tau}^2_{\text{est}}(B) = \frac{1}{B}\int_B \tau^2(\omega)d\omega \tag{9.4.27}$$

The variance of $\bar{\tau}^2_{\text{est}}(B)$ is, by definition,

$$\begin{aligned}
\operatorname{var}\left(\bar{\tau}^2_{\text{est}}(B)\right) &= E\left[\bar{\tau}^2_{\text{est}}(B)\bar{\tau}^2_{\text{est}}(B)\right] - E^2\left[\bar{\tau}^2_{\text{est}}(B)\right] \\
&= \frac{1}{B^2}E\left[\int_B d\omega \int_B d\omega'\tau^2(\omega)\tau^2(\omega')\right] - \left\langle\bar{\tau}^2\right\rangle^2 \\
&= \frac{1}{B}\int_B d\omega'\left\langle\tau^2(\omega)\tau^2(\omega'-\omega)\right\rangle - \left\langle\bar{\tau}^2\right\rangle^2
\end{aligned}$$

where the double integral reduces to a single integral since the integrand depends only on the frequency difference and not absolute value. The last integrand has already been calculated, see (9.4.24). Additionally, it is more relevant to look at the normalized variance so comparisons can be made. Thus, normalizing the variance by $\left\langle\bar{\tau}^2\right\rangle^2$ and computing the integral gives

$$\begin{aligned}
\frac{\operatorname{var}\left(\bar{\tau}^2_{\text{est}}(B)\right)}{\left\langle\bar{\tau}^2\right\rangle^2} = 16\left(\frac{4 - B^2\left\langle\bar{\tau}^2\right\rangle}{B^4\left\langle\bar{\tau}^2\right\rangle^2}\right) + \frac{32}{3}\left(\frac{B^2\left\langle\bar{\tau}^2\right\rangle - 6}{B^4\left\langle\bar{\tau}^2\right\rangle^2}\right)\mathrm{e}^{-B^2\left\langle\bar{\tau}^2\right\rangle/12} \\
+ \frac{16\sqrt{3\pi}}{9}\frac{1}{B\sqrt{\left\langle\bar{\tau}^2\right\rangle}}\operatorname{erf}\left(\frac{B\sqrt{\left\langle\bar{\tau}^2\right\rangle}}{2\sqrt{3}}\right)
\end{aligned}$$

This function is plotted in Fig. 9.12(b). The asymptotic limit for $B\left\langle\bar{\tau}^2\right\rangle > 30$ is

$$\frac{\operatorname{var}\left(\bar{\tau}^2_{\text{est}}(B)\right)}{\left\langle\bar{\tau}^2\right\rangle^2} \simeq \frac{16\sqrt{3\pi}}{9}\frac{1}{B\sqrt{\left\langle\bar{\tau}^2\right\rangle}} \tag{9.4.28}$$

Translation to bandwidth in Hertz and average DGD gives

$$B\sqrt{\langle \bar{\tau}^2 \rangle} = \sqrt{\frac{3\pi^3}{2}}\ (B_f \bar{\tau}) \tag{9.4.29}$$

where $B = 2\pi B_f$. Thus,

$$\frac{\mathrm{var}\left(\bar{\tau}^2_{\mathrm{est}}(B_f)\right)}{\left\langle \bar{\tau}^2 \right\rangle^2} \simeq \frac{16\sqrt{2}}{9\pi}\frac{1}{B_f\bar{\tau}}, \quad B_f\bar{\tau} > 5 \tag{9.4.30}$$

Comparison of this approximate variance formula is given in Fig. 9.12(b).

The expected error is estimated by removing the normalization in expression (9.4.28). Since the estimated quantity is the mean-square of the DGD, plus and minus one standard deviation from the "true" value is $\bar{\tau}^2_{\mathrm{est}} \simeq \left\langle \bar{\tau}^2 \right\rangle \pm \sqrt{\mathrm{var}\left(\bar{\tau}^2_{\mathrm{est}}\right)}$. Substitution of the variance by (9.4.28), translation to cyclic frequency (9.4.29), and converting both sides to mean DGD gives the expression for the estimator uncertainty:

$$\bar{\tau}_{\mathrm{est}}(B_f) \simeq \bar{\tau}\left(1 \pm \frac{0.9}{\sqrt{B_f\,\bar{\tau}}}\right) \tag{9.4.31}$$

where the coefficient in the numerator comes from $\sqrt{16\sqrt{2}/(9\pi)}$. This coefficient agrees with Gisin [22]. Moreover, the expression shows that reduction of the standard deviation of the estimated value of $\bar{\tau}$ is a slow function: $1/\sqrt{B_f}$.

For example, consider an uncertainty of $\pm 10\%$: $B_f\,\bar{\tau} \simeq 110$. For a mean DGD of 10 ps, the required measurement bandwidth is $\sim 11,000$ GHz, or ~ 90 nm. To halve the uncertainty the bandwidth must be quadrupled. It is an open question whether an estimator with a faster convergence can be found. The square-root form for the mean-square estimator suggests an estimator based on the fourth-power of the DGD spectrum. This requires a higher-order autocorrelation function. Another way to increase the certainty of the mean DGD is to take multiple uncorrelated measurements over time. That uncertainty goes as $1/\sqrt{N}$ with N measurements; again a slow function but useful nonetheless.

