Polarization evolution and dispersion in fibers with spatially varying birefringence

C. R. Menyuk and P. K. A. Wai

Department of Electrical Engineering, University of Maryland, Baltimore, Maryland 21228-5398

Received August 13, 1993; revised manuscript received December 15, 1993

A procedure is described that allows one to solve the evolution equations in birefringent optical fibers by using repeated diagonalization. With this approach several practical problems are solved in a unified way. Included are evolution in twisted fibers, sinusoidally rocked fibers, and fibers with randomly varying birefringence. In the last case it is shown that a phenomenological model described by Poole and others applies to fibers whose axes of birefringence can take on any orientation.

1. INTRODUCTION

It is well known that so-called single-mode fibers are actually bimodal because of birefringence.[1-3] Values of $\Delta n/n$ vary in the range $10^{-9}-10^{-3}$. Fibers with high or moderate-to-high values of birefringence play a key role in sensors,[4-7] gyros,[8,9] switches, [10-13] rotators,[14,15] and other devices.[14-16] In these applications the location of the axes of birefringence is known, but random fluctuations along the length of the fiber can degrade the polarization-holding capabilities of the fiber. Kaminow[3] derived a model that describes in an ensemble-averaged sense the decrease in the degree of polarization. Shortly thereafter Rashleigh *et al.*[17] showed that this model applied to single lengths of optical fiber when a broadband source is used. With a sufficiently broadband source excellent agreement is found between theory and experiment.[17,18]

In communication systems, by contrast, the birefringence is typically in the range $\Delta n/n \sim 10^{-7}-10^{-5}$, and the orientation of the axes of birefringence is unknown. While little is known about the details of the birefringence evolution along the fiber, it is generally supposed that the birefringence is locally linear and that the strength and the orientation of the birefringence vary randomly on a length scale whose autocorrelation spectrum may have components from a few centimeters to perhaps hundreds of meters.[19-24] This random variation leads to polarization mode dispersion, which in turn has important implications for long-distance nonreturn-to-zero[25] and soliton[26-28] communication systems. It is not at all apparent that a coupled-mode approach[29] like that of Kaminow[3] will be useful in this context. Yet Poole and co-workers,[30-35] who introduced this model phenomenologically to study polarization mode dispersion in communication fibers, and others who have used this model[36-39] have demonstrated reasonable consistency with the experimental data.

In this paper we introduce a theoretical framework that allows us to derive Poole's phenomenological model from physically reasonable assumptions about how the random variations in optical fibers occur. In particular, we demonstrate that this model remains valid when the birefringence axes can take on any orientation, and we find the theoretical limitations of this model. Additionally, this theoretical framework also allows us to determine the field evolution in a twisted or spun fiber and the evolution in a fiber with rocked axes. While the results that we obtain for twisted and rocked fiber are already known,[5,15,21,22] our approach allows us to obtain these results simply and to deal with the evolution in both these fibers and communication fibers with randomly varying birefringence in a unified way.

The remainder of this paper is organized as follows. In Section 2 we introduce the theoretical approach that we use. In Section 3 we describe its application to twisted and rocked fibers. In Section 4, we describe its application to communication fibers with randomly varying birefringence and obtain Poole's phenomenological model from physically reasonable assumptions. Finally, Section 5 contains the conclusions.

2. THEORETICAL APPROACH

In a perfectly circular fiber the fundamental mode of the fiber is the HE_{11} mode. This mode exists for arbitrarily small index differences between the core and the cladding and, in the limit as the index difference grows smaller, changes to a plane-wave solution propagating along the axis of the fiber. From this physical standpoint it follows that this mode should be doubly degenerate, as indeed is the case. Any ellipticity in the core or stresses at the core–cladding interface will break this degeneracy, leading to two distinct eigenmodes in which the electric and the magnetic fields inside the fiber have a unique configuration. While the same effect that breaks the degeneracy of the HE_{11} mode will lead to some alteration in the transverse profile, this breakdown of the degeneracy is so small—less than one part in a thousand in fibers with the largest possible birefringence—that its effect on the transverse profile can be ignored. Hence the transverse patterns of the two eigenmodes are essentially the same and are simply rotated by 90° with respect to each other.

