Nonlinear polarization-mode dispersion in optical fibers with randomly varying birefringence

P. K. A. Wai,* W. L. Kath,† C. R. Menyuk, and J. W. Zhang

Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250

Randomly varying birefringence leads to nonlinear polarization-mode dispersion (PMD) in addition to the well-known linear PMD. Here we calculate the variance of the field fluctuations produced by this nonlinear PMD. Knowing the size of these fluctuations is useful for assessing when nonlinear PMD is important and for its proper incorporation in fast numerical algorithms. We also derive the equilibrium probability distributions for the PMD coefficients, and we track the evolution of the polarization state's probability distribution from its initial delta-function distribution to its steady-state uniform distribution on the Poincaré sphere. © 1997 Optical Society of America [S0740-3224(97)03411-5]

1. INTRODUCTION

Optical fibers have become an important communication medium because of their waveguiding properties, i.e., their ability to trap light and to transmit information to wherever the fiber is installed.[1] Such transmission is never perfect, of course, and nonideal behavior leads to limitations in the performance of communication systems based on optical fibers. For example, because different frequencies propagate with slightly different group velocities and any signal is necessarily constructed from a range of frequencies, the signal will disperse or spread with distance.[1] One can counteract this dispersion in a number of ways, such as by sending signals at the zero-dispersion wavelength of the optical fiber,[1–3] by using dispersion compensation or management,[4–6] or by using self-phase modulation to turn propagating signal pulses into optical solitons.[2,7] Furthermore, communication fibers are transparent over a wide range of frequencies. Thus any added wideband frequency noise, such as amplified spontaneous emission noise from the erbium-doped fiber amplifiers used to compensate for loss in the fiber, will be propagated unattenuated along with the signal. For optical solitons this frequency noise is subsequently turned into random pulse position fluctuations by their propagation,[8] and unless it is corrected by addition of optical filters,[9,10] this timing jitter increases rapidly enough to impose a limit on the capacity of soliton communication systems.

In addition, the two polarization states in an ideal, perfectly circular optical fiber are degenerate, so the propagation wave number does not depend on the polarization state. Real optical fibers are never perfect, of course, and have slight asymmetries or other perturbations that destroy the degeneracy, leading to two polarization states with slightly different phase and group velocities, a phenomenon known as birefringence.[2] In the situation in which the birefringence axes and magnitude are fixed as a function of distance along the fiber, the evolution of the optical pulse envelope has been shown to be governed by the coupled nonlinear Schrödinger equation.[11–13] In standard communication fibers, however, both the orientation of the principal birefringence axes and the magnitude of the birefringence vary randomly with distance along the fiber. In this case the fiber is on average perfectly circular, and the mean pulse evolution is governed by a special case of the coupled nonlinear Schrödinger equation known as the Manakov equation.[14,15]

When the birefringence varies randomly with distance along the fiber, the deviations in the polarization state from the mean behavior determined by the Manakov equation are also random. Because the two principal polarizations have slightly different group velocities, these polarization fluctuations produce linear or intensity-independent signal pulse broadening and distortion, known as polarization-mode dispersion (PMD). Such linear PMD has been studied in great detail.[16–20] In addition to the linear PMD, however, the polarization fluctuations also produce intensity-dependent signal pulse distortions, known as nonlinear PMD.[21] In this paper we determine the statistics of the pulse distortions produced when nonlinear PMD acts on a propagating signal. These estimates show that nonlinear PMD effects are negligible for present-day, practical communication systems but that they will be significantly more important in proposed very-high-data-rate transmission system operating in the neighborhood of 100 Gbits/s. In addition, we show that nonlinear PMD effects can be enhanced in communication systems that employ optical fiber that has been spun during manufacture to reduce the amount of linear PMD.[22]

Several length scales are important for the discussion of PMD effects. The first is the fiber decorrelation length,[21] h_{fiber}, which is the autocorrelation length scale associated with variations in the principal axes of birefringence and which is an intrinsic property of the fiber. The second is the birefringence beat length, L_B, which is a measure of the birefringence strength and is the length scale over which a 2π-phase shift would be produced between the two principal polarizations if the birefringence

axes were fixed. The calculation for the statistics of the nonlinear PMD fluctuations is shown to simplify considerably when either $h_{\text{fiber}} \ll L_B$ or $h_{\text{fiber}} \gg L_B$. Finally, there is the polarization decorrelation length, h_E, which is the autocorrelation length associated with the electromagnetic field fluctuations induced by the random birefringence variations.[21]

The formulation of the paper is as follows. In Sections 2 and 3 we review the Manakov PMD equations that govern the evolution of nonlinear optical signals in the presence of random birefringence and describe the method that we use to calculate the statistics of the nonlinear PMD fluctuations. In Section 4 we calculate these statistics for the case $h_{\text{fiber}} \ll L_B$, and in Section 5 we do so for the case $h_{\text{fiber}} \gg L_B$. In Section 6 we show how the method can also be used to calculate the well-known result for the linear PMD. In Sections 7 and 8 we discuss how the Fokker–Planck equation for the probability distribution of the polarization state can be simplified when $h_{\text{fiber}} \ll L_B$ and $h_{\text{fiber}} \gg L_B$. Finally, Section 9 gives our conclusions.

2. MANAKOV POLARIZATION-MODE DISPERSION EQUATIONS

The dimensionless set of equations that describe the evolution of the electromagnetic field envelope in a nonlinear optical fiber with random birefringence are the Manakov PMD equations[21,23]

$$i \frac{\partial \bar{\Psi}}{\partial z} \pm \frac{1}{2} \frac{\partial^2 \bar{\Psi}}{\partial t^2} + \frac{8}{9} |\bar{\Psi}|^2 \bar{\Psi} = -ib' \bar{\boldsymbol{\sigma}} \frac{\partial \bar{\Psi}}{\partial t} - \frac{1}{3} (\hat{\mathbf{N}} - \langle \hat{\mathbf{N}} \rangle),$$

$$(1)$$

where $\bar{\Psi} = (\bar{U}, \bar{V})^T$, $\pm$ is for either anomalous or normal second-order dispersion, and $b' = \partial b / \partial \omega$ is the group-delay difference per unit length (recall that $\Delta = 2b$ is the birefringence strength). An alternative measure of the amount of birefringence is the beat length, $L_B = \pi/b$. Also,

$$\bar{\boldsymbol{\sigma}} = \begin{bmatrix} x_1 & x_2 - ix_3 \\ x_2 + ix_3 & -x_1 \end{bmatrix}, \quad (2)$$

$$\hat{N}_1 = z_1{}^2(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} - z_1(z_2 - iz_3)$$
$$\times (2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} - z_1(z_2 + iz_3)\bar{U}^2\bar{V}^*$$
$$- (z_2 - iz_3)^2\bar{V}^2\bar{U}^*, \quad (3a)$$

$$\hat{N}_2 = z_1{}^2(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} + z_1(z_2 + iz_3)$$
$$\times (2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} + z_1(z_2 - iz_3)\bar{V}^2\bar{U}^*$$
$$- (z_2 + iz_3)^2\bar{U}^2\bar{V}^*, \quad (3b)$$

$$\langle \hat{N}_1 \rangle = \frac{1}{3}(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U}, \quad (4a)$$

$$\langle \hat{N}_2 \rangle = \frac{1}{3}(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V}. \quad (4b)$$

The random variables x_j and z_j are the first and the third components, respectively, of the solution of the vector equation

$$\frac{\mathrm{d}}{\mathrm{d}z} \begin{pmatrix} x_j \\ y_j \\ z_j \end{pmatrix} = \begin{bmatrix} 0 & g_\theta & 0 \\ -g_\theta & 0 & 2b \\ 0 & -2b & 0 \end{bmatrix} \begin{pmatrix} x_j \\ y_j \\ z_j \end{pmatrix} \quad (5)$$

for $j = 1, 2, 3$, where $g_\theta = \mathrm{d}\theta/\mathrm{d}z$ is a white-noise process. At $z = 0$ the initial conditions $x_j = \delta_{j1}$, $y_j = \delta_{j2}$, and $z_j = \delta_{j3}$ are satisfied, where δ_{jk} is the Kronecker delta function; i.e., $\delta_{jk} = 1$ if $j = k$, and $\delta_{jk} = 0$ otherwise.

Here $\bar{\Psi}$ is the electric-field envelope measured relative to axes that follow the linear polarization state in the fiber and is related to the electric-field envelope measured relative to fixed axes, $\mathbf{A} = (A_1, A_2)^T$, as follows. First, the linear evolution of the fixed axis polarization state is determined by the solution of [21]

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}z} = ib\boldsymbol{\Sigma}\mathbf{A}, \quad (6)$$

where $\boldsymbol{\Sigma} = \cos\theta\boldsymbol{\sigma}_3 + \sin\theta\boldsymbol{\sigma}_1$, θ is the angle between the principal birefringence directions and fixed coordinate axes, and

$$\boldsymbol{\sigma}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (7)$$

are the Pauli spin matrices. The two polarization components relative to local birefringence axes are then determined from $\mathbf{A}$ by use of the transformation

$$\boldsymbol{\Psi} = \begin{pmatrix} U \\ V \end{pmatrix} = \begin{bmatrix} \cos\theta/2 & \sin\theta/2 \\ -\sin\theta/2 & \cos\theta/2 \end{bmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad (8)$$

so Eq. (6) becomes

$$i \frac{\mathrm{d}\boldsymbol{\Psi}}{\mathrm{d}z} + \tilde{\boldsymbol{\Sigma}}\boldsymbol{\Psi} = 0, \quad \tilde{\boldsymbol{\Sigma}} = \begin{bmatrix} b & -ig_\theta/2 \\ ig_\theta/2 & -b \end{bmatrix}. \quad (9)$$

Stokes parameters provide a good way to track the polarization state of the signal envelope.[21,24] The Stokes parameters relative to the local polarization axes are $\tilde{S}_1 = |U|^2 - |V|^2 = \boldsymbol{\Psi}^\dagger \boldsymbol{\sigma}_3 \boldsymbol{\Psi}$, $\tilde{S}_2 = UV^* + VU^* = \boldsymbol{\Psi}^\dagger \boldsymbol{\sigma}_1 \boldsymbol{\Psi}$, and $\tilde{S}_3 = i(UV^* - VU^*) = \boldsymbol{\Psi}^\dagger \boldsymbol{\sigma}_2 \boldsymbol{\Psi}$. The third Stokes parameter, $\tilde{S}_3$, was used previously with the opposite sign,[21] but the above is consistent with its original definition.[24] These local Stokes parameters satisfy the coupled set of equations[21]

$$\frac{\mathrm{d}}{\mathrm{d}z} \begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \tilde{S}_3 \end{pmatrix} = \begin{bmatrix} 0 & g_\theta & 0 \\ -g_\theta & 0 & 2b \\ 0 & -2b & 0 \end{bmatrix} \begin{pmatrix} \tilde{S}_1 \\ \tilde{S}_2 \\ \tilde{S}_3 \end{pmatrix}. \quad (10)$$

Note that the equation that governs the evolution of the coefficient triplets (x_j, y_j, z_j), Eq. (5), is the same as that which governs the evolution of the Stokes parameters.

