Transient evolution of the polarization-dispersion vector's probability distribution

Yu Tan and Jianke Yang

Department of Mathematics and Statistics, University of Vermont, 16 Colchester Avenue, Burlington, Vermont 05401

William L. Kath

Department of Engineering Sciences and Applied Mathematics, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208

Curtis M. Menyuk

Department of Computer Science and Electrical Engineering, TRC 201A, University of Maryland Baltimore County, Baltimore, Maryland 21228, and Laboratory for Telecommunications Research, c/o U.S. Army Research Laboratory, Building 601, Room 131, 2800 Powder Mill Road, Adelphi, Maryland 20755-1197

Received July 30, 2001; revised manuscript received November 5, 2001

We determine the transient evolution of the probability distribution of the polarization dispersion vector both analytically and numerically, using a physically reasonable model of the fiber birefringence. We show that, for all practical birefringence parameters, the distribution of the differential group delay (DGD), which is the magnitude of the polarization dispersion vector, becomes Maxwellian in just a few kilometers, except in the tail region, where the DGD is large. In this limit, the approach to a Maxwellian distribution takes much longer, of the order of tens of kilometers. In addition, we show that in the transient regime the DGD distribution is very different from Maxwellian. We also find that the probability-distribution function for the polarization-dispersion vector at the output of the fiber depends upon the angle between it and the local birefringence vector on the Poincaré sphere, showing that the DGD remains correlated with the orientation of the local birefringence axes over arbitrarily long distances.

OCIS codes: 060.0060, 260.5430.

1. INTRODUCTION

Polarization mode dispersion (PMD) is caused by the random birefringence present in optical fibers. It can lead to pulse spreading and depolarization, and it is detrimental to system performance. As transmission rates continue to increase, PMD has become a major impairment, thus motivating extensive experimental and theoretical study over the past few years.

PMD is characterized by a three-component polarization dispersion vector $\mathbf{\Omega}$. Its magnitude $|\mathbf{\Omega}|$ gives the differential group delay (DGD) between the principle states, and its direction gives the orientation of the slow principle state of polarization at the output on the Poincaré sphere.[1,2] For short distances, PMD is deterministic, and the DGD probability distribution is a δ function. For long distances, however, previous research using a weak random birefringence model has shown that the three components of the vector $\mathbf{\Omega}$ are independent and Gaussian distributed, so that the DGD distribution is Maxwellian.[1] Similar results are also obtained if one assumes that the fiber birefringence completely randomizes the polarization state over the Poincaré sphere.[3]

Both analysis and numerous numerical and experimental studies have led to the generally accepted wisdom that the asymptotic (long-length limit) distribution function for the DGD due to PMD is Maxwellian. The transient behavior of the distribution, however, is not as well elucidated. Models based upon an analogy with Brownian motion[4] or upon the assumption of random mode coupling[5] have produced distributions with average characteristics that are consistent with more detailed stochastic analysis,[1,6] but experimental results have shown that the actual transient probability distributions themselves do not necessarily follow such models.[7] In particular, if the total fiber length is of the same order as the fiber correlation length, the DGD distribution appears skewed toward larger DGD values. If the fiber length is ten times longer than the correlation length, however, the main part of the DGD distribution appears to be well approximated by a Maxwellian.[7]

While these results are significant, there have been no careful studies of the length required for an asymptotically valid Maxwellian probability distribution to be reached. This question is especially relevant at large DGD values, in the tails of the distribution. These DGD values are the ones most significant for determining outage probabilities due to PMD, since large DGD events are the ones most likely to produce bit errors.

Numerous previous experiments, based on both direct observation[8,9] and observation of the ratio of the nonlinear self- and cross-phase modulation,[10,11] have established that optical fibers are primarily linearly birefringent. Recent experimental research[12] based upon actual measurements of a real fiber's local birefringence has fur-

0740-3224/2002/050992-09$15.00

ther explored these issues. This experimental research has shown that Wai and Menyuk's second birefringence model[6] is the one most appropriate for real fibers. This model assumes that the fiber birefringence is linear with two orthogonal components that are identical, independently distributed Gaussian random variables having a fixed correlation length with respect to distance. Because the ramifications of this birefringence model have not been fully explored, the transient approach to the asymptotic DGD distribution remains an open question.

Here we study the fundamental question of determining the probability distribution of the polarization dispersion vector at any fiber length with general fiber correlation- and beat-length parameters. We will adopt the slightly simpler first birefringence model of Wai and Menyuk.[6] Although the second birefringence model of Wai and Menyuk[6] is more realistic in fibers,[12] Wai and Menyuk have pointed out that both models lead to nearly the same diffusion rates for the polarization states on the Poincaré sphere. Thus it follows that the evolution for the distribution predicted by the first model will be nearly the same as for the second model. In the first model, the fiber is assumed to have a linear birefringence of fixed strength $2b$, and the orientation of the birefringence axes varies randomly with distance along the fiber. Our approach is to numerically solve the Fokker–Planck equation for the probability-density function of the polarization-dispersion vector $\mathbf{\Omega}$ associated with this model, which is a nontrivial task.

The results of our analytical and numerical studies indicate that in current systems, the main part of the DGD probability distribution (i.e., the small and moderate DGD regions) becomes Maxwellian in just a few kilometers, consistent with previous experimental research.[7] The results also show that the tail distribution takes much longer (tens of kilometers) to approach a Maxwellian. Thus the widely used approach of calculating the penalties due to PMD by assuming a Maxwellian DGD distribution should be applied with caution when the fiber length is not long enough to guarantee that the transient behavior has completely died out. Specifically, it will be shown that in the transient regime, the Maxwellian distribution overestimates the probability of large DGD values. In this case, the PMD-induced penalty calculated assuming a Maxwellian distribution will therefore actually be higher than the true penalty due to PMD. In addition, our results also show that the probability-distribution function for the polarization dispersion vector at the output of the fiber depends upon the angle between it and the local-birefringence vector on the Poincaré sphere, even after a distance long enough for the transient behavior to completely decay away. Consequently, the amount of DGD remains correlated with the relative orientation between the output-principal state and the local birefringence axes over arbitrarily long distances.