From the preceding discussion it follows that at a given frequency the complex electric field $\mathcal{E}$ propagating along

0740-3224/94/071288-09$06.00

the fiber axis can be written as

$$\mathcal{E}(z,\mathbf{r}_t) = E_1(z)\mathbf{\Psi}_1(\mathbf{r}_t) + E_2(z)\mathbf{\Psi}_2(\mathbf{r}_t), \tag{1}$$

where z indicates the coordinate along the fiber axis and $\mathbf{r}_t$ indicates the coordinates transverse to the fiber axis. The vectors $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$ differ negligibly in structure except for a 90° rotation. This structure is for all practical purposes the same as the standard HE_{11} structure in circular fibers. The complex scalars E_1 and E_2 contain the z-dependent phase evolution of the electric field. Phase differences between E_1 and E_2 accumulate over lengths that are many orders of magnitude greater than the light's wavelength—indeed, many kilometers in the largest stretches of fiber. Thus even small differences in the rate of change of the phase in the two components E_1 and E_2 will lead to measurable effects. It is this fact that lies at the heart of the observed fiber birefringence. In the absence of nonlinear effects in the optical fiber, the two components will couple linearly; thus we may write

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = i\begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}\begin{pmatrix} E_1 \\ E_2 \end{pmatrix}, \tag{2}$$

which we summarize as

$$\frac{\mathrm{d}\mathbf{E}}{\mathrm{d}z} = i\mathbf{K}\mathbf{E}. \tag{3}$$

The k_{ij} are functions of z whose behavior depends on the details of the variations in the fiber core shape and the stresses that act on the fiber core. Assuming, however, that fiber loss is negligible, the matrix $\mathbf{K}$ must be Hermitian; i.e., k_{11} and k_{22} must be real and $k_{12} = k_{21}{}^*$. In this paper we also assume that the local fiber eigenmodes are linearly polarized. From a physical standpoint this assumption is equivalent to assuming that the electric and magnetic fields have definite orientations on the length scale of a wavelength, so that there is no intrinsic helicity in the material—a physically reasonable assumption in glass. This assumption is consistent with almost all experimental observations to date, although specially manufactured fibers with intrinsic circular birefringence have been reported.[21,22] From a mathematical standpoint this assumption implies that both k_{12} and k_{21} are real, and hence $k_{12} = k_{21}$. We can rephrase these results as follows: defining the standard Pauli matrices,

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\sigma}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{4}$$

we find

$$\mathbf{K} = k_0\mathbf{I} + \kappa_1\boldsymbol{\sigma}_1 + \kappa_2\boldsymbol{\sigma}_2 + \kappa_3\boldsymbol{\sigma}_3, \tag{5}$$

where $k_0 = (k_{11} + k_{22})/2$, $\kappa_1 = (k_{12} + k_{21})/2$, $\kappa_2 = i(k_{12} - k_{21})/2$, and $\kappa_3 = (k_{11} - k_{22})/2$. Demanding that the fiber be lossless is equivalent to demanding that k_0 and all the κ_j be real. Demanding that the local polarization eigenmodes be linear is equivalent to demanding that $\kappa_2 = 0$. We now reduce Eq. (3) by eliminating the phase variation that is not due to the birefringence. We set

$$\mathbf{A} = \mathbf{E}\exp\left[-i\int_0^z k_0(z')\mathrm{d}z'\right] \tag{6}$$

and obtain the equation

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}z} = i\mathbf{\Theta}\mathbf{A}, \tag{7}$$

where $\mathbf{\Theta} = \kappa_3\boldsymbol{\sigma}_3 + \kappa_1\boldsymbol{\sigma}_1$. Defining

$$\kappa_3 = b\sin\theta, \qquad \kappa_1 = b\sin\theta, \tag{8}$$

we find explicitly

$$\mathbf{\Theta} = b\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}, \tag{9}$$

where b can be interpreted physically as the birefringence strength and θ is the angle that the birefringence axes make with respect to a fixed pair of axes.