We obtain the final connection between $\bar{\Psi}$ and $\mathbf{A}$ by using the solution of

$$i \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}z} + \tilde{\boldsymbol{\Sigma}}\mathbf{T} = 0, \quad \mathbf{T}(0) = \mathbf{I} \quad (11)$$

as a transformation defining $\bar{\Psi}$ through

$$\boldsymbol{\Psi}(z, t) = \mathbf{T}(z)\bar{\boldsymbol{\Psi}}. \quad (12)$$

Note that

$$\mathbf{T}(z) \equiv \begin{bmatrix} u_1 & -u_2{}^* \\ u_2 & u_1{}^* \end{bmatrix} \tag{13}$$

is a unitary matrix, so $|u_1|^2 + |u_2|^2 = 1$. It is from this matrix that the coefficients x_j, y_j, and z_j are defined[21]:

$$x_1 = |u_1|^2 - |u_2|^2, \qquad x_2 + ix_3 = -2u_1u_2,$$

$$y_1 = u_1u_2{}^* + u_1{}^*u_2, \qquad y_2 + iy_3 = u_1{}^2 - u_2{}^2,$$

$$z_1 = i(u_1u_2{}^* - u_1{}^*u_2), \qquad z_2 + iz_3 = i(u_1{}^2 + u_2{}^2). \tag{14}$$

If we use the vectors (x_j, y_j, z_k), $j = 1, 2, 3$, to form the columns of a matrix $\mathbf{Q}$:

$$\mathbf{Q} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}, \tag{15}$$

then, because initially $(x_1, y_1, z_1) = (1, 0, 0)$, $(x_2, y_2, z_2) = (0, 1, 0)$, and $(x_3, y_3, z_3) = (0, 0, 1)$, we have $\mathbf{Q} = \mathbf{I}$ at $z = 0$. Furthermore, from Eq. (5) we have

$$\frac{d\mathbf{Q}}{dz} = \mathbf{WQ}, \tag{16}$$

where

$$\mathbf{W} = \begin{bmatrix} 0 & g_\theta & 0 \\ -g_\theta & 0 & 2b \\ 0 & -2b & 0 \end{bmatrix}. \tag{17}$$

Because $\mathbf{W}$ is antisymmetric and $\mathbf{Q}(0) = \mathbf{I}$, $\mathbf{Q}(z)$ is an orthogonal matrix. Thus both the columns and the rows of $\mathbf{Q}$ are orthonormal vectors. This fact is used in Sections 4 and 5.

In the simplest birefringence model it is assumed that the birefringence strength $2b$ remains constant and that the orientation angle θ is driven by the white-noise process:

$$\frac{d\theta}{dz} = g_\theta(z), \tag{18}$$

where

$$\langle g_\theta(z) \rangle = 0, \qquad \langle g_\theta(z)g_\theta(z') \rangle = \sigma_\theta{}^2 \delta(z - z'). \tag{19}$$

This model has been shown to capture all the essentials of the more realistic case in which both the orientation angle and the birefringence strength vary randomly.[20,25] Technically speaking, the averaging here should be done over a random ensemble of fibers, or rather over a random ensemble of the functions $g_\theta(z)$ that determine the birefringence angle $\theta(z)$; however, physical systems such as this one that can be described by Markov processes are generally ergodic,[26,27] in which case ensemble averages can be replaced by spatial averages.

From Eqs. (18) and (19) a diffusion equation for the probability distribution of θ can be obtained[21,28]:

$$\frac{\partial p}{\partial z} = \frac{1}{2} \sigma_\theta{}^2 \frac{\partial^2 p}{\partial \theta^2}, \tag{20}$$

where $p(\theta, z)$ is 2π periodic in θ and $p(\theta, 0) = \delta(\theta - \theta_0)$. Equation (20) allows one of the fundamental

length scales present in the fiber to be determined, namely, the fiber correlation length h_{fiber}. This is the length scale over which the birefringence axes lose memory of their original orientation. From the solution of Eq. (20),

$$p(\theta, z) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} \exp\left(-\frac{1}{2}\sigma_\theta{}^2 n^2 z\right) \cos n(\theta - \theta_0), \tag{21}$$

and in particular from

$$\langle \cos(\theta - \theta_0) \rangle = \exp(-{}^1\!/{}_2\sigma_\theta{}^2 z), \tag{22}$$

we obtain $h_{\text{fiber}} = 2/\sigma_\theta{}^2$.[21] The θ variations drive the birefringence fluctuations because the birefringence matrix $\mathbf{\Sigma}$ depends on θ, and the birefringence fluctuations in turn produce polarization variations in the electromagnetic field envelope $\overline{\mathbf{\Psi}}$ through the coefficients x_j and z_j in Eq. (1). Note, however, that the x_j appear only as linear terms and therefore generate linear PMD, whereas the z_j coefficients appear nonlinearly and thus produce nonlinear PMD.[21]

3. FORMULATION: POLARIZATION-MODE DISPERSION STATISTICS

The statistics of the coefficients x_j and z_j can be determined either from the coefficients' probability distributions or from the equations for their statistical moments. The former method gives complete information about the coefficients' statistics, but finding the probability distribution involves solving a partial differential equation known as the Fokker–Planck equation.[27] We defer this approach until Sections 7 and 8. Using averages and statistical moments of the coefficients is more direct but still yields useful information about the behavior of the linear and nonlinear PMD.

The equations for the means and the second moments of the various coefficients have already been determined; in terms of the Stokes parameters these equations are[21]

$$\frac{d}{dz} \langle x_j \rangle = -{}^1\!/{}_2\sigma_\theta{}^2\langle x_j \rangle, \tag{23a}$$

$$\frac{d}{dz} \langle y_j \rangle = -{}^1\!/{}_2\sigma_\theta{}^2\langle y_j \rangle - 2b\langle z_j \rangle, \tag{23b}$$

$$\frac{d}{dz} \langle z_j \rangle = 2b\langle y_j \rangle, \tag{23c}$$

where $j = 1, 2, 3$, and

$$\frac{d}{dz} \langle x_jx_k \rangle = -\sigma_\theta{}^2(\langle x_jx_k \rangle - \langle y_jy_k \rangle), \tag{24a}$$

$$\frac{d}{dz} \langle y_jy_k \rangle = \sigma_\theta{}^2(\langle x_jx_k \rangle - \langle y_jy_k \rangle) - 4b\langle {}^1\!/{}_2(y_jz_k + y_kz_j) \rangle, \tag{24b}$$

$$\frac{d}{dz} \langle z_jz_k \rangle = 4b\langle {}^1\!/{}_2(y_jz_k + y_kz_j) \rangle, \tag{24c}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}\langle {}^1\!/_2(y_j z_k + y_k z_j)\rangle$$

$$= 2b(\langle y_j y_k\rangle - \langle z_j z_k\rangle)$$

$$- {}^1\!/_2\sigma_\theta^2\langle {}^1\!/_2(y_j z_k + y_k z_j)\rangle, \quad (24\mathrm{d})$$

where $j,\ k = 1, 2, 3$. As might be expected, Eqs. (23) and (24) are the same ones that the means and the second moments of the Stokes parameters satisfy.[21] In particular,

$$\langle x_j\rangle \leftrightarrow \langle \widetilde{S}_1\rangle, \qquad \langle y_j\rangle \leftrightarrow \langle \widetilde{S}_2\rangle, \qquad \langle z_j\rangle \leftrightarrow \langle \widetilde{S}_3\rangle,$$

$$\langle x_j x_k\rangle \leftrightarrow \langle \widetilde{S}_1{}^2\rangle, \qquad \langle y_j y_k\rangle \leftrightarrow \langle \widetilde{S}_2{}^2\rangle, \qquad \langle z_j z_k\rangle \leftrightarrow \langle \widetilde{S}_3{}^2\rangle,$$

$$\langle {}^1\!/_2(y_j z_k + y_k z_j)\rangle \leftrightarrow \langle \widetilde{S}_2\widetilde{S}_3\rangle. \quad (25)$$

Not all the first and the second moments of x_j, y_j, and z_j appear in Eqs. (2) and (3), of course; thus the only terms that are needed are x_1, x_2, x_3, $z_1^2 - (1/3)$, $z_1 z_2$, $z_1 z_3$, $z_2^2 - z_3^2$, and $z_2 z_3$. When one talks about the variations induced in the field envelope $\bar{\boldsymbol{\Psi}}$, however, it is really the statistics of the integrals of these quantities that are needed. The variations in $\bar{\boldsymbol{\Psi}}$ depend on the integrals of the coefficients because the length scale over which the fluctuations in x_j, y_j, and z_j occur (the polarization or field decorrelation length,[21] h_E) is much shorter than the natural length scale over which the pulse envelope evolves, namely, the dispersion length L_{disp}. [Note that in Eq. (1) all spatial variables have been scaled by this length.]

If the birefringence fluctuations produce uniform mixing on the Poincaré sphere, then the field decorrelation length is proportional to the fiber decorrelation length, $h_E \propto h_{\mathrm{fiber}}$.[21] This is the situation when the fiber decorrelation length is much longer than the beat length, $h_{\mathrm{fiber}} \gg L_B$. When $h_{\mathrm{fiber}} \ll L_B$, however, azimuthal mixing on the Poincaré sphere is poor and $L_B^2/(12\pi^2 h_{\mathrm{fiber}})$ becomes the relevant length scale for fluctuations in the coefficients z_j.[21] In this limit, then, nonlinear PMD becomes much more significant than linear PMD. Nevertheless, even for values that lie at the edge of the realistic parameter range, e.g., $h_{\mathrm{fiber}} \approx 0.3$ m and $L_B \approx 100$ m, the nonlinear PMD length scale is smaller than dispersion lengths that are typical in standard communication systems, and the disparate length scales can be exploited to yield the statistics of the induced field fluctuations.