2. DERIVATION OF THE POLARIZATION-MODE-DISPERSION FOKKER–PLANCK EQUATION

The basic dynamical equation for the polarization dispersion vector $\mathbf{\Omega}$ is[2,13]

$$\frac{\partial \mathbf{\Omega}(z,\omega)}{\partial z} = \frac{\partial \mathbf{W}(z,\omega)}{\partial \omega} + \mathbf{W}(z,\omega) \times \mathbf{\Omega}(z,\omega), \quad (1)$$

where the vector $\mathbf{W}$ represents the local fiber birefringence. As stated earlier, the birefringence strength $2b$ is assumed to be fixed. In addition, the rate of change of the birefringence orientation is assumed to be driven by a white-noise process.[6] In other words, $\mathbf{W} = (2b\cos\theta, 2b\sin\theta, 0)$, and $d\theta/dz = g_\theta(z)$, where

$$\langle g_\theta(z)\rangle = 0, \quad \langle g_\theta(z) g_\theta(z')\rangle = \sigma^2 \delta(z - z').$$

The parameter σ^2 is related to the fiber correlation length h_{fiber} through the relation $\sigma^2 = 2/h_{\text{fiber}}$. To help simplify the analysis, we employ the transformation $\tilde{\mathbf{\Omega}} = H_1(z)\mathbf{\Omega}$, where

$$H_1(z) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

so that the dispersion vector will be measured in a coordinate system that follows the local birefringence vector on the Poincaré sphere. With this transformation, the dynamical equation for $\tilde{\mathbf{\Omega}}$ becomes

$$\frac{\partial}{\partial z}\begin{pmatrix} \tilde{\Omega}_1 \\ \tilde{\Omega}_2 \\ \tilde{\Omega}_3 \end{pmatrix} = \begin{pmatrix} \tilde{\Omega}_2 \\ -\tilde{\Omega}_1 \\ 0 \end{pmatrix} g(z) + \begin{pmatrix} 2b' \\ -2b\tilde{\Omega}_3 \\ 2b\tilde{\Omega}_2 \end{pmatrix}. \quad (3)$$

From this stochastic differential equation, one can derive a Fokker–Planck equation[14] that describes the evolution of the probability-distribution function P associated with the local polarization dispersion vector $\tilde{\mathbf{\Omega}}$,

$$\frac{\partial P}{\partial z} + 2b'\frac{\partial P}{\partial \tilde{\Omega}_1} - 2b\left(\tilde{\Omega}_3\frac{\partial P}{\partial \tilde{\Omega}_2} - \tilde{\Omega}_2\frac{\partial P}{\partial \tilde{\Omega}_3}\right) - \frac{1}{2}\sigma^2\left(\tilde{\Omega}_2\frac{\partial}{\partial \tilde{\Omega}_1} - \tilde{\Omega}_1\frac{\partial}{\partial \tilde{\Omega}_2}\right)^2 P = 0. \quad (4)$$

We can simplify this equation by introducing the normalized variables

$$Z = \frac{1}{2}\sigma^2 z = \frac{z}{h_{\text{fiber}}}, \quad \beta = \frac{4b}{\sigma^2} = \frac{4\pi h_{\text{fiber}}}{L_B},$$

$$\tilde{\Omega}_k = \frac{\tilde{\Omega}_k}{4b'/\sigma^2} = \frac{\tilde{\Omega}_k}{2b' h_{\text{fiber}}} \quad (k = 1, 2, 3), \quad (5)$$

where the distance is scaled by the fiber correlation length h_{fiber}, and the DGD is scaled by $2b' h_{\text{fiber}}$. With a PMD coefficient of 1 ps/$\sqrt{\text{km}}$ and a fiber correlation length of 50 m, the normalizing coefficient for the DGD, $2b' h_{\text{fiber}}$, is approximately 0.16 ps. Note the PMD coefficient is defined to be[1,6,13] $\sqrt{8h_{\text{fiber}}}b'$, and the mean accumulated DGD is this coefficient times the square root of the total distance.

Using these normalized variables, we obtain the Fokker–Planck equation,

$$\frac{\partial P}{\partial Z} + \frac{\partial P}{\partial \bar{\Omega}_1} - \beta\left(\bar{\Omega}_3\frac{\partial P}{\partial \bar{\Omega}_2} - \bar{\Omega}_2\frac{\partial P}{\partial \bar{\Omega}_3}\right) - \left(\bar{\Omega}_2\frac{\partial}{\partial \bar{\Omega}_1} - \bar{\Omega}_1\frac{\partial}{\partial \bar{\Omega}_2}\right)^2 P = 0. \quad (6)$$

Note that this equation depends only on one dimensionless parameter β, which measures the relative sizes of the correlation and beat lengths. The fiber beat length L_B is usually in the range of 3 to 30 meters, and the correlation length h_{fiber} is typically between 3 and 100 meters. For these lengths, β falls between 1.2 and 400. Thus β is moderate or large for most fibers. Nevertheless, we also allow β to be small, so that our results will cover all possible fiber parameter values.

The Fokker–Planck equation (6) is difficult to solve analytically. Results are possible in the asymptotic limit of large distance, as shown in Section 3. In addition, averaged quantities can be readily obtained for all distances. First of all, we can show from Eq. (6) that the total probability is conserved,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P d\bar{\Omega}_1 d\bar{\Omega}_2 d\bar{\Omega}_3 = 1, \quad (7)$$

as expected. Secondly, we can show from Eq. (6) that the mean-square DGD is

$$\langle|\boldsymbol{\Omega}|^2\rangle = \langle|\bar{\boldsymbol{\Omega}}|^2\rangle = 2[\exp(-Z) + Z - 1], \quad (8)$$

in agreement with previous results.[6,13] Equation (8) shows that, at short distances ($Z \ll 1$), $\langle|\boldsymbol{\Omega}|^2\rangle = Z^2$, i.e., the mean-square DGD grows quadratically. At large distances ($Z \gg 1$), however, we find that $\langle|\boldsymbol{\Omega}|^2\rangle = 2Z$, showing only linear growth.

Simplification of the Fokker–Planck equation occurs when $\beta \gg 1$, i.e., the fiber correlation length is larger than the beat length. As mentioned earlier, most fibers fall into this regime. When $\beta \gg 1$, the fast rotation represented by the third term in Eq. (6) can be averaged out.[15] Here we present an alternative averaging technique that is simpler. We start with Eq. (3) and employ another variable transformation, $\hat{\boldsymbol{\Omega}} = H_2(z)\tilde{\boldsymbol{\Omega}}$, where

$$H_2(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2bz & \sin 2bz \\ 0 & -\sin 2bz & \cos 2bz \end{bmatrix}, \quad (9)$$

to recast the equation in a coordinate system that rotates with the polarization state on the Poincaré sphere. Under this transformation, Eq. (3) becomes

$$\frac{\partial}{\partial z}\begin{pmatrix}\hat{\Omega}_1\\ \hat{\Omega}_2\\ \hat{\Omega}_3\end{pmatrix} = g(z)\begin{pmatrix}\hat{\Omega}_2\cos 2bz - \hat{\Omega}_3\sin 2bz\\ -\hat{\Omega}_1\cos 2bz\\ \hat{\Omega}_1\sin 2bz\end{pmatrix} + \begin{pmatrix}2b'\\0\\0\end{pmatrix}. \quad (10)$$