In polarization-preserving fibers, in which there are a large birefringence and well-defined axes of birefringence, we may assume that θ is small and randomly varying. In this case, with $\gamma = b\theta$, Eq. (9) becomes

$$\mathbf{\Theta} = \begin{bmatrix} b & \gamma \\ \gamma & -b \end{bmatrix}, \tag{10}$$

where $|\gamma| \ll b$ and varies randomly. This random variation leads to a deterioration in the degree of polarization of injected light. This model applies to many practically important situations, including sensors, gyros, and some switches,[4–10] and its solution has been discussed by Kaminow.[3] One might question what is meant by random in this context, since, for this term to have meaning in a physical application, there must be some ensemble over which one is averaging.[40] It is natural to suppose that the variation of θ as a function of z will yield the necessary ensemble, but experiments indicate that the lengths of fiber being used in at least two of the applications just cited, sensors and gyros, are too short. Rashleigh *et al.*[17] resolved this issue elegantly and completely by showing that an ensemble average over the different frequencies in a broadband source yields excellent agreement between theory and experiment.[17,18] It is possible, indeed advantageous, to use broadband sources in several of these applications. Thus the applicability of Kaminow's model in the context of polarization-preserving fibers is well understood.

The situation with communication fibers differs in important respects. First, the axes of birefringence are unknown and presumably pass through all possible orientations over a sufficiently long length. Second, one does not typically use broadband sources, and one does typically propagate light over long lengths; hence the appropriate ensemble average should be over distance. Despite these differences, Poole and co-workers[32,34] assumed that Kaminow's model still applies and deduced the polarization mode dispersion from this model. While their experimental tests of this model used ensemble averages over frequency and temperature,[33,34] other studies[25–28,37,38] indicate that these measurements remain valid when an appropriate ensemble average is taken over length.

Motivated by these observations, we attempt to solve Eq. (9) by repeated diagonalization. This approach yields a physical basis for understanding when the phenomenological model of Poole and co-workers is expected to apply. Moreover, as is noted in Section 1, this approach is also of use in finding the evolution of light in twisted or rocked fiber with moderate-to-high birefringence.

We begin by carrying out the transformation that diagonalizes $\boldsymbol{\Theta}$ when $\theta_z \equiv \mathrm{d}\theta/\mathrm{d}z = 0$. Since $\boldsymbol{\Theta}$ is Hermitian, this transformation will be unitary, and we find explicitly that

$$\mathbf{U}_1\boldsymbol{\Theta}\mathbf{U}_1^{-1} = b\boldsymbol{\sigma}_3, \tag{11}$$

where

$$\begin{aligned}\mathbf{U}_1 &= \cos(\theta/2)\mathbf{I} + i\sin(\theta/2)\boldsymbol{\sigma}_2,\\ \mathbf{U}_1^{-1} &= \cos(\theta/2)\mathbf{I} - i\sin(\theta/2)\boldsymbol{\sigma}_2.\end{aligned} \tag{12}$$

Making the transformation $\mathbf{B}_1 = \mathbf{U}_1\mathbf{A}$, we conclude that

$$\begin{aligned}\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} &= \left(i\mathbf{U}_1\boldsymbol{\Theta}\mathbf{U}_1^{-1} - \mathbf{U}_1\frac{\mathrm{d}\mathbf{U}_1^{-1}}{\mathrm{d}z}\right)\mathbf{B}_1\\ &= i\left(b\boldsymbol{\sigma}_3 + \frac{\theta_z}{2}\boldsymbol{\sigma}_2\right) \equiv i\boldsymbol{\Psi}_1\mathbf{B}_1.\end{aligned} \tag{13}$$

Physically we are transforming from a fixed frame of reference to a frame that follows the local axes of birefringence. If we assume that $|\theta_z/2| \ll b$ and that $|(b\theta_{zz} - b_z\theta_z)/2| \gg b^3$, where $b_z \equiv \mathrm{d}b/\mathrm{d}z$ and $\theta_{zz} \equiv \mathrm{d}^2\theta/\mathrm{d}z^2$, then Eq. (13) is essentially the same as the phenomenological model of Poole and co-workers, as is discussed in greater detail in Section 4. The appearance of $\boldsymbol{\sigma}_2$ in Eq. (13) as opposed to $\boldsymbol{\sigma}_1$ in Eq. (7) makes no difference in the behavior of the polarization mode dispersion. The first condition $|\theta_z/2| \ll b$ implies that the off-diagonal coupling is weak, while the second condition $|(b\theta_{zz} - b_z\theta_z)/2| \gg b^3$ implies that θ_z changes on a length scale that is short compared with the birefringent beat length, so that ensemble-averaging over length is possible.