Inasmuch as the field decorrelation length is short compared with the dispersion length, the field envelope $\bar{\boldsymbol{\Psi}}$ is able to respond not to any individual changes in these coefficients but rather only to their cumulative effects. Alternatively, as $\mathrm{d}\bar{\boldsymbol{\Psi}}/\mathrm{d}z$ is $O(1)$, the envelope is able to build up appreciable changes only after distances of the order of the dispersion length, and it experiences only small variations over shorter distances. Thus, over a short distance interval from z_0 to $z_0 + l$, in the terms on the right-hand side of Eq. (1), $\bar{\boldsymbol{\Psi}}$ can be replaced by its initial value, $\bar{\boldsymbol{\Psi}}(z_0, t)$, and only the random coefficients themselves are integrated.

By the central-limit theorem,[29] the probability distribution for the integrals of these various nonlinear PMD coefficients will be Gaussian to a good approximation when $l \gg h_E$ because they can be thought of as sums of a large number of effectively independent random variables.

Their means, variances, and covariances can be calculated by extension of the system of moments, Eqs. (23) and (24), to include these new terms (alternatively, the Fokker–Planck equation can be extended). The number of equations grows quickly, however, as additional unneeded moments must be included to make the set of equations complete. However, as long as we are concerned primarily with distance intervals l that are long compared with the mixing length scale of the various coefficients, a simpler method can be employed. Essentially, we calculate

$$\left\langle \int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta \right\rangle = \int_{z_0}^{z_0+l} \langle f(\zeta)\rangle \mathrm{d}\zeta, \quad (26)$$

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta \right] = \left\langle \int_{z_0}^{z_0+l} [f(\zeta) - \langle f(\zeta)\rangle]\mathrm{d}\zeta \right.$$

$$\left. \times \int_{z_0}^{z_0+l} [f(\zeta') - \langle f(\zeta')\rangle]\mathrm{d}\zeta' \right\rangle$$

$$= \int_{z_0}^{z_0+l}\int_{z_0}^{z_0+l} \langle f(\zeta) - \langle f(\zeta)\rangle]$$

$$\times [f(\zeta') - \langle f(\zeta')\rangle]\rangle \mathrm{d}\zeta \mathrm{d}\zeta', \quad (27)$$

where $f(\zeta)$ represents one of the nonlinear PMD coefficients. When z_0 is large compared with the field decorrelation length, however, we can replace $\langle f(\zeta)\rangle$ by its large-distance asymptotic value, $\bar{f}$, say. In this case, the expression for the mean [Eq. (26)] becomes

$$\left\langle \int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta \right\rangle \sim \int_{z_0}^{z_0+l} \bar{f}\mathrm{d}\zeta = \bar{f}l. \quad (28)$$

Note that relation (28) is not expected to be a good approximation within a few field decorrelation lengths of $z_0 = 0$ because in that region the initial polarization state biases the result for $\langle f(\zeta)\rangle$. Once $z_0 \gg h_E$, however, the random motion of the polarization state has acted over a distance long enough for the initial state to have been forgotten.

In a similar way, the variance becomes

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta \right]$$

$$= \int_{z_0}^{z_0+l}\int_{z_0}^{z_0+l} \langle [f(\zeta) - \bar{f}][f(\zeta') - \bar{f}]\rangle \mathrm{d}\zeta \mathrm{d}\zeta'$$

$$= \int_{z_0}^{z_0+l}\int_{z_0}^{z_0+l} R(\zeta - \zeta')\mathrm{d}\zeta \mathrm{d}\zeta', \quad (29)$$

where $R(\zeta - \zeta') = \langle [f(\zeta) - \bar{f}][f(\zeta') - \bar{f}]\rangle$ is the autocorrelation function associated with $f(\zeta)$.[30] Again, within a few field decorrelation lengths of the initial condition at $z_0 = 0$ the autocorrelation function will be more complicated than this because the process is not really stationary in that region; once $z_0 \gg h_E$, however, the process loses memory of the initial condition and the autocorrelation becomes a function only of the difference $\zeta - \zeta'$. Equation (29) simplifies to

$$\int_{z_0}^{z_0+l} \int_{z_0-\zeta}^{z_0+l-\zeta} R(\xi)\mathrm{d}\xi\mathrm{d}\zeta = 2\int_0^l (l-\xi)R(\xi)\mathrm{d}\xi, \quad (30)$$

where we have used $R(-\xi) = R(\xi)$. For $l \gg h_E$ we then have approximately

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta\right] \sim 2l\int_0^\infty R(\xi)\mathrm{d}\xi - 2\int_0^\infty \xi R(\xi)\mathrm{d}\xi, \quad (31)$$

because generally $R(\xi) \to 0$ exponentially for $|\xi| \gg h_E$. If the autocorrelation function is a simple exponential function, $R(\xi) = R(0)\exp(-|\xi|/h_E)$, then Eq. (30) becomes

$$2R(0)h_E{}^2\left[\frac{l}{h_E} + \exp\left(-\frac{l}{h_E}\right) - 1\right], \quad (32)$$

which for $l \gg h_E$ yields

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta\right] \sim 2R(0)h_E l - 2R(0)h_E{}^2, \quad (33)$$

which allows us to define the decorrelation or mixing length[21] h_E for more general cases by looking at the term proportional to l in relation (31):

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} f(\zeta)\mathrm{d}\zeta\right] \sim 2\langle[f(\zeta) - \bar{f}]^2\rangle h_E l = 2R(0)h_E l, \quad (34)$$

so that

$$h_E \equiv \frac{1}{R(0)}\int_0^\infty R(\xi)\mathrm{d}\xi. \quad (35)$$

The autocorrelation function $R(\xi)$ can be calculated in a straightforward manner from the equations for the means and the variances of the random coefficients, Eqs. (23) and (24), when $z_0 \gg h_E$. Again, take $f(z)$ to represent one of the coefficients. If we assume temporarily that $f(z_0)$ is known, then we can use $f(z_0)$ as an initial condition in Eqs. (23) or (24) to calculate the conditional expectation[29] or average of the autocorrelation function given $f(z_0)$:

$$\langle[f(z_0) - \bar{f}][f(z_0 + \xi) - \bar{f}]|f(z_0)\rangle$$
$$= [f(z_0) - \bar{f}]\langle[f(z_0 + \xi) - \bar{f}]|f(z_0)\rangle. \quad (36)$$

We then find the autocorrelation function by averaging over all possible states $f(z_0)$:

$$R(\xi) = \langle\langle[f(z_0) - \bar{f}][f(z_0 + \xi) - \bar{f}]|f(z_0)\rangle\rangle. \quad (37)$$

Here the inner average is done assuming that $f(z_0)$ is given and the outer average is done over all possible states of $f(z_0)$. If $z_0 \gg h_E$, then the polarization state has become completely randomized on the Poincaré sphere and this last step becomes a simple angular averaging. Note, however, that Eqs. (24) for the second moments are still sufficiently complicated that additional approximations are desirable. It is shown that considerable simplification results in the limits $h_\mathrm{fiber} \ll L_B$ (Section 4) and $h_\mathrm{fiber} \gg L_B$ (Section 5).

Finally, a similar asymptotic expression for the pairwise covariances of the integrated nonlinear PMD coefficients can be obtained in terms of the integral of the cross-correlation function.[30] This expression is omitted

here, as an argument is given in Section 4 that all such asymptotic covariances are zero.

The central-limit theorem states that the probability distribution associated with these integrated terms will be close to a joint Gaussian distribution[29] when $l \gg h_E$. Thus, once all means, variances, and covariances are known, it is relatively simple to numerically generate realizations of the integrated terms that have the required statistics. As a result, it is possible to propose a method for numerically integrating Eq. (1) by identifying l with the numerical step size Δz. We start by integrating the unperturbed Manakov equation, i.e., the deterministic left-hand side of the equation, using some numerical method.[2] The PMD effects are then included by the addition, after each numerical step, of a single sample of a set of random variables with the same statistics as the integrated terms. In the case of the nonlinear PMD terms this can likely be done in such a way as to preserve the same quantities that the continuous equation conserves; for example, a small additive term $i\Delta\Phi\bar{\boldsymbol{\Psi}}$ can be written as the multiplicative term $\exp(i\Delta\Phi)\bar{\boldsymbol{\Psi}}$ to the same degree of accuracy, but the latter method preserves $|\bar{\boldsymbol{\Psi}}|^2$.

The size of the numerical step that can be taken with this method is, of course, limited both by the accuracy required by the numerical method when one is solving the deterministic terms and by the requirement that the effects of PMD per step be small. Nevertheless, it appears that relatively large steps should be able to be taken with this method, in a manner similar to the coarse-step method used for the coupled nonlinear Schrödinger equation.[14,15,23]

4. NONLINEAR POLARIZATION-MODE DISPERSION STATISTICS WHEN $h_\mathrm{fiber} \ll L_B$

In the limit $h_\mathrm{fiber} \ll L_B$, we have $\sigma_\theta{}^2 \gg b$ in Eqs. (24). The quasi-steady approximation to Eq. (24a) then shows that for $z \gg h_\mathrm{fiber}$ we have $\langle x_j x_k\rangle \approx \langle y_j y_k\rangle$. Also, because the vectors (x_j, y_j, z_j) are the columns of an orthogonal matrix $\mathbf{Q}$, we have $x_j x_k + y_j y_k + z_j z_k = \delta_{jk}$, i.e., 1 if $j = k$ and 0 otherwise. When these facts are used in Eq. (24d), and the limit $\sigma_\theta{}^2 \gg b$ is again applied, we find that

$$\langle\tfrac{1}{2}(y_j z_k + y_k z_j)\rangle \approx \frac{2b}{\sigma_\theta{}^2}[\delta_{jk} - 3\langle z_j z_k\rangle]. \quad (38)$$

Finally, substituting relation (38) into Eq. (24c), we have

$$\frac{\mathrm{d}}{\mathrm{d}z}\langle z_j z_k\rangle \approx \frac{24b^2}{\sigma_\theta{}^2}[\tfrac{1}{3}\delta_{jk} - \langle z_j z_k\rangle]. \quad (39)$$

From relation (39) it is seen that the decay length for the second moments of z is $\sigma_\theta{}^2/24b^2$, which in this limit is much longer than $2/\sigma_\theta{}^2 = h_\mathrm{fiber}$.