The resulting Fokker–Planck equation for the probability-density function P of the polarization dispersion vector $\breve{\boldsymbol{\Omega}}$ is[14]

$$\begin{aligned}\frac{\partial P}{\partial z} = \frac{1}{2}\sigma^2\Big\{&(\hat{\Omega}_2\cos 2bz - \hat{\Omega}_3\sin 2bz)^2 P_{\hat{\Omega}_1\hat{\Omega}_1}\\ &+ \hat{\Omega}_1^2\Big(\cos^2 2bz P_{\hat{\Omega}_2\hat{\Omega}_2} + \sin^2 2bz P_{\hat{\Omega}_3\hat{\Omega}_3}\\ &- 2(\hat{\Omega}_1\hat{\Omega}_2 P)_{\hat{\Omega}_1\hat{\Omega}_2}\cos^2 2bz\\ &- 2(\hat{\Omega}_1\hat{\Omega}_3 P)_{\hat{\Omega}_1\hat{\Omega}_3}\sin^2 2bz\\ &+ \sin 4bz[(\hat{\Omega}_1\hat{\Omega}_3 P)_{\hat{\Omega}_1\hat{\Omega}_2} + (\hat{\Omega}_1\hat{\Omega}_2 P)_{\bar{\Omega}_1\hat{\Omega}_3}\\ &- \hat{\Omega}_1^2 P_{\hat{\Omega}_2\hat{\omega}_3}] + (\hat{\Omega}_1 P)_{\hat{\Omega}_1} + \cos^2 2bz(\hat{\Omega}_2 P)_{\hat{\Omega}_2}\\ &+ \sin^2 2bz(\hat{\Omega}_3 P)_{\hat{\Omega}_3} - \frac{1}{2}\sin 4bz(\hat{\Omega}_3 P_{\hat{\Omega}_2}\\ &+ \hat{\Omega}_2 P_{\hat{\Omega}_3})\Big\} - 2b' P_{\hat{\Omega}_1}. \end{aligned} \quad (11)$$

When the system is in the limit $\beta \gg 1$, i.e., $2b \gg \sigma^2/2$, only the averages of the rapidly oscillating coefficients are significant. Replacing these terms by their averages, the reduced Fokker–Planck equation is

$$\frac{\partial P}{\partial z} = \frac{\sigma^2}{4}\left[\left(\hat{\Omega}_1\frac{\partial}{\partial \hat{\Omega}_2} - \hat{\Omega}_2\frac{\partial}{\partial \hat{\Omega}_1}\right)^2 + \left(\hat{\Omega}_1\frac{\partial}{\partial \hat{\Omega}_3} - \hat{\Omega}_3\frac{\partial}{\partial \hat{\Omega}_1}\right)^2\right]P - 2b'\frac{\partial P}{\partial \hat{\Omega}_1}. \quad (12)$$

Now employing the same variable scalings (5) with $\tilde{\Omega}_k$ replaced by $\hat{\Omega}_k$ and $\bar{\omega}_k$ replaced by $\breve{\Omega}_k$, the normalized form of Eq. (12) is then

$$\frac{\partial P}{\partial Z} = \frac{1}{2}\left[\left(\breve{\Omega}_1\frac{\partial}{\partial \breve{\Omega}_2} - \breve{\Omega}_2\frac{\partial}{\partial \breve{\Omega}_1}\right)^2 + \left(\breve{\Omega}_1\frac{\partial}{\partial \breve{\Omega}_3} - \breve{\Omega}_3\frac{\partial}{\partial \breve{\Omega}_1}\right)^2\right]P - \frac{\partial P}{\partial \breve{\Omega}_1}. \quad (13)$$

It is important to note that this dimensionless equation no longer has any free parameters. Thus the probability-distribution function P is the same for all fiber parameters as long as the assumption $\beta \gg 1$ is satisfied. This result shows that even though the full Fokker–Planck equation (4) has two length scales, h_{fiber} and L_B, in the limit $L_B \ll h_{\text{fiber}}$ the beat length averages out. In this limit, therefore, only h_{fiber} is important for the evolution of the probability distribution.

It may appear that Eq. (13) is not much simpler than Eq. (6), but this is not true. Consider the spherical coordinates,

$$\breve{\Omega}_1 = \tau\cos\phi, \quad \breve{\Omega}_2 = \tau\sin\phi\cos\psi, \quad \breve{\Omega}_3 = \tau\sin\phi\sin\psi, \quad (14)$$

where the $\breve{\Omega}_1$ axis is aligned with the orientation of the birefringence vector on the Poincaré sphere, $\tau = |\breve{\Omega}|$, ϕ is the angle between $\breve{\boldsymbol{\Omega}}$ and the $\breve{\Omega}_1$ axis (i.e., ϕ is the angle between the dispersion vector and birefringence vector on

the Poincaré sphere), and ψ is the azimuthal angle. In this coordinate system, Eq. (13) becomes

$$\frac{\partial P}{\partial Z} = \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P}{\partial\phi}\right) + \frac{1}{2}\left(\frac{1}{\sin^2\phi} - 1\right)\frac{\partial^2 P}{\partial\psi^2} - \left(\cos\phi\frac{\partial P}{\partial\tau} - \frac{\sin\phi}{\tau}\frac{\partial P}{\partial\phi}\right). \quad (15)$$

This equation shows that if the initial condition is independent of the azimuthal angle ψ, then P will be independent of the azimuthal angle ψ at all distances. In this case, all ψ derivatives in Eq. (15) can be dropped, and Eq. (15) simply reduces to

$$\frac{\partial P}{\partial Z} = \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P}{\partial\phi}\right) - \left(\cos\phi\frac{\partial P}{\partial\tau} - \frac{\sin\phi}{\tau}\frac{\partial P}{\partial\phi}\right). \quad (16)$$

Since this reduced equation has one less dimension than the full Fokker–Planck equation (6), it is much easier to study, particularly in numerical simulations.

3. ASYMPTOTIC ANALYSIS OF THE REDUCED FOKKER–PLANCK EQUATION

In this section we show analytically that the DGD probability distribution determined by the reduced Fokker–Planck equation, Eq. (15) or Eq. (16), approaches a Maxwellian distribution as $Z \to \infty$. We will also demonstrate that the probability-distribution function for the PMD vector at the output of the fiber depends upon the angle ϕ between it and the local birefringence vector on the Poincaré sphere. For convenience, we will work with Eq. (16), but similar results can be obtained by use of Eq. (15) as well. The mathematical technique we will use is the standard multiple-scale perturbation method.[16]

When the distance Z is large, the DGD values τ also become large. In addition, the dependence of the probability distribution P upon τ becomes smoother and flatter so that the size of derivatives with respect to τ decreases. Thus the terms inside the last parentheses of Eq. (16) become small compared with the other terms in the equation. This observation is the basis of our perturbation analysis. To specifically examine this limit, we first let $\hat{\tau} = \epsilon\tau$, where $\hat{\tau}$ is $O(1)$ and ϵ is a small parameter.[16] Equation (16) then becomes

$$\frac{\partial P}{\partial Z} = \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P}{\partial\phi}\right) - \epsilon\left(\cos\phi\frac{\partial P}{\partial\hat{\tau}} - \frac{\sin\phi}{\hat{\tau}}\frac{\partial P}{\partial\phi}\right). \quad (17)$$

We then expand the solution of Eq. (17) using a multiple-scale perturbation series.[16]

$$P = P_0(\hat{\tau}, \phi, Z, Z_2) + \epsilon P_1(\hat{\tau}, \phi, Z, Z_2) + \epsilon^2 P_2 + \ldots, \quad (18)$$

where $Z_2 = \epsilon^2 Z$ is the slow length scale.