Returning to Boroditsky *et al.*, the authors show that average DGD estimated from the magnitude SOPMD spectrum gives both an unbiased estimator and reduces the measurement uncertainty by 30%. The reduction in measurement uncertainty is equivalent to effectively doubling the measurement bandwidth. They further show that average DGD estimated from the PDCD spectrum along yields a better estimate of average DGD compared to direct mean-square DGD spectrum analysis. However, the magnitude SOPMD spectrum fluctuates roughly twice as fast as the corresponding DGD spectrum, which in turn requires greater care in measurement. The vector MPS technique should produce sufficiently accurate measurements. Moreover, the width of the PDCD density is only 1/9 that of the magnitude SOPMD spectrum, so again, care must be used in obtaining a sufficiently accurate measurement to effectively employ these techniques.

9.4.4 Discrete Waveplate Model

The analytic developments of this chapter are derived from a Brownian motion model of the local birefringence vector. The powerful tools of stochastic calculus and partial-differential equations are then employed to derive statistical properties of polarization and PMD. However, the cascaded waveplate model is very often used instead. The waveplate model concentrates differential delay into homogeneous segments and then abruptly mode-mixes between adjacent segments. The waveplate model is suitable as a good approximation in certain regimes as long as it is correctly constructed. While there are several variations, the model below converges to the correct statistics.

In the regime $L \gg L_C \gg L_B$, where L is the fiber length, the waveplate model illustrated in Fig. 9.13(a) gives a reasonable approximation for the PMD. In particular, the rms DGD statistics follow (9.4.1). The model uses N equal-length waveplates where each plate is $L_C/2$ long and there are $N_c = 2L/L_C$ waveplates in total. The statistics track for $N_c \gg 30$. The physical waveplate orientation is uniformly distributed on $[-\pi/2, \pi/2]$ and zero chirality is asserted. The birefringence (magnitude) of each plate is a random variable selected from a Rayleigh distribution.

There are two aspects to be worked out. One relates to the frequency bandwidth and step size and the other to the gaussian distributions of the cartesian components of the birefringence. First the frequency grid. To derive a good statistic, uncorrelated DGD values over a sufficiently wide bandwidth must be calculated. At the discrete level, the total bandwidth B_f comes from N_f points of step size Δf: $B_f = N_f \Delta f$. The minimum uncorrelated bandwidth for an average DGD $\bar{\tau}$ is $\Delta f\ \bar{\tau} = 2/\pi$, so the bandwidth-mean-DGD product is $B_f \bar{\tau} = 2N_f/\pi$. Substitution into the mean-DGD estimate (9.4.31) gives

$$\bar{\tau}_{\text{est}}(N_f) \simeq \bar{\tau} \left(1 \pm \frac{1.13}{\sqrt{N_f}} \right) \tag{9.4.32}$$

This is the basis on which N_f is set. For instance, $N_f = 500$ gives a standard deviation of 5%.

Next the birefringent distribution is determined. The mean-square DGD as a function of length (9.4.1) is rewritten as

$$\left\langle \tau^2(N_c) \right\rangle = \left\langle \beta_\omega^2 \right\rangle L_C^2 N_c \tag{9.4.33}$$

where $N_c = 2z/L_C$ and $\tau_c = \beta_\omega L_C$. The PMD distribution is Maxwellian, so a $3\pi/8$ scale factor relates its first and second moments. The birefringent distribution is Rayleigh, so the second-moment $\left\langle \beta_\omega^2 \right\rangle$ is a factor of two larger than the variance of the underlying i.i.d. gaussian cartesian-component distributions. Putting these together gives the variance of the component gaussian:

$$\sigma_{\beta,k}^2 = \frac{3\pi\bar{\tau}^2}{16 L_C^2 N_c} \tag{9.4.34}$$

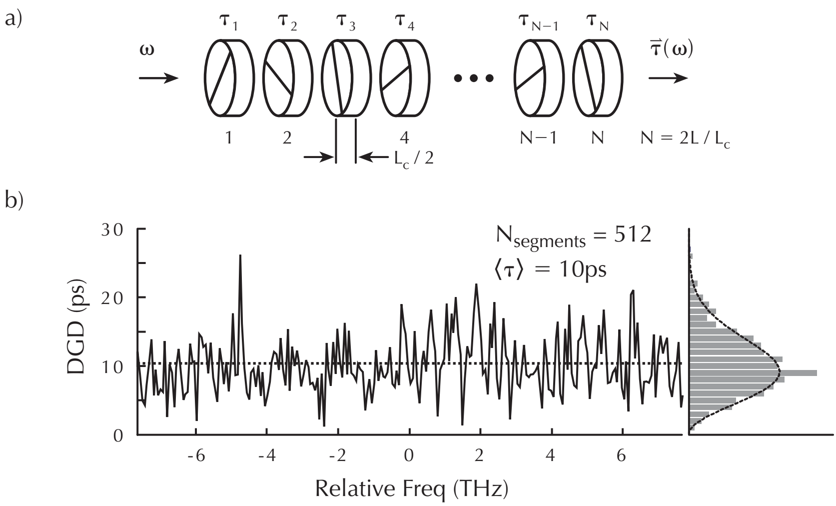

Fig. 9.13. Waveplate model of a fiber, good for the $L \gg L_C \gg L_B$ regime. a) Waveplates have uniformly distributed e-axis orientations and are each $L_C/2$ long. There are $N_c = 2L/L_C$ segments in total. The DGD per segment is determined from a Rayleigh distribution. b) Realization of a DGD spectrum for $N_c = 512$ and $N_f = 600$, given $\langle \tau \rangle = 10$ ps. The DGD distribution is shown to the right. The calculation uses large-enough frequency steps so that DGD values are statistically uncorrelated.