The means and the variances of the integrated coefficients can now be found. First, we can calculate the long-term mean of $\tfrac{1}{3}\delta_{jk} - \langle z_j z_k\rangle$ needed in relation (28) by interpreting the transpose of the orthogonal matrix $\mathbf{Q}$ as a rotation. In particular, the last row (z_1, z_2, z_3) is the rotation $\mathbf{Q}^T$ applied to the initial vector $(0, 0, 1)$, and, as rotations preserve lengths, (z_1, z_2, z_3) has unit length.

This vector thus represents some arbitrary point on the unit sphere, and relative to arbitrary fixed axes this point can be represented as

$$\langle z_1, z_2, z_3 \rangle = (\cos \varphi, \ \sin \varphi \cos \psi, \ \sin \varphi \sin \psi) \tag{40}$$

for some spherical angles φ and ψ. Thus

$$z_1{}^2 - (1/3) = \cos^2 \varphi - (1/3), \tag{41a}$$

$$2z_1 z_2 = 2 \cos \varphi \sin \varphi \cos \psi = \sin 2\varphi \cos \psi, \tag{41b}$$

$$2z_1 z_3 = 2 \cos \varphi \sin \varphi \sin \psi = \sin 2\varphi \sin \psi, \tag{41c}$$

$$z_2{}^2 - z_3{}^2 = \sin^2 \varphi(\cos^2 \psi - \sin^2 \psi) = \sin^2 \varphi \cos 2\psi, \tag{41d}$$

$$2z_2 z_3 = 2 \sin^2 \varphi \cos \psi \sin \psi = \sin^2 \varphi \sin 2\psi. \tag{41e}$$

Moreover, because the fixed axes are completely arbitrary the various components of this representation can also be rotated cyclically when necessary: $z_1 \to z_2$, $z_2 \to z_3$, and $z_3 \to z_1$.

When $z_0 \gg h_E$ the probability distribution associated with the polarization state will be uniformly distributed over the Poincaré sphere, and we therefore have

$$\langle z_1{}^2 - 1/3 \rangle = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} (\cos^2 \varphi - 1/3)\sin \varphi \mathrm{d}\varphi \mathrm{d}\psi = 0, \tag{42}$$

$$\langle z_1 z_2 \rangle = \langle z_2 z_3 \rangle = \langle z_2 z_3 \rangle$$

$$= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} (\cos \varphi \sin \varphi \cos \psi)\sin \varphi \mathrm{d}\varphi \mathrm{d}\psi$$

$$= 0, \tag{43}$$

so, as expected, the means of the various nonlinear PMD terms do not grow with distance. The variances of the integrated terms are found from relation (31) and Eq. (37). From relation (39) we have

$$\langle z_j z_k - 1/3 \delta_{jk} | z_j(z_0) z_k(z_0) \rangle$$

$$= [z_j(z_0) z_k(z_0) - 1/3]\exp\left[-\frac{24b^2}{\sigma_\theta{}^2} (z - z_0) \right], \tag{44}$$

so that

$$R(\xi) = \langle [z_j(z_0) z_k(z_0) - 1/3 \delta_{jk}]^2 \rangle \exp\left(-\frac{24b^2}{\sigma_\theta{}^2} \xi \right). \tag{45}$$

From Eq. (45) we can see that

$$R(0) = \langle [z_j(z_0) z_k(z_0) - 1/3 \delta_{jk}]^2 \rangle$$

$$= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} (z_j z_k - 1/3 \delta_{jk})^2 \sin \varphi \mathrm{d}\varphi \mathrm{d}\psi. \tag{46}$$

Because

$$\langle (\cos^2 \varphi - 1/3)^2 \rangle = 4/45, \tag{47a}$$

$$\langle (\cos \varphi \sin \varphi \sin \psi)^2 \rangle = 1/15, \tag{47b}$$

$$\langle (\sin^2 \varphi \cos 2\psi)^2 \rangle = 4/15, \tag{47c}$$

we can therefore apply relation (33), using the mixing length $h_E = \sigma_\theta{}^2/24b^2 = L_B{}^2/(12\pi^2 h_{\text{fiber}})$, to obtain

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} (z_1{}^2 - 1/3)\mathrm{d}\zeta \right] \sim \frac{8}{45} h_E(l - h_E), \tag{48}$$

$$\mathrm{Var}\left(\int_{z_0}^{z_0+l} z_j z_k \mathrm{d}\zeta \right) \sim \frac{2}{15} h_E(l - h_E) \tag{49}$$

for $(j, k) = (1, 2)$, $(j, k) = (1, 3)$, or $(j, k) = (2, 3)$, and

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l} \left(z_2{}^2 - z_3{}^2 \right)\mathrm{d}\zeta \right] \sim \frac{8}{15} h_E(l - h_E). \tag{50}$$

In addition, if relation (32) is used, relations (48)–(50) become

$$\mathrm{Var} \sim \sigma_E{}^2 \left[\frac{l}{h_E} + \exp\left(-\frac{l}{h_E} \right) - 1 \right], \tag{51}$$

where $\sigma_E{}^2 = 8/45 h_E{}^2$, $2/15 h_E{}^2$, or $8/15 h_E{}^2$.

Each of these integrated nonlinear PMD coefficients yields the same mixing length scale or field decorrelation length, $h_E = \sigma_\theta{}^2/24b^2 = L_B{}^2/(12\pi^2 h_{\text{fiber}})$. This expression shows one of the somewhat counterintuitive features of nonlinear PMD when $b \ll \sigma_\theta{}^2$, namely, that the amount of nonlinear PMD increases as the birefringence decreases. We can understand this fact by examining Eq. (5), where the changes in θ are seen to produce random rotations about the z_j axis, while the birefringence b is proportional to the rotation rate about the x_j axis. When b is small, therefore, relatively slow z_j variations result, and latitudinal mixing on the Poincaré sphere is poor. As the z_j coefficients appear in the fluctuating cross-phase modulation terms in Eq. (1) and the averaged effect of these terms is polarization dependent whenever the polarization state is not uniformly distributed on the Poincaré sphere, the reduced mixing that arises when b decreases produces an increase in the amount of nonlinear PMD. From a physical perspective this result is not surprising. As the birefringence strength tends to zero, the length scale over which the field remains correlated along a fixed set of axes increases.[20,25] The ellipticity of the field varies in accordance with this length scale even when the field is measured with respect to the rapidly varying local frame.

The covariances of the various integrated coefficients can also be calculated in a similar manner, but it is clear from the above arguments that in each case the asymptotic behavior will be proportional to the cross-correlation function evaluated at zero delay, i.e., the pairwise covariances of the nonintegrated coefficients. It can easily be seen from Eqs. (41) that the averages of all pairwise combinations of the various coefficients are zero, however, and therefore the asymptotic covariances of all the various integrated coefficients vanish. The integrated coefficients are therefore statistically independent of one another.

Figure 1 shows a comparison between the numerical estimate for the variance of $\int (z_1{}^2 - 1/3)\mathrm{d}z$ obtained by solution of Eq. (5) and the theoretical result, relation (51). Several values of the ratio h_{fiber}/L_B were considered, and for each case 5000 separate solutions of Eq. (5) were cal-

culated, with each separate run using a different random sequence to simulate the white-noise process. The agreement is seen to be very good when h_{fiber}/L_B is small, but significant discrepancies begin to appear when this ratio is as large as 1/10. It is shown in Section 5 that this discrepancy is due mainly to errors in the estimate for h_E. Similar curves were produced for the variances of the integrals of the other coefficients, but they are almost identical to those shown here. Thus the theoretical result that the variances are in ratio to one another in the amounts $8/45$:$2/15$:$8/15$ is a good estimate. Similarly, the numerical results for both the means and the covariances of the various integrals were found to be negligibly small, as expected.

In Section 3 it was indicated that when $l \gg h_E$ the probability distributions for the various integrated nonlinear PMD terms should be Gaussian. For short distances, however, the probability distributions will be close to the steady-state distributions of the integrands. We can calculate these steady-state distributions by first recalling that the steady-state distribution of the polarization state is uniform on the Poincaré sphere, so $p(\varphi, \psi) = 1/4\pi$, and that the average of any quantity is determined by means of $\langle F \rangle = \int\int F p(\varphi, \psi)\sin\varphi \, d\varphi \, d\psi$. We determine the probability distribution of $u = z_1{}^2 - 1/3 = \cos^2\varphi - 1/3$, for example, from the marginal distribution $p(\varphi) = 1/2$ (obtained by integration of ψ of out of the joint distribution) by using[29]

$$p(u) = p(\varphi)\sin\varphi \, \frac{d\varphi}{du} = \frac{1}{4\sqrt{u + 1/3}},$$
$$-1/3 < u < 2/3. \quad (52)$$

The steady-state probability distributions of the other integrands can be determined similarly. Figure 2 shows the numerically determined probability distribution of $\int(z_1{}^2 - 1/3)dz$ for various distances l when $L_B/h_{\text{fiber}} = 10$. For short lengths the distribution clearly reflects
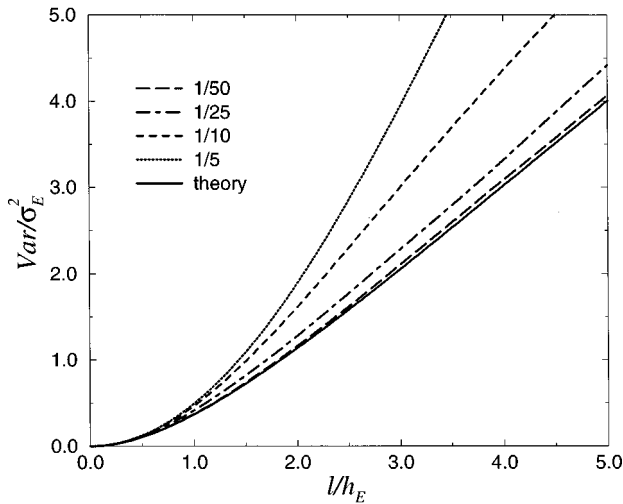

Fig. 1. Comparison between the numerical and the theoretical estimates for the variance of $\int_0^l (z_1{}^2 - 1/3)dz$ for several values of h_{fiber}/L_B when the ratio is small. Each curve represents the results of 5000 solutions of Eq. (5). The theoretical curve is Eq. (51) with $\sigma_E{}^2 = 8/45 h_E{}^2$ and $h_E = L_B{}^2/(12\pi^2 h_{\text{fiber}})$.
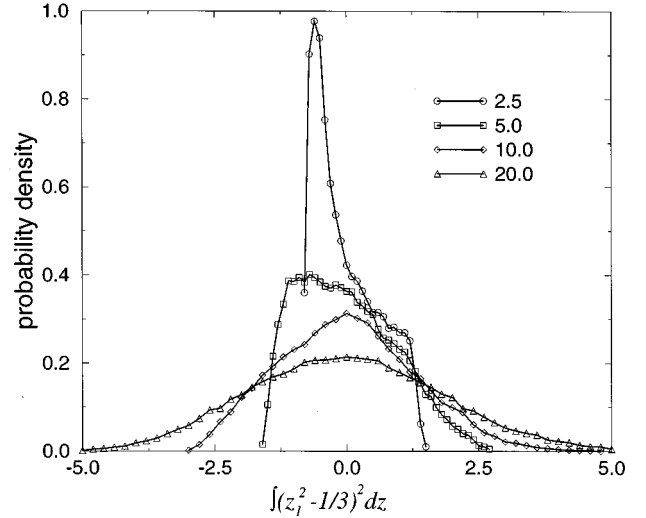

Fig. 2. Numerically determined probability distribution of $\int(z_1{}^2 - 1/3)dz$ for $l/h_{\text{fiber}} = 2.5$, 5.0, 10.0, and 20.0, showing the approach to a Gaussian distribution for large l. Here $L_B/h_{\text{fiber}} = 10$ (so $h_E \approx 0.84 h_{\text{fiber}}$), and each curve is the result of 50,000 runs.

the asymmetric steady-state distribution of $z_1{}^2 - 1/3$, but for longer distances it relatively quickly becomes Gaussian. Other values of the ratio L_B/h_{fiber} produce similar results.