At the zeroth order in ϵ, we obtain

$$\frac{\partial P_0}{\partial Z} = \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P_0}{\partial\phi}\right). \quad (19)$$

The exact solution to this equation is

$$P_0 = \sum_{n=0}^{\infty} P_{0n}(\hat{\tau}, Z_2)\exp[-n(n+1)Z/2]L_n(\cos\phi), \quad (20)$$

where $L_n(x)$ is the nth-order Legendre polynomial. As we can see from this formula, all terms with $n \geq 1$ decay rapidly with distance. Thus the solution P_0 rapidly approaches the first term $P_{00}(\hat{\tau}, Z_2)$, which is independent of the angle ϕ and fast distance variable Z. Since we are considering the large distance limit, we will simply neglect the transient and take $P_0 = P_{00}(\hat{\tau}, Z_2)$ in the analysis that follows. The evolution of P_0 on the slow-distance scale Z_2 will be determined when we carry out the perturbation expansion to order ϵ^2.

At the next order, order ϵ, we have

$$\frac{\partial P_1}{\partial Z} - \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P_1}{\partial\phi}\right) = -\cos\phi\frac{\partial P_0}{\partial\hat{\tau}}. \quad (21)$$

To suppress linear growth in Z of the solution P_1, the right-hand side of the above equation must be orthogonal to $\sin\phi$, i.e.,[16]

$$\int_0^{\pi}\sin\phi\cos\phi\frac{\partial P_0}{\partial\hat{\tau}}d\phi = 0. \quad (22)$$

This condition is automatically satisfied here. Again, neglecting the initial transient, the bounded solution for P_1 is

$$P_1 = -\cos\phi\frac{\partial P_0}{\partial\hat{\tau}}. \quad (23)$$

Note that this solution is dependent of the local polarization angle ϕ on the Poincaré sphere.

At order ϵ^2, the equation for P_2 is

$$\frac{\partial P_2}{\partial Z} - \frac{1}{2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial P_2}{\partial\phi}\right) = -\left(\frac{\partial P_0}{\partial Z_2} + \cos\phi\frac{\partial P_1}{\partial\hat{\tau}} - \frac{\sin\phi}{\hat{\tau}}\frac{\partial P_1}{\partial\phi}\right). \quad (24)$$

The condition for suppression of linear growth in P_2 is now

$$\int_0^{\pi}\left(\frac{\partial P_0}{\partial Z_2} + \cos\phi\frac{\partial P_1}{\partial\hat{\tau}} - \frac{\sin\phi}{\hat{\tau}}\frac{\partial P_1}{\partial\phi}\right)\sin\phi\,d\phi = 0. \quad (25)$$

When Eq. (23) for P_1 is substituted into the above condition, the evolution equation for P_0 is obtained, namely,

$$\frac{\partial P_0}{\partial Z_2} = \frac{1}{3}\left(\frac{\partial^2 P_0}{\partial\hat{\tau}^2} + \frac{2}{\hat{\tau}}\frac{\partial P_0}{\partial\hat{\tau}}\right). \quad (26)$$

In terms of the original variables Z and τ, the equation is

$$\frac{\partial P_0}{\partial Z} = \frac{1}{3}\left(\frac{\partial^2 P_0}{\partial\tau^2} + \frac{2}{\tau}\frac{\partial P_0}{\partial\tau}\right). \quad (27)$$

Equation (27) governs the evolution of the leading order probability-density function P_0 at large distances. When a delta-function initial condition is used for P_0, it is easy to check that the solution that satisfies Eq. (27) is

$$P_0(\tau, Z) = \left(\frac{3}{4\pi Z}\right)^{3/2} \exp\left(-\frac{3\tau^2}{4Z}\right), \quad Z \gg 1. \tag{28}$$

Recalling Eq. (23), the leading two-term approximation to the probability-density function P is then

$$P(\tau, \phi, Z) \approx P_0(\tau, Z) - \cos\phi \frac{\partial P_0(\tau, Z)}{\partial \tau}, \quad Z \gg 1, \tag{29}$$

where P_0 is given by Eq. (28). To obtain the DGD probability distribution $p(\tau, Z)$, we need only to average the full probability-distribution function P for the polarization dispersion vector over a sphere of radius τ to obtain the marginal distribution

$$p(\tau, Z) = 2\pi\tau^2 \int_0^\pi P \sin\phi \mathrm{d}\phi. \tag{30}$$

When approximation (29) is substituted into Eq. (30), the second term drops out because it has a zero average, and the first term shows that the DGD distribution is Maxwellian:

$$p(\tau, Z) \approx \frac{3\sqrt{3}\tau^2}{2\sqrt{\pi}Z^{3/2}} \exp\left(-\frac{3\tau^2}{4Z}\right) = p_M(\tau, Z), \quad Z \gg 1. \tag{31}$$

This result is consistent with previous research based on other birefringence models.[1,3,13] A new feature of our results, however, is that the probability-density function P for the polarization dispersion vector depends on the angle ϕ between it and the birefringence vector on the Poincaré sphere. Indeed, rewriting approximation (29) as

$$P(\tau, \phi, Z) \approx P_0(\tau - \cos\phi, Z), \quad Z \gg 1, \tag{32}$$

shows that the probability distribution is shifted toward larger τ values when $\phi = 0$ and toward smaller τ values when $\phi = \pi$. The amount of shift when ϕ varies from 0 to π is 2 DGD units in dimensionless variables. In dimensional quantities this amount is $4b'h_{\text{fiber}}$, which is 0.32 ps for the example parameters discussed earlier, i.e., for a PMD coefficient of 1 ps/$\sqrt{\text{km}}$ and $h_{\text{fiber}} = 5$ m). Intuitively, if we think of the fiber as consisting of sections of length h_{fiber} randomly oriented with respect to one another, the differential group delay accumulates at the rate $2b'$ in any constant birefringence section. Hence the last section increases the total DGD by an additional $2b'h_{\text{fiber}}$ if $\mathbf{\Omega}$ is aligned with the birefringence vector on the Poincaré sphere and decreases it by the same amount if $\mathbf{\Omega}$ points in the opposite direction.