Each cartesian component of the birefringence is randomly picked from a normal distribution with density $\rho_{\beta,k} = N\left(0, \sigma^2_{\beta,k}\right)$. The segment birefringence magnitude is then $\beta_\omega = \sqrt{\beta^2_{\omega,1} + \beta^2_{\omega,2}}$, with a corresponding segment DGD of $\tau_c = \beta_\omega L_C$.

Finally, the average Stokes-vector rotation per frequency step $\Delta\omega\bar{\tau}_c$ is estimated by combining (9.4.33) recast in terms of $\bar{\tau}$ and $\bar{\tau}_c$, and the uncorrelated frequency step $\Delta f \bar{\tau}$. The result is

$$\Delta\omega\, \bar{\tau}_c = \sqrt{\frac{3\pi^2}{2N_c}} \tag{9.4.35}$$

For instance, with $N_c = 512$, the average Stokes rotation per frequency step is $\Delta\omega\, \bar{\tau}_c \simeq 9.7°$. This is a good check because a single step in excess of $\Delta\omega\, \bar{\tau}_c > \pi$ creates an ambiguity as to whether the PMD completed more than a half-revolution in one direction or less than half in the other direction.

A cropped spectral window of a DGD spectrum constructed in the manner outlined is illustrated in Fig. 9.13(b). The waveplate cascade was made with $N_c = 512$ waveplates calculated at $N_f = 600$ uncorrelated frequency points. The resulting distribution and its Maxwellian fit are plotted on the right. One instance of the concatenation using this number of waveplates and frequency points is not sufficient to derive high-quality statistics, especially on the tail of the Maxwellian. But multiple runs of this cascade using different realizations of the birefringence vector will build up the statistics as a rate of $\sqrt{N}$. Note that varying the waveplate length as well as the aforementioned factors does not improve the statistics and works only to lower the convergence rate.

9.4.5 Karhunen-Loève Expansion of Brownian Motion

Other than the waveplate model above, the derivations in this chapter have relied on Brownian motion as the driving term for the evolution of various parameters. Brownian motion is often modelled on a microscopic, step-by-step level where the displacement for each step comes from choosing a random value from a gaussian density. This approach works but has no analytic expression.

A useful alternative is the Karhunen-Loeve (KL) expansion of Brownian motion [59]. The KL expansion gives a macroscopic view of the motion on an interval and guarantees the proper covariance. For Brownian motion the KL expansion on $[0, 1]$ is

$$B_z = \sum_{k=1}^{\infty} \frac{\sqrt{2}}{\pi\left(k+\frac{1}{2}\right)}\, \xi_k \sin\left(\pi\left(k+\tfrac{1}{2}\right) z\right) \tag{9.4.36}$$

where ξ_k are random variables with density $N(0, 1)$. Each term in the summation spans the entire interval. Higher values of k produce higher oscillations but with lower amplitudes. In practice the sum is taken large enough to fill in the necessary spatial resolution and is thereafter truncated. Figure 9.14(a) shows four sample paths generated by (9.4.36).

The KL expansion is the function-space analogue of a Markov process at the discrete level. On this level a Markov process is determined purely by its covariance matrix A. The eigenvectors and values are found from the equation $Ax = \lambda x$. The spectral theorem gives the entries in A in terms of its eigenvectors and values: $A = \sum_{k=1}^{n} \lambda_k v_k v_k^T$. On the continuous level, the eigenvalue equation is

$$\int_0^1 K(z, y)\varphi(y)dy = \lambda\varphi(z) \tag{9.4.37}$$

where the covariance of the process is $K(z, y) = z \wedge y$. The symbol $\wedge$ is means the minimum of the two quantities. The spectral theorem says that the entries of the covariance function, entries which are now functions not vectors, are

$$K(z, y) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(y)\varphi_k(z) \tag{9.4.38}$$

where λ_k are the eigenvalues of (9.4.37) and $\varphi_k(z)$ are its eigenvectors.

The eigenvalue equation is solved by substituting in the covariance of Brownian motion. This gives

$$\int_0^1 (z \wedge y)\, \varphi(y) dy = \lambda \varphi(z)$$

The integral is separated into two pieces to make an integral equation:

$$\int_0^z y\varphi(y)dy + \int_z^1 z\varphi(y)dy = \lambda\varphi(z) \tag{9.4.39}$$

This is an integral equation which can be solved by taking successive derivatives. Looking ahead a couple of steps, the differential equation produced by the above integral equation is second order, and therefore requires two boundary conditions to be solved uniquely. One boundary condition is found directly from this integral equation: when $z = 0$ then $\varphi(0) = 0$.