The origin of the second constraint can best be understood by noting that the transformation from Eq. (7) to Eq. (13) is the first in an infinite series of transformations that can be used to completely diagonalize Eq. (7) and hence solve it when the variation of θ is sufficiently smooth. We obtain the second member of this series of transformations by setting

$$b = c_2\cos\phi_2, \qquad \theta_z/2 = c_2\sin\phi_2, \tag{14}$$

defining $\mathbf{U}_2 = \cos(\phi_2/2)\mathbf{I} - i\sin(\phi_2/2)\boldsymbol{\sigma}_1$, and letting $\mathbf{B}_2 = \mathbf{U}_2\mathbf{B}_1$. We then find

$$\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} = i\boldsymbol{\Psi}_2\mathbf{B}_1, \tag{15}$$

where $\boldsymbol{\Psi}_2 = c_2\boldsymbol{\sigma}_3 - (\phi_{2,z}/2)\boldsymbol{\sigma}_1$ and

$$\phi_{2,z} = \frac{b^2}{b^2 + \theta_z^2/4}\frac{\mathrm{d}}{\mathrm{d}z}\frac{\theta_z}{2b}, \tag{16}$$

so that the second constraint amounts to demanding that $\phi_{2,z}$ be large. The physical meaning of the transformation from Eq. (13) to Eq. (15) is not nearly so transparent as the physical meaning of the transformation from Eq. (7) to Eq. (13). Nonetheless, these two transformations are closely analogous, and we can repeat this procedure an arbitrarily large number of times, obtaining an infinite sequence of transformations. At successive transformations successively higher derivatives of θ and b appear. When the product of these transformations converges, i.e., when the infinite product

$$\mathbf{V} = \cdots\mathbf{U}_3\mathbf{U}_2\mathbf{U}_1 \tag{17}$$

exists, we find $\mathbf{A} = \mathbf{V}^{-1}\mathbf{B}_\infty$, where

$$\frac{\mathrm{d}\mathbf{B}_\infty}{\mathrm{d}z} = i\boldsymbol{\Psi}_\infty\mathbf{B}_\infty \tag{18}$$

and $\boldsymbol{\Psi}_\infty = \mathbf{V}\boldsymbol{\Theta}\mathbf{V}^{-1} \equiv c_\infty\boldsymbol{\sigma}_3$ is diagonal. Writing the two components of $\mathbf{B}_\infty$ as $B_{\infty,1}$ and $B_{\infty,2}$, we obtain

$$\begin{aligned}B_{\infty,1}(z) &= B_{\infty,1}(0)\exp\left[i\int_0^z c_\infty(z')\mathrm{d}z'\right],\\ B_{\infty,2}(z) &= B_{\infty,2}(0)\exp\left[-i\int_0^z c_\infty(z')\mathrm{d}z'\right].\end{aligned} \tag{19}$$

The quantity $c_\infty > b$ is greater than zero. Hence, in an optical fiber whose statistical properties are stationary, it follows that the phase difference between the two components of $\mathbf{B}_\infty$ will grow linearly with z on average, as long as θ and b are sufficiently smooth functions of z to ensure that the infinite product (17) converges. The polarization mode dispersion, which is given by the frequency derivative of the phase difference, will also grow linearly with z. This result differs from both the result of the phenomenological model of Poole and co-workers as well as the applicable experiments, which yield a polarization mode dispersion that is ultimately proportional to $z^{1/2}$.

It is not surprising that when the product given by Eq. (17) is convergent the dispersion is ultimately proportional to z, since well-defined modes exist that are uncoupled, although their relationship to the local eigenmodes can be quite complicated and, in particular, they can be in any state of polarization as a function of z, not just the linearly polarized states. We thus conclude that the assumption that there is random mode coupling in the fiber depends on the breakdown of convergence of Eq. (17), which corresponds physically to the presence of abrupt changes in some of the derivatives of b and θ. We will postpone further discussion of the physical conditions under which we can treat mode coupling as random to Section 4. We merely point out here that the assumption that θ_z is small in magnitude while θ_{zz} is large is not unreasonable. It corresponds to many small, abrupt twitches in the drawing mechanism as the fiber is being drawn. This sort of behavior is quite common in nature—Brownian motion being a classical example—and it is connected to the fact that the acceleration of physical objects can change abruptly, but their velocities cannot.