5. NONLINEAR POLARIZATION-MODE DISPERSION STATISTICS WHEN $h_{\text{fiber}} \gg L_B$

In the limit $h_{\text{fiber}} \gg L_B$, or equivalently $\sigma_\theta{}^2 \ll b$, it is convenient to let

$$\alpha = \langle x_j x_k \rangle - 1/2 \langle y_j y_k + z_j z_k \rangle, \quad (53a)$$

$$\beta = \langle y_j y_k - z_j z_k \rangle, \quad (53b)$$

$$\gamma = \langle y_j z_k + y_k z_j \rangle, \quad (53c)$$

which yields the equation

$$\frac{d}{dz}\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{bmatrix} -3/2\sigma_\theta{}^2 & 3/4\sigma_\theta{}^2 & 0 \\ \sigma_\theta{}^2 & -1/2\sigma_\theta{}^2 & -4b \\ 0 & 4b & -1/2\sigma_\theta{}^2 \end{bmatrix}\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \quad (54)$$

In addition, using $x_j x_k + y_j y_k + z_j z_k = \delta_{jk}$, we obtain

$$\langle z_j z_k \rangle = 1/3 \delta_{jk} - 1/3\alpha - 1/2\beta. \quad (55)$$

Equation (54) is more convenient to analyze than Eq. (5) when $\sigma_\theta{}^2 \ll b$ because in the (α, β, γ) coordinates it is easier to separate the rapid oscillation about the x_j axis produced by the large birefringence b from the relatively slower variations caused by the random rotations about the z_j axis. In the new coordinates the evolution is a rapid rotation of β and γ about the α axis combined with a slow exponential decay of all three components. This effect can be seen more explicitly when we make the transformation

$$\begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{bmatrix} \cos 4\beta\xi & -\sin 4b\xi \\ \sin 4b\xi & \cos 4b\xi \end{bmatrix}\begin{pmatrix} \beta' \\ \gamma' \end{pmatrix}, \quad (56)$$

where $\xi = z - z_0$, which yields

$$\frac{d\alpha}{d\xi} = -{}^3/_2\sigma_\theta{}^2\alpha + {}^3/_4\sigma_\theta{}^2(\beta'\cos 4b\xi - \gamma'\sin 4b\xi),$$
(57a)

$$\frac{d\beta'}{d\xi} = \sigma_\theta{}^2\alpha\cos 4b\xi - {}^1/_2\sigma_\theta{}^2\beta',$$
(57b)

$$\frac{d\gamma'}{d\xi} = -\sigma_\theta{}^2\sin 4b\xi - {}^1/_2\sigma_\theta{}^2\gamma'.$$
(57c)

Inasmuch as $b \gg \sigma_\theta{}^2$, the sinusoidal terms in Eqs. (57) oscillate rapidly and contribute only a small amount to the solution, so that

$$\alpha \approx \alpha(z_0)\exp(-{}^3/_2\sigma_\theta{}^2\xi),$$
(58a)

$$\beta' \approx \beta(z_0)\exp(-{}^1/_2\sigma_\theta{}^2\xi),$$
(58b)

$$\gamma' \approx \gamma(z_0)\exp(-{}^1/_2\sigma_\theta{}^2\xi).$$
(58c)

Thus Eq. (55) yields

$$\langle z_j z_k | (x_l, y_l, z_l)_{z_0}\rangle - {}^1/_3\delta_{jk}$$

$$= -{}^1/_3[x_j x_k - {}^1/_2(y_j y_k + z_j z_k)]_{z_0}\exp(-{}^3/_2\sigma_\theta{}^2\xi)$$

$$- {}^1/_2[y_j y_k - z_j z_k]_{z_0}\exp(-{}^1/_2\sigma_\theta{}^2\xi)\cos 4b\xi$$

$$+ {}^1/_2[y_j z_k + y_k z_j]_{z_0}\exp(-{}^1/_2\sigma_\theta{}^2)\sin 4b\xi.$$
(59)

Equation (59) must be integrated with respect to ξ, and the rapidly oscillating terms do not contribute substantially to the integral. Therefore it is appropriate to use as the integrand the averaged solution

$$\langle z_j z_k | (x_l, y_l, z_l)_{z_0}\rangle - {}^1/_3\delta_{jk}$$

$$\approx {}^1/_2({}^1/_3\delta_{jk} - x_j x_k)_{z_0}\exp(-{}^3/_2\sigma_\theta{}^2\xi).$$
(60)

Equation (60) does not provide a good pointwise approximation to the autocorrelation, because in general $R(\xi)$ will combine an exponential decay with a rapid oscillation, and Eq. (60) gives only the exponential decay. When Eq. (59) is integrated, however, it is only the exponential decay that contributes to the nonlinear PMD variances at leading order, because the rapid oscillations average out. Note it is the rapid rotation around the x_j axis induced by the relatively large birefringence b that produces the rapid oscillations and which causes the average value of $\langle z_j z_k\rangle$ to depend only on $x_j x_k$, or equivalently on $y_j y_k + z_j z_k$, at z_0. The rapid oscillation is also what leads to the factor of 1/2 that appears in Eq. (60). Essentially, first the rapid rotation about the x_j axis produces strong mixing in the y_j and z_j directions and then the slower random rotations about the z_j axis cause variations in the x_j direction to grow.

Using Eqs. (30), (37), and (60), we can find how expression (32) is modified, namely,

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l}(z_j z_k - {}^1/_3)d\zeta\right]$$

$$= 2R(0)h_E{}^2\left[\frac{l}{h_E} + 4\exp\left(-\frac{l}{4h_E}\right) - 4\right],$$
(61)

where $h_E = 1/(6\sigma_\theta{}^2) = {}^1/_{12}h_{\text{fiber}}$ is the field decorrelation length defined by Eq. (35) and $R(0) = \langle[x_j x_k - {}^1/_3\delta_{jk}]^2\rangle = \langle[z_j z_k - {}^1/_3\delta_{jk}]^2\rangle$. The extra factor of 4 comes from the 1/2 in Eq. (60). As a result,

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l}(z_1{}^2 - {}^1/_3)d\zeta\right] \sim {}^8/_{45}h_E(l - 4h_E),$$
(62)

$$\mathrm{Var}\left(\int_{z_0}^{z_0+l}z_j z_k d\zeta\right) \sim {}^2/_{15}h_E(l - 4h_E)$$
(63)

for $(j, k) = (1, 2)$, $(j, k) = (1, 3)$, or $(j, k) = (2, 3)$, and

$$\mathrm{Var}\left[\int_{z_0}^{z_0+l}(z_2{}^2 - z_3{}^2)d\zeta\right] \sim {}^8/_{15}h_E(l - 4h_E).$$
(64)

As in Section 4, we find that all the asymptotic covariances of the integrated nonlinear PMD terms are zero.

Figure 3 shows comparisons between the numerical estimate for the variance of $\int(z_1{}^2 - {}^1/_3)dz$ obtained by solution of Eq. (5) and the scaled theoretical result, Eq. (61) with $\sigma_E{}^2 = 2R(0)h_E{}^2$, for several values of h_{fiber}/L_B. The agreement is seen to be very good when h_{fiber}/L_B is large, but the effect of the oscillatory terms that appear in Eqs. (56) and (57) becomes evident when this ratio decreases, apparently because when it is not large the terms do not properly average out. As was found in Section 4 for the case $h_{\text{fiber}} \ll L_B$, the numerically determined curves for the variances of the other integrated coefficients are almost identical to those shown here, so when $h_{\text{fiber}} \gg L_B$ the variances are also in ratio to one another in the amounts ${}^8/_{45}:{}^2/_{15}:{}^8/_{15}$, as expected. Similarly, the numerical results for the means and covariances of the various integrals were found to be negligibly small.