Differentiating both sides of (9.4.39) with respect to z makes

$$z\varphi(z) + \int_z^1 \varphi(y)dy - z\varphi(z) = \lambda\varphi'(z)$$

where $\varphi'(z)$ denotes the first derivative with respect to z. After cancelling the two terms in the left, a second boundary condition is determined: for $z = 1$ $\varphi'(1) = 0$. Differentiating again gives

$$\varphi(z) = -\lambda\varphi''(z) \tag{9.4.40}$$

This ODE is solved subject to the boundary conditions $\varphi(0) = \varphi'(1) = 0$. The general solution is

$$\varphi(z) = A\sin(az) + B\cos(bz)$$

where the derivatives yield

$$\begin{aligned}\varphi'(z) &= aA\cos(az) - bB\sin(bz)\\ \varphi''(z) &= -a^2A\sin(az) - b^2B\cos(bz)\end{aligned}$$

The boundary conditions restrict the four unknown coefficients in the following way:

$$\begin{aligned}\varphi(0) = 0 \quad &\longrightarrow \qquad B = 0\\ \varphi'(1) = 0 \quad &\longrightarrow \quad aA\cos(a) = 0\end{aligned}$$

Substitution of $\varphi(z)$ into the differential equation (9.4.40) gives the definition for a: $a = 1/\sqrt{\lambda}$. Summarizing these restrictions, the solution thus far is

$$\varphi(z) = A\lambda^{-1/2}\sin\left(z\lambda^{-1/2}\right) \tag{9.4.41}$$

subject to the condition that

$$A\lambda^{-1/2}\cos\left(\lambda^{-1/2}\right) = 0$$

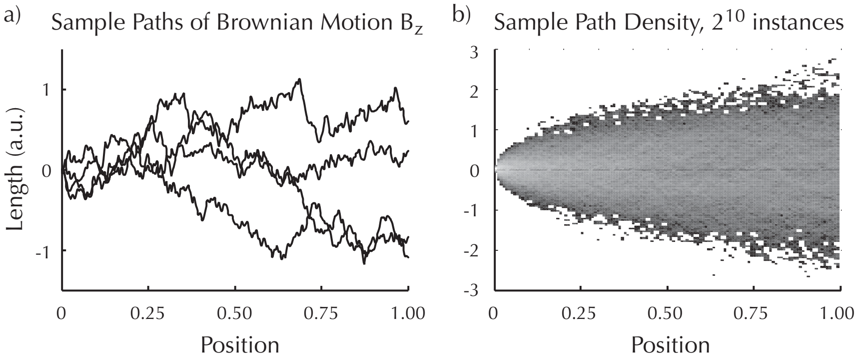

Fig. 9.14. Sample paths of Brownian motion created by the Karhunen-Loeve expansion. a) Four sample paths on the interval $[0,1]$. b) Density of 2^{10} sample paths on the interval. As expected, the density width increases as square-root of the length.

This boundary condition is satisfied when

$$\frac{1}{\sqrt{\lambda_k}} = \pi\left(k+\tfrac{1}{2}\right), \qquad k \in \mathbb{Z}$$

Rearranging, the eigenvalue is

$$\lambda_k = \left(\pi\left(k+\tfrac{1}{2}\right)\right)^{-2} \tag{9.4.42}$$

Finally, the coefficient A is determined from the orthogonality condition of $\varphi(z)$:

$$\int_0^1 \varphi_k^2(y)dy = 1$$

This is satisfied when $A=\sqrt{2}$. The eigenvector function is therefore

$$\varphi_k(z) = \sqrt{2}\sin\left(\pi\left(k+\tfrac{1}{2}\right)z\right), \qquad z \in [0,1], k \in \mathbb{Z} \tag{9.4.43}$$

The KL expansion for a general gaussian process G_z is

$$G_z = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\,\xi_k\,\varphi_k(z) \tag{9.4.44}$$

where the random variable ξ is characterized by $E\left[\xi_k\right]=0$ and $E\left[\xi_k^2\right]=1$. One can show that $E\left[G_y G_z\right]$ reproduces the correct covariance (9.4.38). Substitution of (9.4.43) into (9.4.44) makes the expression (9.4.36) stated at the beginning.

Figure 9.14(b) shows the density of B_z on the interval $[0,1]$ for 2^{10} sample paths. As expected the density grows as the square-root of the length.

9.5 PDL Statistics

The statistics of polarization-dependent loss are derived in an analogous manner as those for polarization and PMD. The local differential loss is modelled as a white-noise process and the cumulative PDL is determined by the diffusion of the PDL vector $\vec{\Gamma}$ along the fiber. Concomitant with the assumption of no chiral birefringence, circular dichroism is excluded from the PDL model.