The product given by Eq. (17) will evidently be convergent when it terminates, as is the case when b equals a constant and θ equals some polynomial in z. When $\theta = \epsilon \sin \Omega z$, it is not apparent that the sequence will be convergent, but it will be useful if $\Omega \ll b$. Both these cases are of great practical interest, and the transformation approach that we have just described gives a simple, powerful, unified way to determine the evolution. This issue is discussed in Section 3.

3. APPLICATION TO NONRANDOM VARIATIONS

Twisted fibers with strong birefringence play an important role in a wide variety of fiber devices, including sensors, rotators, and filters,[5,15,21,22] and thus have long been an object of study. The evolution of the light's state of polarization is usually studied on the Poincaré sphere. A complex vector $\mathbf{A}$ is related to a point (X, Y, Z) on the Poincaré sphere by the relations[41–43]

$$X = \mathbf{A}^\dagger \boldsymbol{\sigma}_3 \mathbf{A}, \qquad Y = \mathbf{A}^\dagger \boldsymbol{\sigma}_1 \mathbf{A}, \qquad Z = \mathbf{A}^\dagger \boldsymbol{\sigma}_1 \mathbf{A} \tag{20}$$

This relation can be regarded as a case of the well-known SU(2) $\leftrightarrow$ O(3) homomorphism.[44] An equation of motion corresponding to Eq. (7) can be derived for motion on the Poincaré sphere.[32,33] We have found it useful instead to solve Eq. (7) directly, yielding a solution in the form

$$\mathbf{A}(z) = \mathsf{T}(z)\mathbf{A}_0 , \tag{21}$$

where $\mathbf{A}_0 = \mathbf{A}(z = 0)$ and $\mathsf{T}(z)$ must be unitary. Defining now $R_j = \mathbf{A}^\dagger \boldsymbol{\sigma}_j \mathbf{A}$, one can show that

$$\sum_{j=1}^{3} R_j \boldsymbol{\sigma}_j = \mathsf{T} \sum_{j=1}^{3} R_{j,0} \boldsymbol{\sigma}_j \mathsf{T}^\dagger , \tag{22}$$

where $R_{j,0} = R_j(z = 0)$, and we note that

$$\mathbf{A}\mathbf{A}^\dagger = \frac{1}{2}\mathsf{I} + \frac{1}{2}\sum_{j=1}^{3} \boldsymbol{\sigma}_j \mathbf{A}^\dagger \boldsymbol{\sigma}_j \mathbf{A} . \tag{23}$$

Equation (22) allows us to trace the motion on the Poincaré sphere.

We now consider three simple examples.

A. Constant Angle ($\theta = \theta_0$)

In this case $\boldsymbol{\Theta} = b(\cos \theta_0)\boldsymbol{\sigma}_3 + b(\sin \theta_0)\boldsymbol{\sigma}_1$ in Eq. (7), and the matrix U_1 that diagonalizes $\boldsymbol{\Theta}$ is given by $\mathsf{U}_1 = \cos(\theta_0/2)\mathsf{I} + i \sin(\theta_0/2)\boldsymbol{\sigma}_2$. Letting $\mathbf{B}_1 = \mathsf{U}_1 \mathbf{A}$, we obtain the equation

$$\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} = i\boldsymbol{\Psi}_1 \mathbf{B}_1 , \tag{24}$$

where $\boldsymbol{\Psi}_1 = b\boldsymbol{\sigma}_3$. Since Eq. (24) is diagonal, it is easily integrated to yield

$$\mathbf{B}_1(z) = [\cos(bz)\mathsf{I} + i \sin(bz)\boldsymbol{\sigma}_3]\mathbf{B}_1(0) . \tag{25}$$

Using the relation $\mathbf{A} = \mathsf{U}_1^{-1}\mathbf{B}_1$ to return to the original coordinate system, we find

$$\begin{aligned} \mathsf{T} &= \mathsf{U}_1^{-1}[\cos(bz)\mathsf{I} + i \sin(bz)\boldsymbol{\sigma}_3]\mathsf{U}_1 \\ &= \cos(bz)\mathsf{I} + i \cos \theta_0 \sin(bz)\boldsymbol{\sigma}_3 \\ &\quad + i \sin \theta_0 \sin(bz)\boldsymbol{\sigma}_1 . \end{aligned} \tag{26}$$