An attempt can be made at constructing a uniform approximation for the variance of the integrated coefficients by combining the results obtained in Section 4 for the case in which h_{fiber}/L_B is small with the results obtained above
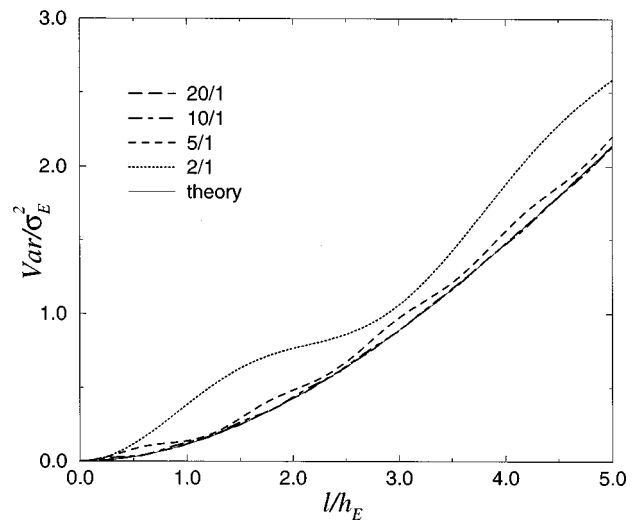

Fig. 3. Comparison between numerical and theoretical estimates for the variance of $\int_0^l(z_1{}^2 - {}^1/_3)dz$ for several values of h_{fiber}/L_B when the ratio is large. Each curve represents the results of 5000 solutions of Eq. (5), and the theoretical curve is Eq. (61) with $R(0) = 4/45$ and $h_E = {}^1/_{12}h_{\text{fiber}}$.
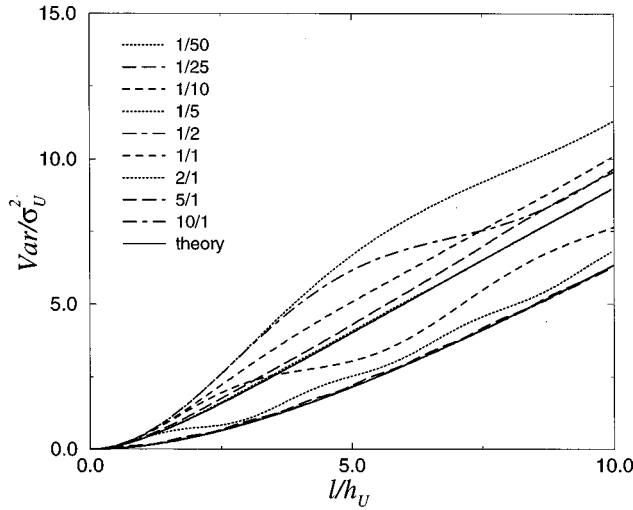
Fig. 4. Comparison between numerical and uniform theoretical estimates for the variance of $\int_0^l(z_1{}^2 - {}^1\!/\!_3)\mathrm{d}z$ for several values of h_{fiber}/L_B. Each curve represents the results of 5000 solutions of Eq. (5). In this case two theoretical curves are shown, namely, Eqs. (51) and (61), with $R(0) = 4/45$ and h_E replaced by h_U. The uniform mixing length h_U is given by Eq. (65).

for the case in which h_{fiber}/L_B is large. A proper uniform treatment would involve determining the full solution of Eqs. (24) and then averaging over all initial conditions to obtain the autocorrelation function, but this cannot be done easily because the closed-form solution of Eqs. (24) is extremely complicated. We can construct a first guess at such a uniform approximation by defining a uniform mixing length:

$$h_U = \frac{L_B{}^2}{12\pi^2 h_{\text{fiber}}} + \frac{1}{12}\, h_{\text{fiber}}, \qquad (65)$$

and an associated variance scale $\sigma_U{}^2 = 2R(0)h_U{}^2$. These expressions have the correct behavior in the two limits and should be at least close to correct between the limits. Figure 4 shows the results of comparing Eq. (65) with the numerical solution of Eq. (5). The agreement between the two is now seen to be improved in comparison with Fig. 1. In particular, the slope of the curves for large l appears to be in good agreement for all the values of h_{fiber}/L_B, even though the transient behavior for small l and the asymptotic intercept for large l vary with h_{fiber}/L_B. The two theoretical curves (for large and small h_{fiber}/L_B) are also shown.

6. LINEAR POLARIZATION-MODE DISPERSION STATISTICS

The method outlined in Section 3 and used in Sections 4 and 5 to evaluate the magnitude of the nonlinear PMD can also be used in a slightly modified form to evaluate the amount of linear PMD. To do this we consider the z derivative and the linear perturbing terms in Eq. (1):

$$\frac{\partial \overline{\boldsymbol{\Psi}}}{\partial z} = -b' \,\overline{\boldsymbol{\sigma}}\, \frac{\partial \overline{\boldsymbol{\Psi}}}{\partial t}, \qquad (66)$$

where $\overline{\boldsymbol{\sigma}}$ is as given in Eq. (2). By integrating Eq. (66) from z_0 to $z_0 + l$, where l is small, we obtain

$$\overline{\boldsymbol{\Psi}}(z_0 + l,\, t) \approx \overline{\boldsymbol{\Psi}}(z_0,\, t) - b' \int_{z_0}^{z_0+l} \overline{\boldsymbol{\sigma}}\mathrm{d}z \, \frac{\partial \overline{\boldsymbol{\Psi}}}{\partial t}(z_0,\, t). \qquad (67)$$

In the deterministic case, the linear term merely produces a polarization-dependent shift in a pulse's position. Assuming this also to be true when the birefringence is random, we let $\overline{\boldsymbol{\Psi}}(z_0,\, t) = f(t)\mathbf{U}$, where $|\mathbf{U}|^2 = 1$ and $\overline{\boldsymbol{\Psi}}(z_0 + l,\, t) = f(t - \Delta\tau)\mathbf{U}$. Expanding this expression for $\Delta\tau \ll 1$, we therefore must have

$$\Delta\tau\mathbf{U} = b' \int_{z_0}^{z_0+l} \overline{\boldsymbol{\sigma}}\mathrm{d}z\,\mathbf{U}. \qquad (68)$$

Equation (68) is an eigenvalue condition for both the time delay $\Delta\tau$ and the accompanying polarization state $\mathbf{U}$. From Eq. (2), the characteristic equation is just

$$\Delta\tau^2 = b'^2 \sum_{j=1}^{3} \left(\int_{z_0}^{z_0+l} x_j \mathrm{d}z \right)^2. \qquad (69)$$

The magnitude of the linear PMD, τ_D, is defined to be $\Delta\tau_{\max} - \Delta\tau_{\min}$, or $\tau_D = 2\Delta\tau$. Therefore

$$\langle \tau_D{}^2 \rangle = 4b'^2 \sum_{j=1}^{3} \left\langle \left(\int_{z_0}^{z_0+l} x_j \mathrm{d}z \right)^2 \right\rangle. \qquad (70)$$

Equation (23a) gives $\langle x_1 | x_1(z_0) \rangle = x_1(z_0)\exp(-{}^1\!/\!_2\sigma_\theta{}^2\xi)$, with similar results for x_2 and x_3. As $\langle x_j \rangle = 0$, expression (32) yields

$$\left\langle \left(\int_{z_0}^{z_0+l} x_1 \mathrm{d}z \right)^2 \right\rangle = {}^2\!/\!_3 h_E{}^2 \left[\frac{l}{h_E} + \exp\!\left(-\frac{l}{h_E} \right) - 1 \right]$$

$$\sim {}^2\!/\!_3 h_E(l - h_E), \qquad (71)$$

where $h_E = 2/\sigma_\theta{}^2 = h_{\text{fiber}}$ and $\langle x_1{}^2 \rangle = 1/3$. Adding the contributions in Eq. (70) from x_1, x_2, and x_3 (which are statistically independent of one another by the same argument as that used for the nonlinear PMD terms), we finally obtain

$$\langle \tau_D{}^2 \rangle = 8b'^2 h_E{}^2 \left[\frac{l}{h_E} + \exp\!\left(-\frac{l}{h_E} \right) - 1 \right], \qquad (72)$$

which is identical to results obtained previously.[16–19]

The long-term behavior of the linear PMD is seen to be independent of whether the birefringence beat length is larger or smaller than the fiber decorrelation length, $h_{\text{fiber}} = 2/\sigma_\theta{}^2$. Because the linear PMD arises from just the x_j coefficients, the rate of its growth depends only on the rate of longitudinal mixing on the Poincaré sphere. This rate is always proportional to $\sigma_\theta{}^2$, independently of whether $b \ll \sigma_\theta{}^2$ or $b \gg \sigma_\theta{}^2$. Figure 5 shows the numerical estimate for the variance of $\int x_1 \mathrm{d}z$ obtained by solution of Eq. (5) for nine values of h_{fiber}/L_B. The agreement with the theoretical result, relation (71), is seen to be good for all values of h_{fiber}/L_B.
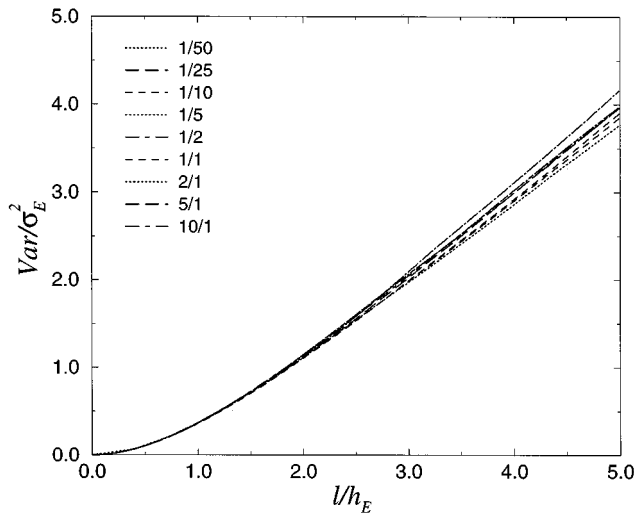
Fig. 5.   Numerical estimates for the variance of $\int x_1 \mathrm{d}z$ for various values of h_{fiber}/L_B.   Each curve represents the results of 5000 solutions of Eq. (5).   Here $\sigma_E^2 = {}^2\!/{}_3 h_E^2$.

7. POLARIZATION-STATE PROBABILITY-DISTRIBUTION FUNCTION WHEN $h_{\text{fiber}} \ll L_B$

Rather than working with the means and the variances of the various linear and nonlinear PMD coefficients, one can alternatively work with the Fokker–Planck equation for the polarization state's probability-distribution function.   When we omit the distribution with respect to θ, the appropriate generator for the probability distribution P associated with Eq. (10) is[21]

$$\frac{\partial P}{\partial z} = \left[\frac{1}{2}\,\sigma_\theta{}^2\left(\widetilde{S}_2\,\frac{\partial}{\partial \widetilde{S}_1} - \widetilde{S}_1\,\frac{\partial}{\partial \widetilde{S}_2}\right)^2 - 2b\left(\widetilde{S}_3\,\frac{\partial}{\partial \widetilde{S}_1} - \widetilde{S}_1\,\frac{\partial}{\partial \widetilde{S}_3}\right)\right]P. \tag{73}$$

Note that the first operator in parentheses corresponds to flow about the $\widetilde{S}_3$ axis, so that the squared operator is a diffusion process, and that the second operator corresponds to flow about the $\widetilde{S}_2$ axis.   This point can be made more explicit if we transform to spherical coordinates, using

$$\widetilde{S}_1 = R\,\sin\,\phi\,\cos\,\psi, \qquad \widetilde{S}_2 = R\,\sin\,\phi\,\sin\,\psi,$$

$$\widetilde{S}_3 = R\,\cos\,\phi. \tag{74}$$

With these variables, Eq. (73) becomes

$$\frac{\partial P}{\partial z} = \frac{1}{2}\,\sigma_\theta{}^2\,\frac{\partial^2 P}{\partial \psi^2} - 2b\left(\frac{\cos\,\phi\,\cos\,\psi}{\sin\,\phi}\,\frac{\partial P}{\partial \psi} + \sin\,\psi\,\frac{\partial P}{\partial \phi}\right), \tag{75}$$

and the diffusion about the $\widetilde{S}_3$ axis in the ψ direction is now clearly shown.