When immersed in a random birefringence medium, the axes of minimum and maximum transmission of a local differential loss element $\vec{\alpha}$ are scrambled from one point to another. Recall the evolution of the cumulative PDL vector in the presence of birefringence from (8.3.16) on page 377:

$$\frac{d\vec{\Gamma}}{dz} = \vec{\beta} \times \vec{\Gamma} + \vec{\alpha} - \left(\vec{\alpha} \cdot \vec{\Gamma}\right) \vec{\Gamma} \tag{9.5.1}$$

The cross-product term spins the cumulative PDL vector about the local birefringent axis $\vec{\beta}$, scrambling its orientation. Propagation through multiple randomly oriented birefringent elements drives the PDL vector toward isotropic coverage of the Poincaré sphere. The second term pulls the cumulative PDL vector toward the local element, while the last term governs the growth and decay of $\vec{\Gamma}$.

The statistics for PDL reported in the literature are based on PDL immersed in random birefringence [10, 19, 38, 64]. PMD statistics, by contrast, were derived in the absence of PDL. The reason PMD is included in PDL statistics is because a long concatenation of pure PDL is not likely in a telecommunications link. The consequence of PMD inclusion is that the local differential loss is treated as three-dimensional i.i.d. white noise in Stokes space. The correlations of the white-noise vector $\vec{\alpha}$ are

$$\langle \alpha_j(z) \rangle = 0, \qquad \langle \alpha_j(z) \alpha_k(z') \rangle = \frac{\sigma_\alpha^2}{3}\, \delta_{j,k}(z - z') \tag{9.5.2}$$

where σ_α^2 is the strength of the disturbance. A differential Brownian vector is defined as $d\vec{B}_z = \vec{\alpha} dz$ such that $d\vec{B}_z \cdot d\vec{B}_z = \sigma_\alpha^2 dz$.

The evolution equation (9.5.1) with the cross-produce removed (as its effect averages to zero in the isotropic PDL model) is rewritten in SDE form as

$$d\vec{\Gamma} = \left(I - \vec{\Gamma}\vec{\Gamma}\cdot\right) d\vec{B}_z$$

This diffusion equation is interpreted in the Stratonovich sense and must be translated to Itô form in order to use Itô calculus. The translation makes [38]

$$d\vec{\Gamma} = -\frac{\sigma_\alpha^2}{3} \left(2 - \Gamma^2\right) \vec{\Gamma}\, dz + \left(I - \vec{\Gamma}\vec{\Gamma}\cdot\right) d\vec{B}_z$$

The diffusion generator for this equation is

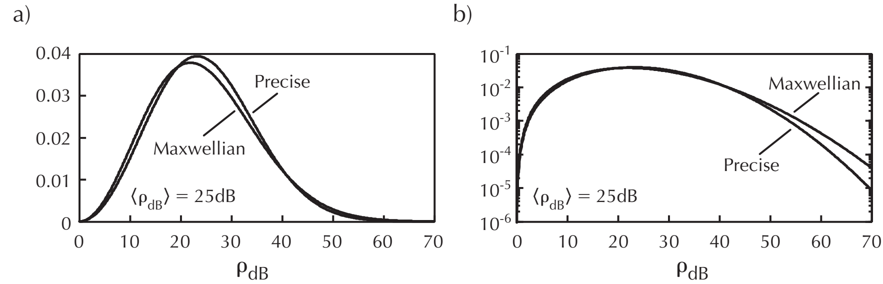

Fig. 9.15. PDL probability density and Maxwellian approximation verses decibel value, linear and semi-log scales. The log scale is in $\log_{10}$.

$$G = -\frac{\sigma_\alpha^2 \left(2-\Gamma^2\right)}{3} \left(\sum_{i=1}^{3} \Gamma_i \frac{\partial}{\partial \Gamma_i} + \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \Gamma_i \Gamma_j \frac{\partial^2}{\partial \Gamma_i \partial \Gamma_j} \right) + \frac{\sigma_\alpha^2}{6} \sum_{i=1}^{3} \frac{\partial^2}{\partial \Gamma_i^2} \tag{9.5.3}$$

where $\left(\sigma\sigma^T\right) = \left(I - \vec{\Gamma}\vec{\Gamma}^T\right)\left(I - \vec{\Gamma}\vec{\Gamma}^T\right)^T = \left(I - (2-\Gamma^2)\vec{\Gamma}\vec{\Gamma}^T\right)$. One can now calculate expectations of the diffusion using Kolmogorov's backward equation (9.2.5) on page 394.

It is an oddity of PDL that diffusions of Γ_k, Γ^n, $\vec{\Gamma}^n$ and the like are difficult to solve while those of the logarithm of $T_{\max}/T_{\min}$ make closed solutions. It would appear that the $(2-\Gamma^2)$ coefficient is only cleanly removed when an logarithmic function is used. Fukada does, however, succeed in expressing the probability densities in linear terms [18]. For the present, the moments of the PDL magnitude expressed in decibels are used. Recall the definition:

$$\rho_{\mathrm{dB}} = 10 \log_{10} \left(\frac{1+\Gamma}{1-\Gamma} \right) \tag{9.5.4}$$

It is particular to PDL that even though the cartesian components of the local PDL vectors are modelled as isotropic i.i.d. Gaussian random variables the cumulative distribution is not strictly Maxwellian. Shtaif and Mecozzi show that for low cumulative PDL (25 dB or less, although indeed this is extremely high for a lightwave system) the distribution is approximately Maxwellian [38]. Galtarossa and Palmieri use the diffusion generator to calculate the PDL distribution exactly and validate the Shtaif and Mecozzi approximation [19].