The above asymptotic analysis has shown that at large distances, the DGD probability-density function becomes Maxwellian, and the probability-density function of the polarization dispersion vector depends upon the angle between it and the direction of the birefringence vector on the Poincaré sphere at the fiber output. One question that cannot be resolved by this asymptotic analysis, however, is how long it takes for the DGD to approach a Maxwellian distribution and for the probability P of the PMD vector to approach the approximation given by approximations (29) or (32). This question will be addressed in the next two sections by means of direct numerical simulations of the reduce and full Fokker–Planck equations. These direct numerical simulations will also serve to verify the above analytical results.

4. SIMULATION OF THE REDUCED FOKKER–PLANCK EQUATION

In this section, we simulate the reduced Fokker–Planck equation, Eq. (16), using a split-step method. The terms on the right-hand side of this equation are of two types: a diffusion term and a convection term. The solution to the equation containing just the diffusion term may be obtained by expanding the solution in terms of Legendre polynomials [see Eq. (20)]. The solution to the equation containing just the convection term is a translation that can be performed numerically with a two-dimensional spline interpolation. A second-order (Strang-splitting) scheme[17] is used to combine these two methods and thus integrate Eq. (16) in Z. For numerical efficiency, we have also used adaptive gridding along the τ direction. The Z step size changes accordingly to maintain high accuracy. This split-step method is unconditionally stable and efficient.

Ideally, the initial condition for Eq. (6) would be a δ function. Practically, we take the initial condition to be

$$P|_{Z=0} = \frac{D^3}{\pi^{3/2}} \exp(-D^2|\bar{\mathbf{\Omega}}|^2), \tag{33}$$

where D is large, so that $P|_{Z=0}$ simulates a δ function. Note that the coefficient in this initial condition is chosen so that the total probability (7) integrates to 1.

The results of simulations for $D = 50$ are shown in Fig. 1. For sufficiently large D values, the results are virtually identical, except at very short distances ($Z < 0.2$). Figures 1(a)–1(g) shows the DGD probability function as determined by Eq. (30) at various distances. In Figs. 1(f) and 1(g), a Maxwellian distribution, Eq. (31), is also shown (dashed curves) for comparison. We observe in Fig. 1 that, initially, the distribution is a very sharp Maxwellian function due to our choice of initial conditions, Eq. (33) [see Fig. 1(a)]. As Z increases, p becomes more or less symmetric in shape [see Fig. 1(b)]. As Z increases further, p becomes skewed toward large τ values [see Eqs. 1(c) and 1(d)], a behavior in contrast to that for a Maxwellian distribution. At $Z \approx 9$, p becomes almost symmetric again [see Fig. 1(e)]. For larger values of Z, p becomes skewed toward smaller τ values, finally tending toward the asymptotic Maxwellian distribution [see Figs. 1(f) and 1(g)]. At $Z \approx 30$, p is close to the Maxwellian distribution, except in the tail region where DGD is large, as described below and shown in Fig. 2.

To interpret these results in physical units, let us take a relatively large fiber correlation length, e.g., 100 m. Then the results in Fig. 1 indicate that, for this birefringence model, if the distance is hundreds of meters, then the DGD distribution is quite different from Maxwellian. Within a couple of kilometers, however, the distribution approaches a Maxwellian except in its tail region. We note that at large distances, the above numerical results confirm the analytical results obtained in Section 3,

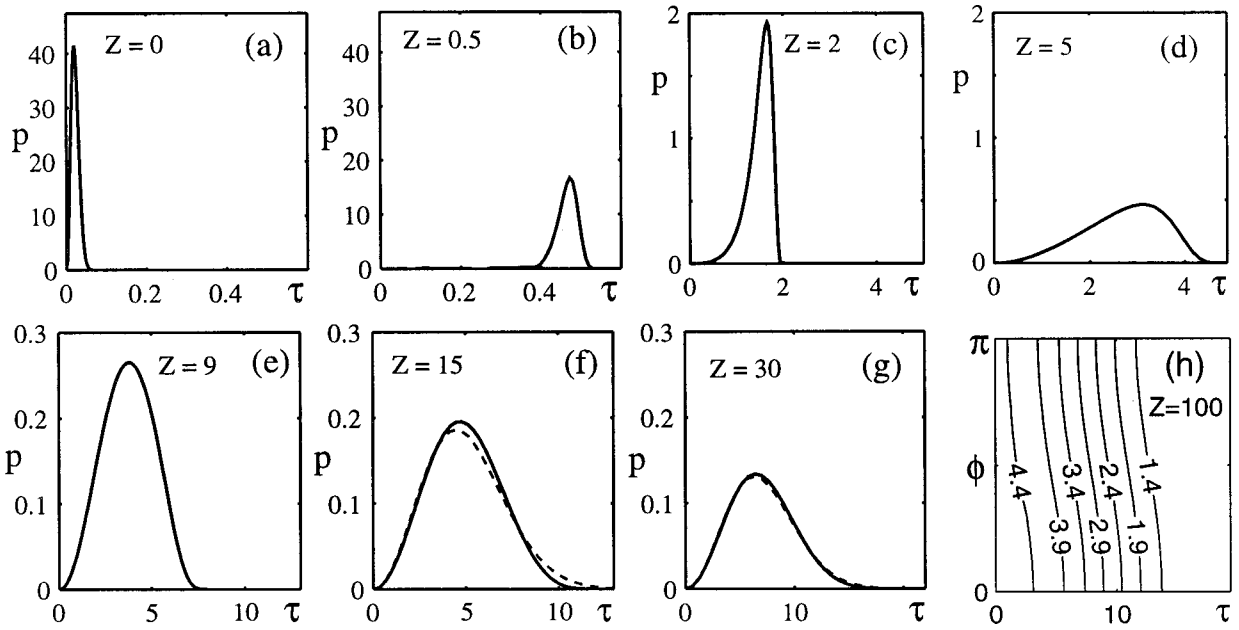


Fig. 1. (a)–(g) Probability-distribution function $p(\tau,Z)$ of the differential group delay at various distances when $\beta \gg 1$. The dotted curves in (f) and (g) are Maxwellian distributions. The distance is normalized to the fiber correlation length h_{fiber}, and the DGD is normalized to $2b'h_{\text{fiber}}$. (h) Contour plot showing the angular dependence of the probability distribution P for the dispersion vector $\vec{\Omega}$ at $Z = 100$. The contour levels have been magnified 40 000 times.