Equation (75) cannot be solved in general, but we can exploit widely different length scales, such as when $\sigma_\theta{}^2 \gg b$ or $\sigma_\theta{}^2 \ll b$, to obtain approximate analytical solutions.[28]   First, we consider the case $\sigma_\theta{}^2 \gg b$.   We let $^1\!/_2\sigma_\theta{}^2 z = \zeta$ and define $\epsilon = 4b/\sigma_\theta{}^2$.   We also expand the probability distribution in a multiple-scale perturbation expansion that separates the rapidly and slowly varying parts of the evolution[28,31]:

$$P = P_0(\phi, \psi, \zeta, \widetilde{\zeta}) + \epsilon P_1(\phi, \psi, \zeta, \widetilde{\zeta})$$
$$+ \epsilon^2 P_2(\phi, \psi, \zeta, \widetilde{\zeta}) + \dots, \tag{76}$$

where $\widetilde{\zeta} = \epsilon^2 \zeta$ is the slow variable.   At leading order, we find that

$$\frac{\partial P_0}{\partial \zeta} = \frac{\partial^2 P_0}{\partial \psi^2}, \tag{77}$$

which has the solution

$$P_0 = \sum_{-\infty}^{\infty} p_n(\phi, \widetilde{\zeta})\exp(in\psi - n^2\zeta). \tag{78}$$

In particular, we note that $P_0 \to p_0(\phi, \widetilde{\zeta})$ as $\zeta \to \infty$.

At order ϵ we obtain the equation

$$\frac{\partial P_1}{\partial \zeta} - \frac{\partial^2 P_1}{\partial \psi^2} = -\left(\frac{\cos\,\phi\,\cos\,\psi}{\sin\,\phi}\,\frac{\partial P_0}{\partial \psi} + \sin\,\psi\,\frac{\partial P_0}{\partial \phi}\right)$$
$$\equiv f_1(\phi, \psi, \zeta) = \sum_{-\infty}^{\infty} f_n^{(1)}\,\exp(in\psi). \tag{79}$$

We also expand P_1 in a Fourier series in ψ, and all the terms remain bounded as $\zeta \to \infty$, except possibly for $n = 0$.   The $n = 0$ term will grow linearly with ζ unless a solvability condition is satisfied.   In general, at $O(\epsilon^k)$ this condition is

$$\lim_{\zeta \to \infty} f_0^{(k)}(\phi, \zeta) = \lim_{\zeta \to \infty} \frac{1}{2\pi} \int_0^{2\pi} f_k(\phi, \psi, \zeta)\mathrm{d}\psi = 0. \tag{80}$$

At order ϵ with $k = 1$ this condition is automatically satisfied, but in general the linear growth with ζ is not desirable because it will disorder the perturbation expansion by producing terms proportional to $\epsilon^n \zeta$, where n is the order at which the linear growth first appears.   Such behavior can be thought of as arising from a nonuniform perturbation expansion in powers of ϵ.   For a uniform result one does not want to expand any dependencies of the form $\epsilon^n \zeta$.   The multiple-scale expansion takes care of this by defining new slow variables that keep such combinations of ϵ and ζ from being separated.   We can find the steady-state behavior at $O(\epsilon)$ by taking the limit $\zeta \to \infty$ in Eq. (79), which yields

$$-\frac{\partial^2 P_1}{\partial \psi^2} = -\sin\,\psi\,\frac{\partial p_0}{\partial \phi}, \tag{81}$$

so that, as $\zeta \to \infty$,

$$P_1 \to -\sin\,\psi\,\frac{\partial p_0}{\partial \phi}. \tag{82}$$

At $O(\epsilon^2)$, the equation is

$$\frac{\partial P_2}{\partial \zeta} - \frac{\partial^2 P_2}{\partial \psi^2} = \frac{\partial P_0}{\partial \widetilde{\zeta}} - \left(\frac{\cos \phi \cos \psi}{\sin \phi} \frac{\partial P_1}{\partial \psi} + \sin \psi \frac{\partial P_1}{\partial \phi} \right)$$

$$\equiv f_2(\phi, \psi, \zeta) = \sum_{-\infty}^{\infty} f_n^{(2)} \exp(in\psi). \qquad (83)$$

As before, this equation will have unbounded solutions as $\zeta \to \infty$ unless the solvability condition, Eq. (80) with $k = 2$, is satisfied. Applying this condition, we obtain

$$\frac{\partial p_0}{\partial \widetilde{\zeta}} = \frac{1}{2} \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial p_0}{\partial \phi} \right). \qquad (84)$$

Equation (77) describes the short-distance evolution of the probability distribution on an $O(1)$ length scale, whereas Eq. (84) describes the long-distance evolution on an $O(1/\epsilon^2)$ length scale. It is also possible to combine the two to arrive at an equation that uniformly describes the evolution of the probability distribution. We reverse the multiple-scale expansion, Eq. (76), by noting that

$$\frac{\partial P}{\partial \zeta} = \frac{\partial P_0}{\partial \zeta} + \epsilon^2 \frac{\partial P_0}{\partial \widetilde{\zeta}} + \dots \qquad (85)$$

and by also noting that the $\widetilde{\zeta}$ derivative is important only when ζ is large, in which case $P_0 \approx p_0$. Adding the two equations, using $P = P_0 + \dots$, and converting from ζ back to z, we thus obtain the approximate equation

$$\frac{\partial P}{\partial z} = \frac{1}{2} \sigma_\theta^2 \frac{\partial^2 P}{\partial \psi^2} + \frac{4b^2}{\sigma_\theta^2} \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial P}{\partial \phi} \right). \qquad (86)$$

Equation (86) can be solved exactly by separation of variables.[32]

The multiple-scale perturbation solution of Eq. (86) is uniformly close to that of Eq. (75); i.e., the error between the two solutions is $O(\epsilon)$ uniformly for all ζ. In addition, Eq. (86) clearly shows the rapid mixing of the probability distribution in longitude (ψ) and the relatively slow mixing with respect to latitude (ϕ) that occur when $b \ll \sigma_\theta^2$. This equation is a Fokker–Planck equation in its own right, which allows us to replace the original stochastic differential equation for the PMD coefficients, Eq. (10), with another equation that generates the same probability distribution with only small errors. This equation is

$$\frac{d}{dz} \begin{pmatrix} \widetilde{S}_1 \\ \widetilde{S}_2 \\ \widetilde{S}_3 \end{pmatrix} = \begin{bmatrix} 0 & g_\theta & g_1 \\ -g_\theta & 0 & g_2 \\ -g_1 & -g_2 & 0 \end{bmatrix} \begin{pmatrix} \widetilde{S}_1 \\ \widetilde{S}_2 \\ \widetilde{S}_3 \end{pmatrix}, \qquad (87)$$

where $g_\theta(z)$ is the same white-noise process as defined in Eqs. (19) and $g_1(z)$ and $g_2(z)$ are two additional independent white-noise processes that satisfy

$$\langle g_j(z) \rangle = 0, \qquad \langle g_j(z) g_k(z') \rangle = \frac{8b^2}{\sigma_\theta^2} \delta(z - z') \delta_{jk}. \qquad (88)$$

Equation (87) generates the Fokker–Planck equation[21]

$$\frac{\partial P}{\partial z} = \left[\frac{1}{2} \sigma_\theta^2 + \frac{4b^2}{\sigma_\theta^2} \left(\frac{1}{\sin^2 \phi} - 1 \right) \right] \frac{\partial^2 P}{\partial \psi^2}$$

$$+ \frac{4b^2}{\sigma_\theta^2} \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial P}{\partial \phi} \right). \qquad (89)$$

For $\sigma_\theta^2 \gg b$, the second term inside the brackets in Eq. (89) is negligible compared with the first, except possibly for a small region near the poles of the Poincaré sphere, and therefore Eq. (89) is for all intents and purposes the same as Eq. (86).

Equation (87) can be used as the starting point for an alternative numerical solution of Eq. (10). Rather than combining a deterministic rotation about the $\widetilde{S}_1$ axis at rate $2b$ with a random rotation about the $\widetilde{S}_3$ axis as indicated in Eq. (10), we can use random rotations about all three axes as indicated in Eq. (87). Because $4b^2/\sigma_\theta^2 \ll 1/2\sigma_\theta^2$, the random rotations about the $\widetilde{S}_1$ and $\widetilde{S}_2$ axes can be taken much more infrequently than those about the $\widetilde{S}_3$ axis. Also, if b/σ_θ^2 is small enough, it is even possible that the random rotations about the $\widetilde{S}_1$ and $\widetilde{S}_2$ axes can be done more infrequently than would be necessary to resolve the deterministic rotation about the $\widetilde{S}_1$ axis in Eq. (10), thus resulting in an overall computational savings.

8. POLARIZATION-STATE PROBABILITY-DISTRIBUTION FUNCTION WHEN $h_{\text{fiber}} \gg L_B$

In the limit $b \gg \sigma_\theta^2$, we first transform Eq. (10) to remove the rapid rotation produced by the relatively large birefringence b, using

$$\begin{pmatrix} \widetilde{S}_1 \\ \widetilde{S}_2 \\ \widetilde{S}_3 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2bz & \sin 2bz \\ 0 & -\sin 2bz & \cos 2bz \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}. \qquad (90)$$

In addition, we let $\zeta = 2bz$, so Eq. (10) becomes

$$\frac{d}{d\zeta} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} = \epsilon h(\zeta) \begin{bmatrix} 0 & \cos \zeta & \sin \zeta \\ -\cos \zeta & 0 & 0 \\ -\sin \zeta & 0 & 0 \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}, \qquad (91)$$

where $\langle h(\zeta) h(\zeta') \rangle = \delta(\zeta - \zeta')$ and $\epsilon = \sigma_\theta/(2\sqrt{b})$.