The details of the Galtarossa and Palmieri calculation are laborious but indeed elegant. Central to their calculations are the functionals $\psi = \rho^{2k}$ and $\psi' = \rho^{2k+1}/\Gamma$. From this they construct the characteristic function (CF) of the ρ^2 density and, after proof of convergence, inverse-Fourier transform the CF to the density proper. Subsequent conversion to the density of ρ $(= \rho_{\mathrm{dB}})$ makes

$$\rho_{\mathsf{\rho}}(p,\tilde{z}) = \frac{2p^2}{\gamma^3\sqrt{2\pi\tilde{z}^3}}\,\frac{\sinh(p/\gamma)}{(p/\gamma)}\,\exp\left(-\frac{p^2}{2\gamma^2\tilde{z}}\right)\,\exp\left(-\frac{\tilde{z}}{2}\right) \tag{9.5.5}$$

for $p \geq 0$, where $\tilde{z} = z\sigma_\alpha^2/3$ and $\gamma = 20\log 10\mathrm{e} \simeq 8.868$. This density is plotted on linear and semi-log scale in Fig. 9.15. The first and second moments are

$$\langle \mathsf{\rho}(\tilde{z})\rangle = \gamma\left(\sqrt{\frac{2\tilde{z}}{\pi}}\,\mathrm{e}^{-\tilde{z}/2} + (1+\tilde{z})\,\mathrm{erf}\left(\frac{\tilde{z}}{2}\right)\right) \tag{9.5.6a}$$

$$\left\langle \mathsf{\rho}^2(\tilde{z})\right\rangle = \gamma^2\,(\tilde{z}+3)\,\tilde{z} \tag{9.5.6b}$$

In the limit of large $\tilde{z}$ the cumulative PDL grows linearly with $\tilde{z}$. The Shtaif and Mecozzi second moment is, by comparison,

$$\left\langle \mathsf{\rho}^2(\tilde{z})\right\rangle = \frac{9\gamma^2}{2}\left(\mathrm{e}^{2\tilde{z}/3} - 1\right) \tag{9.5.7}$$

Both expressions are equal to second order in $\tilde{z}$.

The Maxwellian approximation to the PDL density function (9.5.5) written in terms of the second moment is

$$\rho_{\mathsf{\rho}}(p,\tilde{z}) \simeq \frac{2p^2}{\sqrt{2\pi\left(\langle\mathsf{\rho}^2(\tilde{z})\rangle/3\right)^3}}\,\exp\left(-\frac{p^2}{2\left(\langle\mathsf{\rho}^2(\tilde{z})\rangle/3\right)}\right), \quad p \geq 0 \tag{9.5.8}$$

Figure 9.15 shows a comparison between the Maxwellian approximation and the precise PDL density. The cumulative mean PDL is $\langle\mathsf{\rho}\rangle = 25$ dB; for lower mean PDL's the approximation improves. However, the extreme case plotted here indicates the divergence of the true distribution from the Maxwellian: the tails of the precise distribution fall faster for high PDL values. This means that the cartesian-component distributions of the logarithm PDL vector $\vec{\mathsf{\rho}}_{\mathrm{dB}}$ have slightly shorter tails that the Gaussian distributions of $\vec{B}_z$.

Finally, PDL and PMD are complementary regarding partial polarization. PMD tends to depolarize a perfectly polarized input, while PDL tends to repolarize a perfectly unpolarized input. While not presented here, the statistics of PDL-induced repolarization are derived by Menyuk *et al.* and have an approximate Maxwellian form as well [41].

References

1. L. Arnold, *Stochastic Differential Equations: Theory and Applications.* Malabar, Florida: Krieger Publishing Company, 1992, reprinted from original 1974 edition.
2. G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, no. 2, pp. 310–312, Feb. 2002.
3. G. Biondini, W. Kath, and C. Menyuk, "Importance sampling for polarization-mode dispersion: techniques and applications," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.
4. M. Boroditsky, M. Brodsky, N. J. Frigo, P. Magill, and M. Shtaif, "Improving the accuracy of mean DGD estimates by analysis of second-order PMD statistics," *IEEE Photonics Technology Letters*, vol. 16, no. 3, pp. 792–794, Mar. 2004.
5. M. Brodsky, P. Magill, and N. J. Frigo, "Polarization-mode dispersion of installed recent vintage fiber as a parametric function of temperature," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 209–211, Jan. 2004.
6. X. Chen, M. Li, and D. A. Nolan, "Polarization mode dispersion of spun fibers: An analytical solution," *Optics Letters*, vol. 27, no. 5, pp. 294–296, Mar. 2002.
7. F. Curti, B. Daino, G. de Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," *Journal of Lightwave Technology*, vol. 8, no. 8, pp. 1162–1166, Aug. 1990.
8. W. B. Davenport, *Probability and Random Processes.* New York: McGraw-Hill, Inc., 1970.
9. M. C. de Lignie, H. Nagel, and M. van Deventer, "Large polarization mode dispersion in fiber optic cables," *Journal of Lightwave Technology*, vol. 12, no. 8, pp. 1325–1329, Aug. 1994.
10. A. El Amari, N. Gisin, B. Perny, H. Zbinden, and C. W. Zimmer, "Statisitcal prediction and experimental verification of concatenations of fiber optic components with polarization dependent loss," *Journal of Lightwave Technology*, vol. 16, no. 3, pp. 332–339, Mar. 1998.
11. S. L. Fogal, G. Biondini, and W. L. Kath, "Correction to: Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, pp. 1487–1489, 2002.
12. ——, "Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 14, no. 9, pp. 1273–1275, Sept. 2002.
13. E. Forestieri, "A fast and accurate method for evaluating joint second-order PMD statistics," *Journal of Lightwave Technology*, vol. 21, no. 11, pp. 2942–2952, Nov. 2003.
14. G. J. Foschini, L. Nelson, R. Jopson, and H. Kogelnik, "Probability densities of the second order polarization mode dispersion including polarization dependent chromatic dispersion," *IEEE Photonics Technology Letters*, vol. 12, no. 3, pp. 293–295, Mar. 2000.
15. G. J. Foschini, R. M. Jopson, L. E. Nelson, and H. Kogelnik, "The statistics of PMD-induced chromatic fiber dispersion," *Journal of Lightwave Technology*, vol. 17, no. 9, pp. 1560–1565, Sept. 1999.
16. G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1882–1886, Dec. 1991.