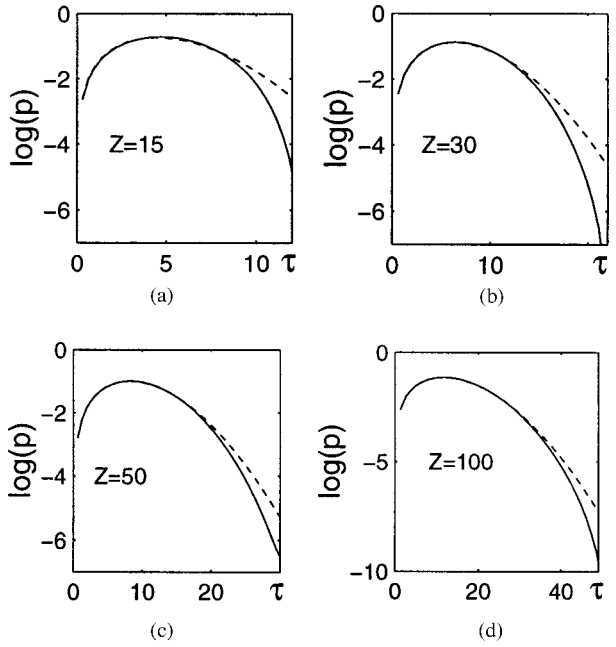


Fig. 2. Comparison of the true DGD probability distribution (solid curves) with a Maxwellian approximation (dashed curves) at various distances with a log scale to emphasize the behavior in the tails of the distributions. Differences at large DGD values are clearly seen, showing that hundreds of correlation lengths can be required for the distribution's tail to become Maxwellian.

which showed that the DGD probability distribution approaches a Maxwellian asymptotically.

The DGD distribution function in the tail region where the DGD is large is important for the assessment of system penalties due to PMD. The tail of the distribution function takes much longer to approach a Maxwellian than does the main part. To better illustrate the difference in the tails between our numerically obtained DGD distribution and the Maxwellian distribution, we have plotted both curves on a logarithmic scale at distances of $Z = 15$, 30, 50, and 100 in Fig. 2. Here the numerical distribution is shown with solid curves, and the Maxwellian distribution is denoted with dashed curves. Again, the distance is in terms of fiber correlation lengths, h_{fiber}. We see that at each distance, the difference between the true DGD distribution and Maxwellian distribution is most pronounced in the tails, and this difference can be many orders of magnitude larger. This difference is visible even at hundreds of fiber correlation lengths, corresponding to tens of kilometers [see Fig. 2(d)]. Thus the assessment of system penalties due to PMD based on the assumption that the DGD obeys a Maxwellian distribution may not be accurate. The use of a Maxwellian distribution overestimates the system penalty as the true probability of large DGD events is much smaller than the Maxwellian distribution predicts.

Consistent with our analytical results, our numerical simulations also reveal an angular (ϕ) dependence of the probability function P even after hundreds of fiber correlation lengths. To demonstrate this dependence, we show in Fig. 1(h) a contour plot of P in the (τ, ϕ) plane at $Z = 100$. We see that P is larger along the $\phi = 0$ direction (positive $\check{\Omega}_1$ axis) and smaller along the $\phi = \pi$ direction (negative $\check{\Omega}_1$ axis). The amount that the probability distribution shifts when ϕ varies from 0 to π in Fig. 1(h) is almost exactly 2 DGD units in dimensionless variables, in agreement with the analysis in Section 3. This angular dependence indicates that at the output of the fiber the DGD will be correlated with the angle between the polarization dispersion vector $\boldsymbol{\Omega}$ and the local birefringence vector on the Poincaré sphere. In particular, the expected DGD will be larger when $\boldsymbol{\Omega}$ is aligned with the bi-

refringence vector on the Poincaré sphere. A heuristic interpretation of this angular dependence has already been given in Section 3.

5. SIMULATION OF THE FULL FOKKER–PLANCK EQUATION

The reduced Fokker–Planck equation (16) is valid only when $\beta \gg 1$. Even though most installed fibers fall into this parameter regime, there are cases where $\beta \sim 1$. For instance, if we take the fiber correlation length to be 3 m and beat length to be 30 m, then $\beta = 1.26$. Thus to completely determine the DGD probability distribution for all fiber parameters, we are motivated to solve the full Fokker–Planck equation (6) for arbitrary values of β.

As before, we use a split-step method to solve Eq. (6) numerically. This equation can again be split into two parts: a convection and a rotation part corresponding to the second and third terms, and a diffusion part corresponding to the fourth term. The equation that contains just the diffusion term is solved by the pseudospectral method in spherical coordinates:

$$\bar{\Omega}_1 = \tau \sin\Theta \cos\Phi, \quad \bar{\Omega}_2 = \tau \sin\Theta \sin\Phi, \quad \bar{\Omega}_3 = \tau \cos\Theta. \tag{34}$$

Here, $\tau = |\bar{\mathbf{\Omega}}|$, Θ is the longitudinal angle, and Φ is the azimuthal angle. In these coordinates, the diffusion part of Eq. (6) is simply

$$\frac{\partial P}{\partial Z} = \frac{\partial^2 P}{\partial \Phi^2}. \tag{35}$$

The equation that contains just the convection and rotation terms is solved in Cartesian $\bar{\mathbf{\Omega}}$ coordinates by a translation and rotation of the axes with a semi-Lagrangian method.[18] The interpolation involved in this solution is handled by a reduced cubic interpolation.[19] We have also used the Strang-splitting method[17] so that our scheme is second-order accurate in Z. This split-step method is unconditionally stable and is insensitive to the size of the parameter value β. To further improve simulation efficiency, we have used adaptive gridding along the τ direction. The Z step size changes accordingly to maintain accuracy. We have tested our code by comparing the numerically determined and exact values of the total probability (which is one) and the average of τ^2 [see Eq. (8)], and have found that the code performs very well. We have also performed convergence tests to make sure that the parameters used in the numerical scheme were properly chosen; varying them to increase the accuracy did not significantly affect the results.

We have performed numerical simulations of the full Fokker–Planck equation (6) for various β values. In each case, we used $D = 50$ in the initial condition (33). For larger D values, the results are the same except at very short distances. We chose the grid numbers, $257 \times 111 \times 111$, on the following domain:

$$\Phi \times \Theta \times \tau = [0,2\pi) \times [0,\pi] \times [0,\tau_1],$$

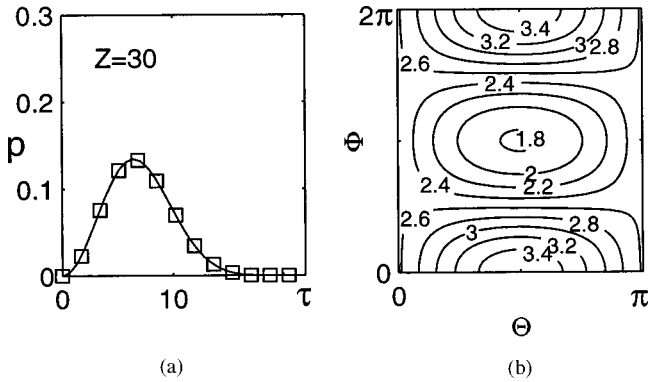
with τ_1 initially taken to be 0.1. Note that our code is adaptive along the τ direction. The τ interval $[0,\tau_1]$ is doubled when the solution P becomes nonnegligible on the spherical surface of radius τ_1. The Z step size was initially taken as $\Delta Z = 0.0005$ and changed to $1.5\,\Delta Z$ when τ_1 doubles. The quantities monitored in our simulations are the DGD probability-density function $p(\tau, Z)$, which is $P(\tau, \Phi, \Theta, Z)$ averaged over a spherical surface of radius τ,

$$p(\tau, Z) = \int_0^{2\pi} \int_0^{\pi} P \tau^2 \sin\Theta \, d\Theta \, d\Phi, \tag{36}$$

and the distribution of P on the spherical surface of radius $\tau_{\max}(Z)$, where $p(\tau, Z)$ attains its maximum value at a given distance Z.