Equation (91) is now weakly random because of the ϵ on the right-hand side, and all the coefficients are oscillatory with zero mean. It is therefore in the proper form to be averaged.[28,33,34] To summarize the result of this averaging, if we write Eq. (91) in the form

$$\frac{dR_i}{d\zeta} = \epsilon \sum_j B_{ij}(\zeta) R_j, \qquad (92)$$

then the averaged Fokker–Planck equation is[28,33,34]

$$\frac{dP}{d\zeta} = \epsilon^2 \sum_{i,j,k,l} \langle B_{ij} B_{kl} \rangle R_j \frac{\partial}{\partial R_i} \left(R_l \frac{\partial P}{\partial R_k} \right). \qquad (93)$$

When we use the specific form of B_{ij} and change back to z derivatives, Eq. (93) becomes

$$\frac{dP}{dz} = \frac{\sigma_\theta^2}{4} \left[\left(R_1 \frac{\partial}{\partial R_2} - R_2 \frac{\partial}{\partial R_1} \right)^2 \right.$$
$$\left. + \left(R_1 \frac{\partial}{\partial R_3} - R_3 \frac{\partial}{\partial R_1} \right)^2 \right] P. \tag{94}$$

We can solve Eq. (94) exactly by first converting to the spherical coordinates $R_1 = \cos\varphi$, $R_2 = \sin\varphi\cos\vartheta$, and $R_3 = \sin\varphi\sin\vartheta$ and then using separation of variables.[32]

This averaged Fokker–Planck equation can be thought of as arising from independent random rotations about the R_2 and R_3 axes; i.e.,[32]

$$\frac{d}{dz} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} = \begin{bmatrix} 0 & g_1 & g_2 \\ -g_1 & 0 & 0 \\ -g_2 & 0 & 0 \end{bmatrix} \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}, \tag{95}$$

where $g_1(z)$ and $g_2(z)$ satisfy

$$\langle g_j(z) \rangle = 0, \qquad \langle g_j(z) g_k(z') \rangle = (\sigma_\theta^2/2) \delta(z - z') \delta_{jk}. \tag{96}$$

Because the analytical transformation, Eq. (90), is deterministic, it should be possible to solve Eq. (95) numerically when $b \gg \sigma_\theta^2$ with less computational effort than would be required to solve Eq. (10), as the rapid rotation about the $\widetilde{S}_1$ axis would no longer need to be resolved numerically.

9. CONCLUSIONS

In this paper we have calculated the means and the variances of the nonlinear PMD fluctuations produced by random birefringence in an optical fiber. It was shown that, for distances that are long compared with the polarization decorrelation length h_E, the nonlinear PMD corrections have a Gaussian distribution with zero mean and a variance given by $2R(0)h_E l$, where $R(0)$ is an $O(1)$ constant ranging from 1/15 to 4/15, depending on the particular correction term being considered. The mixing length h_E was found to be $L_B^2/(12\pi^2 h_{\text{fiber}})$ when $h_{\text{fiber}} \ll L_B$ and $h_{\text{fiber}}/12$ when $h_{\text{fiber}} \gg L_B$.

Situations in which the nonlinear PMD is not negligible occur when the mixing length h_E is not small in comparison with the dispersive or nonlinear length scales, so that incomplete mixing on the Poincaré sphere results. As fiber decorrelation lengths for standard communication fiber are typically in the range 0.3–300 m and beat lengths range from 10 to 100 m, such a situation is not likely to occur in the limit $h_{\text{fiber}} \gg L_B$, where $h_E = h_{\text{fiber}}/12$, because the resulting mixing lengths are quite short. Even when $h_{\text{fiber}} \ll L_B$ and extreme values for h_{fiber} and L_B are considered, the resulting amount of nonlinear PMD is still negligible under conditions typical for practical, present-day communication systems. For example, for $h_{\text{fiber}} = 0.3$ m and $L_B = 100$ m, one obtains $h_E = 280$ m. This is much shorter than present-day dispersive and nonlinear scale lengths, which are typically hundreds of kilometers.

Nonlinear PMD effects are much more likely to become important in proposed very-high-data-rate transmission systems operating in the neighborhood of 100 Gbits/s, however, particularly in those that employ solitons.

First, in such systems pulse widths must be smaller by an order of magnitude, and thus the dispersive and nonlinear length scales are in the kilometer range. Second, with shorter pulse widths such systems are more susceptible to effects such as linear PMD. Equation (72) shows that reducing linear PMD, however, means either increasing the beat length L_B or reducing the fiber decorrelation length h_{fiber}. Increasing the beat length is equivalent to reducing the magnitude of the birefringence, which can be done by more careful control of the optical fiber during manufacture to reduce perturbations from the ideal circular state. Decreasing the fiber decorrelation length means shortening the length scale associated with such variations, which is likely what occurs when the fiber is spun during the manufacturing process, a method that has already been used to reduce linear PMD.[22]

Reducing linear PMD by either method, however, has the side effect of increasing the polarization decorrelation length h_E and with it the amount of nonlinear PMD. In essence, decreasing the amount of linear PMD also reduces the rate of mixing on the Poincaré sphere, which in turn increases the nonlinear PMD. As an example, if h_{fiber} is merely reduced from 30 to 5 cm by the spinning process (which does not seem unreasonable because spin rates in the range of 20–50 turns/m have been reported), then h_E increases to 1.5 km, which is now comparable with the dispersive and nonlinear length scales in a high-bit-rate transmission system. Recent research by Arend *et al.*[35] shows experimental evidence of this effect. As a result, we conclude that, although nonlinear PMD poses no threat to standard communication systems, it will in principle affect very-high-data-rate systems and may make it impossible to deploy them in practice.

*Permanent address, Department of Electronic Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

†Permanent address, Department of Engineering Sciences and Applied Mathematics, McCormick School of Engineering and Applied Science, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208-3125.

REFERENCES

1. D. Marcuse, *Light Transmission Optics*, 2nd ed. (Van Nostrand Reinhold, New York, 1982).
2. G. P. Agrawal, *Nonlinear Fiber Optics*, 2nd ed. (Academic, San Diego, Calif., 1995).
3. P. K. A. Wai, C. R. Menyuk, Y. C. Lee, and H. H. Chen, "Nonlinear pulse propagation in the neighborhood of the zero-dispersion wavelength of monomode optical fibers," Opt. Lett. **11**, 464–466 (1986).
4. C. D. Poole, J. M. Wiesenfeld, A. R. McCormick, and K. T. Nelson, "Broadband dispersion compensation by using the higher-order spatial mode in a two-mode fiber," Opt. Lett. **17**, 985–987 (1992).
5. A. Yariv, D. Fekete, and D. M. Pepper, "Compensation for channel dispersion by nonlinear optical phase conjugation," Opt. Lett. **4**, 52–54 (1979).
6. N. Bergano, "Circulating loop transmission experiments for the study of long-haul transmission systems using erbium-doped fiber amplifiers," J. Lightwave Technol. **13**, 879–888 (1995).
7. A. Hasegawa and Y. Kodama, *Solitons in Optical Communications* (Oxford U. Press, Oxford, 1995).
8. J. P. Gordon and H. A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission," Opt. Lett. **11**, 665–667 (1986).
9. A. Mecozzi, J. Moores, H. Haus, and Y. Lai, "Soliton transmission control," Opt. Lett. **16**, 1841–1843 (1991).
10. Y. Kodama and A. Hasegawa, "Generation of asymptotically stable optical solitons and suppression of the Gordon–Haus effect," Opt. Lett. **17**, 31–33 (1992).
11. C. R. Menyuk, "Nonlinear pulse propagation in birefringent optical fibers," IEEE J. Quantum Electron. **23**, 174–176 (1987).
12. C. R. Menyuk, "Stability of solitons in birefringent optical fibers. I. Equal propagation amplitudes," Opt. Lett. **12**, 614–616 (1987).
13. C. R. Menyuk, "Stability of solitons in birefringent optical fibers. II. Arbitrary amplitudes," J. Opt. Soc. Am. B **5**, 392–402 (1988).
14. P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," Opt. Lett. **16**, 1231–1233 (1991).
15. S. G. Evangelides, Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," J. Lightwave Technol. **10**, 28–35 (1992).
16. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," Opt. Lett. **13**, 687–689 (1988).
17. C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," Opt. Lett. **14**, 523–525 (1989).
18. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," J. Lightwave Technol. **9**, 1439–1456 (1991).
19. C. R. Menyuk and P. K. A. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," J. Opt. Soc. Am. B **11**, 1288–1296 (1994).
20. P. K. A. Wai, and C. R. Menyuk, "Anisotropic diffusion of the state of polarization in optical fibers with randomly varying birefringence," Opt. Lett. **20**, 2493–2495 (1995).
21. P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation and diffusion in optical fibers with randomly varying birefringence," J. Lightwave Technol. **14**, 148–157 (1996).
22. A. J. Barlow and J. J. Ramskov-Ha, "Birefringence and polarization mode-dispersion in spun single-mode fibers," Appl. Opt. **20**, 2962–2968 (1981).
23. D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to the studies of signal propagation in optical fibers with randomly varying birefringence," J. Lightwave Technol. (to be published).
24. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980).
25. P. K. A. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," Opt. Lett. **19**, 1517–1519 (1994).
26. A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer-Verlag, New York, 1983).
27. N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).
28. G. C. Papanicolaou, "Introduction to the asymptotic analysis of stochastic equations," in *Modern Modeling of Continuum Phenomena*, R. C. DiPrima, ed., Vol. 16 of Lectures in Applied Mathematics (American Mathematical Society, Providence, R.I., 1977), pp. 109–147.
29. M. H. DeGroot, *Probability and Statistics* (Addison-Wesley, Menlo Park, Calif., 1975).
30. L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, N.J., 1962).
31. J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics* (Springer-Verlag, New York, 1981).
32. T. Ueda and W. L. Kath, "Dynamics of pulses in randomly birefringent nonlinear optical fibers," Physica D **55**, 166–181 (1992).
33. T. Ueda and W. L. Kath, "Stochastic simulation of pulses in nonlinear optical fibers with random birefringence," J. Opt. Soc. Am. B **11**, 818–825 (1994).
34. G. Papanicolaou and J. B. Keller, "Stochastic differential equations with applications to random harmonic oscillators and wave propagation in random media," SIAM J. Appl. Math. **21**, 287–305 (1971).
35. M. F. Arend, M. L. Dennis, I. N. Duling III, E. A. Golovchenko, A. N. Pilipetskii, and C. R. Menyuk, "Simulations of a nonlinear optical loop mirror demultiplexor using random birefringence fiber," in *Optical Fabrication and Testing*, Volume 6 of 1996 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1996), pp. 226–227.