17. G. J. Foschini and C. D. Poole, “Statistical theory of polarization dispersion in single mode fibers,” *Journal of Lightwave Technology*, vol. 9, no. 11, p. 1439, Nov. 1991.
18. Y. Fukada, “Probability density function of polarization dependent loss (pdl) in optical transmission system composed of passive devices and connecting fibers,” *Journal of Lightwave Technology*, vol. 20, no. 6, pp. 953–964, June 2002.
19. A. Galtarossa and L. Palmieri, “The exact statistics of polarization-dependent loss in fiber-optic links,” *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 57–59, Jan. 2003.
20. A. Galtarossa, L. Palmieri, and D. Sarchi, “Measure of spin period in randomly birefringent low-PMD fibers,” *IEEE Photonics Technology Letters*, vol. 16, no. 4, pp. 1131–1133, Apr. 2004.
21. A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, “Statistical characterization of fiber random birefringence,” *Optics Letters*, vol. 25, no. 18, pp. 1322–1324, Sept. 2000.
22. N. Gisin, B. Gisin, J. Von der Weid, and R. Passy, “How accurately can one measure a statistical quantity like polarization-mode dispersion?” *IEEE Photonics Technology Letters*, vol. 8, no. 12, pp. 1671–1673, 1996.
23. N. Gisin, “Solutions of the dynamical equation for polarization dispersion,” *Optics Communications*, vol. 86, pp. 371–373, 1991.
24. N. Gisin and J. P. Pellaux, “Polarization mode dispersion: Time versus frequency domains,” *Optics Communications*, vol. 89, pp. 316–323, May 1992.
25. J. P. Gordon and H. Kogelnik, “PMD fundamentals: Polarization mode dispersion in optical fibers,” *Proceedings of National Academy of Sciences*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000. [Online]. Available: http://www.pnas.org
26. S. Hadjifaradji, L. Chen, D. S. Waddy, and X. Bao, “Autocorrelation function of the principal state of polarization vector for systems having PMD,” *IEEE Photonics Technology Letters*, vol. 16, no. 6, pp. 1489–1491, June 2004.
27. B. Huttner, J. Reecht, N. Gisin, R. Passy, and J. Weid, “Distributed beatlength measurement in single-mode fibers with optical frequency-domain reflectometry,” *Journal of Lightwave Technology*, vol. 20, no. 5, pp. 828–835, May 2002.
28. E. Ibragimov, G. Shtengel, and S. Suh, “Statistical correlation between first and second order PMD,” *Journal of Lightwave Technology*, vol. 20, no. 4, pp. 586–590, 2002.
29. I. P. Kaminow, “Polarization in optical fibers,” *IEEE Journal of Quantum Electronics*, vol. QE-17, no. 1, pp. 15–22, 1981.
30. M. Karlsson, “Probability density functions of the differential group delay in optical fiber communication systems,” *Journal of Lightwave Technology*, vol. 19, no. 3, pp. 324–331, Mar. 2000.
31. M. Karlsson and J. Brentel, “Autocorrelation function of the polarization-mode dispersion vector,” *Optics Letters*, vol. 24, no. 14, pp. 939–941, July 1999.
32. P. J. Leo, G. R. Gray, G. J. Simer, and K. B. Rochford, “State of polarization changes: Classification and measurement,” *Journal of Lightwave Technology*, vol. 21, no. 10, pp. 2189–2193, Oct. 2003.
33. I. T. Lima, A. O. Lima, J. Zweek, and C. R. Menyuk, “Efficient computation of outage probabilities due to polarization effects in a WDM system using a reduced stokes model and importance sampling,” *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 45–47, Jan. 2003.