We take first $\beta = 10$. In this case, the averaging of Section 2 applies, so we expect that the DGD distribution function $p(\tau, Z)$ remains the same as that shown in Fig. 1 except at very short distances where the averaging is not yet valid. This is indeed the case. Numerically, we found that for $Z > 0.5$, the DGD distribution of the full Fokker–Planck equation is almost indistinguishable from that in Fig. 1. To demonstrate, we plot $p(\tau, Z)$ determined from the full Fokker–Planck equation at $Z = 30$ (squares) in Fig. 3(a). In the same plot, $p(\tau, Z)$ from the reduced Fokker–Planck equation (16) is simultaneously shown with the solid curve for comparison [see Fig. 1(g)]. It can be seen that the two results agree with each other very well in both the tail and bulk regions. This agreement verifies that the averaging in Section 2 for $\beta \gg 1$ is justified. In view of Fig. 1, the DGD distribution for large β values indeed approaches a Maxwellian in ~30 correlation lengths, except in the tail where the approach to Maxwellian is much slower (see Fig. 2).

At the same distance, $Z = 30$, the angular dependence of the probability P on a sphere of radius $\tau_{\max} = 6.5$ in the (τ, Φ, Θ) space is shown in Fig. 3(b). At this value of $\tau_{\max}$, the DGD distribution $p(\tau)$ reaches its maximum [see Fig. 3(a)]. We note that, on this spherical surface, P is maximal at $\Theta = \pi/2$ and $\Phi = 0$ (or 2π), which is along the direction of the positive $\bar{\Omega}_1$ axis. The minimum of P occurs at $\Theta = \pi/2$ and $\Phi = \pi$, which is along the direc-



(a) (b)

Fig. 3. Simulation result for the full Fokker–Planck equation (6) with $\beta = 10$ and $Z = 30$. (a) DGD probability distribution $p(\tau)$ (squares). Also plotted in this figure is the same quantity from simulation of the reduced Fokker–Planck equation (16) (solid curves) for comparison [see Fig. 1(g)]. (b) Contour plot of $P(\tau,\Theta,\Phi)$ at $\tau = \tau_{\max} = 6.5$, where $p(\tau)$ reaches maximum [see (a)]. The contour levels have been magnified 10 000 times.

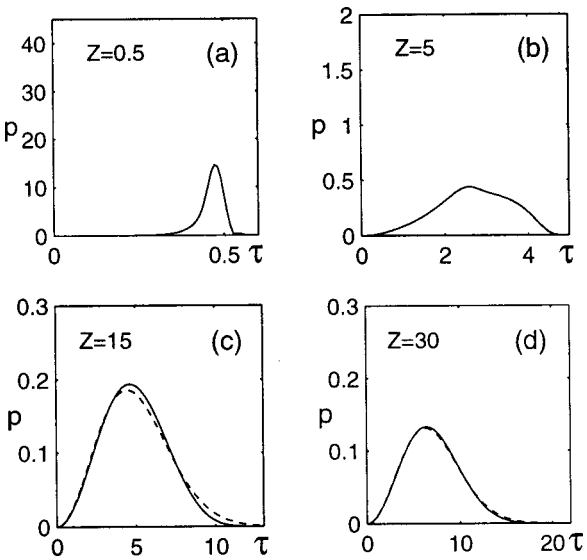


Fig. 4. DGD probability distribution $p(\tau,Z)$ at various distances for $\beta = 1$. In (c) and (d), the Maxwellian distribution (31) (dashed curves) is also shown for comparison.

tion of the negative $\bar{\Omega}_1$ axis. This result is consistent with the reduced Fokker–Planck analysis and the simulation in Sections 3 and 4.

Next we take $\beta = 1$, where the averaging of Section 2 is not necessarily expected to be valid. In this case, we are interested to know if the DGD distribution still approaches a Maxwellian at large distances, and, if so, is the approach faster or slower? To answer these questions, we plot the DGD distribution function $p(Z, \tau)$ at various distances ($Z = 0.5$, 5, 15, and 30) in Fig. 4. In Figs. 4(c) and 4(d), the Maxwellian distribution [approximation (31)] is also plotted for comparison (dashed curves). We find in this case that the DGD distribution still approaches a Maxwellian asymptotically, and that, even at this value of β, the approach to a Maxwellian distribution occurs roughly at the same rate as for large β values. The similar rates of approach are visible in the comparison of Fig. 1 with Fig. 4. Our numerical simulations at $\beta = 1$ also show that the probability distribution P of the polarization dispersion vector depends on the angles Φ and Θ in a way that is similar to that shown in Fig. 3(b), but the contours in this case are slightly skewed. This angular dependence again persists at all distances.

Our simulations for other β values between 1 and 10 produced similar results. Thus we conclude that, for $\beta \gtrsim 1$, the DGD distribution approaches a Maxwellian distribution in ~30 correlation lengths, except in the tail region, where the approach to Maxwellian occurs of the order of hundreds of correlation lengths.

For the sake of completeness, we have also explored the parameter regime $0 < \beta < 1$, or, equivalently, $h_{\text{fiber}} < L_B/4\pi$. Although fibers manufactured by traditional means do not typically fall within this regime, this parameter range may be relevant for low-PMD fiber manufactured by spinning the fiber during the drawing process.[20,21] Such spinning is likely to reduce the effective fiber correlation length. Since spin rates of the order of 20 turns per meter have been reported,[22,23] short correlation lengths can be expected.