34. I. T. Lima, R. Khosravani, P. Ebrahimi, E. Ibragimov, C. R. Menyuk, and A. E. Willner, "Comparison of polarization mode dispersion emulators," *Journal of Lightwave Technology*, vol. 19, no. 12, pp. 1872–1881, Dec. 2001.
35. Q. Lin and G. P. Agrawal, "Correlation theory of polarization mode dispersion in optical fibers," *Journal of the Optical Society of America B*, vol. 20, no. 2, pp. 292–301, Feb. 2003.
36. D. Marcuse, C. R. Menyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 15, no. 9, pp. 1735–1746, Sept. 1997.
37. A. Mecozzi, private communication, 2004.
38. A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 3, pp. 313–315, Mar. 2002.
39. ——, "Study of the two-frequency moment generating function of the PMD vector," *IEEE Photonics Technology Letters*, vol. 15, no. 12, pp. 1713–1715, Dec. 2003.
40. C. R. Menyuk and P. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," *Journal of the Optical Society of America B*, vol. 11, no. 7, pp. 1288–1296, July 1994.
41. C. Menyuk, D. Wang, and A. Pilipetskii, "Repolarization of polarization-scrambled optical signals due to polarization dependent loss," *IEEE Photonics Technology Letters*, vol. 9, no. 9, pp. 1247–1249, Sept. 1997.
42. L. Nelson, R. Jopson, H. Kogelnik, and G. J. Foschini, "Measurement of depolarization and scaling associated with second-order polarization mode dispersion in optical fibers," *IEEE Photonics Technology Letters*, vol. 11, no. 12, pp. 1614–1617, Dec. 1999.
43. D. Nolan, X. Chen, and M.-J. Li, "Fibers with low polarization-mode dispersion," *Journal of Lightwave Technology*, vol. 22, no. 4, pp. 1066–1077, Apr. 2004.
44. B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th ed. New York: Springer, 1998.
45. D. L. Peterson, B. C. Ward, K. B. Rochford, P. J. Leo, and G. Simer, "Polarization mode dispersion compensator field trial and field fiber characterization," *Optics Express*, vol. 10, no. 14, pp. 614–621, July 2002. [Online]. Available: http://www.opticsexpress.org/
46. S. M. Pietralunga, M. Ferrario, P. Martelli, and M. Martinelli, "Direct observation of local birefringence and axis rotation in spun fiber with centimetric resolution," *IEEE Photonics Technology Letters*, vol. 16, no. 1, pp. 212–214, Jan. 2004.
47. A. Pizzinat, L. Palmieri, B. S. Marks, C. R. Menyuk, and A. Galtarossa, "Analytical treatment of randomly birefringent periodically spun fibers," *Journal of Lightwave Technology*, vol. 21, no. 12, pp. 3355–3363, Dec. 2003.
48. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Optics Letters*, vol. 13, no. 8, pp. 687–689, Aug. 1988.
49. ——, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," *Optics Letters*, vol. 14, no. 10, pp. 523–525, 1989.
50. C. D. Poole and C. R. Giles, "Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber," *Optics Letters*, vol. 13, no. 2, pp. 155–157, 1988.

51. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," *Optics Letters*, vol. 16, no. 6, pp. 372–374, 1991.
52. C. D. Poole and D. L. Favin, "Polarization-mode dispersion measurements based on transmission spectra through a polarizer," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 917–929, 1994.
53. S. C. Rashleigh, "Origins and control and polarization effects in single-mode fiber," *Journal of Lightwave Technology*, vol. LT-1, no. 2, pp. 312–331, 1983.
54. S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single mode fibers," *Optics Letters*, vol. 3, no. 2, pp. 60–62, 1978.
55. H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. New York: Springer, 1989.
56. M. Shtaif and A. Mecozzi, "Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states," *IEEE Photonics Technology Letters*, vol. 12, no. 1, pp. 53–55, Jan. 2000.
57. ——, "Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion," *Optics Letters*, vol. 25, no. 10, pp. 707–709, May 2000.
58. Y. Tan, J. Yang, W. L. Kath, and C. R. Menyuk, "Transient evolultion of the polarization-dispersion vector's probability distribution," *Journal of the Optical Society of America B*, vol. 19, no. 5, pp. 992–1000, May 2002.
59. E. Vanden-Eijnden, private communication, Courant Institute of Mathematical Sciences, New York University, N.Y., 2003.
60. P. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 19, no. 19, pp. 1517–1519, Oct. 1994.
61. ——, "Anisotropic diffusion of the state of polarization in optical fibers with randomly varying birefringence," *Optics Letters*, vol. 20, no. 24, pp. 2493–2495, Dec. 1995.
62. ——, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *Journal of Lightwave Technology*, vol. 14, no. 2, pp. 148–157, Feb. 1995.
63. J. Yang, W. L. Kath, and C. R. Menyuk, "Polarization mode dispersion probability distribution for arbitrary distances," *Optics Letters*, vol. 26, no. 19, pp. 1472–1474, Oct. 2001.
64. M. Yu, C. Kan, M. Lewis, and A. Sizmann, "Statistics of polarization-dependent loss, insertion loss, and signal power in optical communication systems," *IEEE Photonics Technology Letters*, vol. 14, no. 12, pp. 1695–1697, Dec. 2002.