Analytically, it can again be shown that when $\beta \ll 1$, the DGD probability distribution approaches a Maxwellian asymptotically at large distances. Our method is similar to that used in Section 3, and so the details will not be presented here. The main difference is that in this limit both h_{fiber} and L_B are important, and neither averages out. A careful analysis of a closely related problem, the Fokker–Planck equation for the evolution of the polarization state's probability distribution, shows that in this limit, behavior on both a fast and a slow length scale, namely, h_{fiber} and $L_B^2/4\pi^2 h_{\text{fiber}}$, is seen.[6,15,24] Note that when h_{fiber} is small, the second length scale can be much larger than the first. These two length scales are also expected to be visible in the evolution of the DGD's probability distribution. Numerically, our simulation results for Eq. (6) verify that for very small β values, the approach of the DGD distribution to a Maxwellian can be much slower when measured in terms of the fiber correlation length than for large β values. For instance, if $\beta = 0.1$, the DGD distribution approaches Maxwellian in the bulk region only after over 250 fiber correlation lengths. This observation is in contrast with Figs. 1 and 4 for moderate and large β values. However, this slow approach does not necessarily mean that the approach to a Maxwellian distribution is slower in physical distance units. The reason is that smaller β values typically correspond to shorter fiber correlation lengths. For instance, let us suppose that the beat length is 30 m; then $\beta = 0.1$ implies $h_{\text{fiber}} = 0.24$ m. As a result, the Maxwellian distribution is reached in the bulk region after 250 $h_{\text{fiber}} = 60$ m, which is still a short distance. This result also compares reasonably well with the theoretical estimate, $L_B^2/4\pi^2 h_{\text{fiber}} \approx 95$ m. For even smaller fiber correlation lengths, however, the distance required for the DGD distribution to become Maxwellian might be expected to be even longer.

6. SUMMARY AND DISCUSSION

In this paper we have studied the transient evolution of the probability distribution of the polarization dispersion vector using a realistic model of the fiber birefringence. First, we have shown analytically that when the fiber correlation length is of the same order or larger than the beat length, the DGD distribution becomes Maxwellian at large distances. Second, we have simulated numerically the Fokker–Planck equation for the probability function of the polarization dispersion vector and found that, for all practical fiber parameters, it becomes Maxwellian within a couple of kilometers, except in the distribution's tails where the DGD values are large. The tail of the DGD probability distribution also becomes Maxwellian, but much more slowly. Lastly, we have shown that the probability distribution for the polarization dispersion vector depends upon the angle between it and the local birefringence vector on the Poincaré sphere, showing that the amount of DGD remains correlated with the orientation of the local birefringence vector over arbitrarily long distances.

Generally speaking, the transient evolution of the type observed here is to be expected because linear fiber birefringence does not produce isotropic mixing on the Poin-

caré sphere.[6] Here we have demonstrated this effect using the first model of Wai and Menyuk,[6] where the magnitude of the birefringence is constant but its orientation varies randomly. In practice, the second model of Wai and Menyuk, where both the orientation and magnitude of the birefringence are randomly distributed, is a more appropriate description of the behavior of optical fibers.[12] Because the analysis of this second model is significantly more complicated, it has not been performed here. The same basic picture is still expected to hold for the second model, but the numerical factors are not likely to be precisely the same.

ACKNOWLEDGMENTS

J. Yang's and Y. Tan's research was supported in part by the U.S. Air Force Office of Scientific Research under contract USAF F49620-99-1-0174. W. L. Kath's research was supported by the U.S. Air Force Office of Scientific Research and the National Science Foundation. C. R. Menyuk's research was supported by the U.S. Department of Energy and the National Science Foundation.

REFERENCES

1. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," J. Lightwave Technol. **9**, 1439–1456 (1991).
2. J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization-mode dispersion in optical fibers," Proc. Natl. Acad. Sci. U.S.A. **97**, 4541–4550 (2000).
3. N. Gisin, "Solutions of the dynamic equation for polarization dispersion," Opt. Commun. **86**, 371–373 (1991).
4. F. Curti, B. Daino, G. D. Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," J. Lightwave Technol. **8**, 1162–1170 (1990).
5. N. Gisin, J.-P. Von der Weid, and J.-P. Pellaux, "Polarization mode dispersion of short and long single-mode fibers," J. Lightwave Technol. **9**, 821–827 (1991).
6. 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).
7. N. Gisin, R. Passy, P. Blasco, M. O. Van Deventer, R. Distl, H. Gilgen, B. Perny, R. Keys, E. Krause, C. C. Larsen, K. Morl, J. Pelayo, and J. Vobian, "Definition of polarization mode dispersion and first results of the COST 241 round-robin measurements," Pure Appl. Opt. **4**, 511–522 (1995).
8. I. P. Kaminow, "Polarization in optical fibers," IEEE J. Quantum Electron. **QE-17**, 15–22 (1981).
9. R. E. Schuh, X. Shan, and A. S. Siddiqui, "Polarization mode dispersion in spun fibers with different linear birefringence and spinning parameters," J. Lightwave Technol. **16**, 1583–1588 (1998).
10. J. Botineau and R. H. Stolen, "Effect of polarization on spectral broadening in optical fibers," J. Opt. Soc. Am. **72**, 1592–1596 (1982).
11. M. N. Islam, "Ultrafast all-optical logic gates based on soliton trapping in fibers," Opt. Lett. **14**, 1257–1259 (1989).
12. A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, "Statistical characterization of fiber random birefringence," Opt. Lett. **25**, 1322–1324 (2000).
13. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," Opt. Lett. **16**, 372–374 (1991).
14. L. Arnold, *Stochastic Differential Equations, Theory and Applications* (Wiley, New York, 1974).
15. J. W. Zhang, P. K. A. Wai, W. L. Kath, and C. R. Menyuk, "Nonlinear polarization-mode dispersion in optical fibers with randomly varying birefringence," J. Opt. Soc. Am. B **16**, 2967–2979 (1997).
16. A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).
17. G. Strang, "On the construction and comparison of difference schemes," SIAM J. Numer. Anal. **5**, 506–517 (1968).
18. D. R. Durran, *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics* (Springer-Verlag, New York, 1999).
19. H. Ritchie, C. Temperton, A. Simmons, M. Hortal, T. Davies, D. Dent, and M. Hamrud, "Implementation of the semi-Lagrangian method in a high-resolution version of the ECMWF model," Mon. Weather. Rev. **123**, 489–514 (1995).
20. A. J. Barlow, D. N. Payne, M. R. Hadley, and R. J. Mansfield, "Production of single-mode fibers with negligible intrinsic birefringence and polarization mode dispersion," Electron. Lett. **17**, 725–726 (1981).
21. A. J. Barlow and J. J. Ramskov-Ha, "Birefringence and polarization mode-dispersion in spun single-mode fibers," Appl. Opt. **20**, 2962–2968 (1981).
22. M. J. Li and D. A. Nolan, "Fiber spin-profile designs for producing fibers with low polarization mode dispersion," Opt. Lett. **23**, 1659–1662 (1998).
23. R. E. Schuh, X. Shan, and A. S. Siddiqui, "Polarization mode dispersion in spun fibers with different linear birefringence and spinning parameters," J. Lightwave Technol. **16**, 1583–1588 (1998).
24. T. Ueda and W. L. Kath, "Dynamics of optical pulses in randomly birefringent fibers," Physica D **55**, 166–181 (1992).