Using Eq. (22), we finally obtain

$$\begin{aligned} X &= [\cos^2(bz) + \cos(2\theta_0)\sin^2(bz)]X_0 \\ &\quad + \sin(2\theta_0)\sin^2(bz)Y_0 - \sin \theta_0 \sin(2bz)Z_0 , \\ Y &= \sin(2\theta_0)\sin^2(bz)X_0 \\ &\quad + [\cos^2(bz) - \cos(2\theta_0)\sin^2(bz)]Y_0 \\ &\quad + \cos \theta_0 \sin(2bz)Z_0 , \\ Z &= \sin \theta_0 \sin(2bz)X_0 \\ &\quad - \cos \theta_0 \sin(2bz)Y_0 + \cos(2bz)Z_0 . \end{aligned} \tag{27}$$

Equations (26) and (27) can be easily generalized to deal with the case in which b is z dependent; one simply replaces bz by $\int_0^z b(z')\mathrm{d}z'$.

If we initially set $X_0 = \cos \theta_0$, $Y_0 = \sin \theta_0$, and $Z_0 = 0$, or we initially set $X_0 = -\cos \theta_0$, $Y_0 = -\sin \theta_0$, and $Z_0 = 0$, corresponding to an input wave whose input polarization state is one of the fiber eigenmodes, then there is no motion on the Poincaré sphere. Otherwise the trajectory of Eqs. (27) is a circle about the axis of the eigenmodes. This circular motion is most readily apparent when $\theta_0 = 0$, in which case Eqs. (27) become

$$\begin{aligned} X &= X_0 , \\ Y &= \cos(2bz)Y_0 + \sin(2bz)Z_0 , \\ Z &= -\sin(2bz)Y_0 + \cos(2bz)Z_0 . \end{aligned} \tag{28}$$

The periodic evolution of the state of polarization predicted by Eqs. (28) is a well-known feature of evolution in any birefringent material, not just optical fibers,[41] and plays an important role in a variety of optical devices. The period in Eqs. (28) is half the period in which the components of $\mathbf{A}$ evolve. The reason is that after a distance $z = \pi b$, $\mathbf{A} \to -\mathbf{A}$, and both $\mathbf{A}$ and $-\mathbf{A}$ correspond to the same point on the Poincaré sphere.

B. Constant Rate of Angular Change ($\theta = \Omega z$)

In the case of a constant angular rotation, one must carry out two transformations as described in Section 2. These transformations yield, in sequence,

$$\begin{aligned} \frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} &= i\left(b\boldsymbol{\sigma}_3 + \frac{\Omega}{2}\boldsymbol{\sigma}_2\right)\mathbf{B}_1 , \\ \frac{\mathrm{d}\mathbf{B}_2}{\mathrm{d}z} &= ic_2\boldsymbol{\sigma}_3\mathbf{B}_2 , \end{aligned} \tag{29}$$

where $c_2 = (b^2 + \Omega^2/4)^{1/2}$. The last equation can be integrated immediately to yield

$$\mathbf{B}_2(z) = \mathsf{T}_2\mathbf{B}_{2,0} ,$$

where $\mathsf{T}_2 = \cos(c_2 z)\mathsf{I} + i \sin(c_2 z)\boldsymbol{\sigma}_3$, and $\mathbf{B}_{2,0} = \mathbf{B}_2(z = 0)$. We then find that $\mathbf{B}_1 = \mathsf{T}_1\mathbf{B}_{1,0}$, where $\mathsf{T}_1 = \mathsf{U}_2^{-1}\mathsf{T}_2\mathsf{U}_2$ and $\mathsf{U}_2 = \cos(\phi_2/2)\mathsf{I} - i \sin(\phi_2/2)\boldsymbol{\sigma}_1$. We recall that $c_2 \sin \phi_2 = \Omega/2$ and $c_2 \cos \phi_2 = b$. We find explicitly that

$$\mathsf{T}_1 = \cos(c_2 z)\mathsf{I} + i \cos \phi_2 \sin(c_2 z)\boldsymbol{\sigma}_3 + i \sin \phi_2 \sin(c_2 z)\boldsymbol{\sigma}_2 . \tag{30}$$

Defining, by analogy to Eqs. (20), $X_1 = \mathbf{B}_1{}^\dagger \boldsymbol{\sigma}_3 \mathbf{B}_1$, $Y_1 = \mathbf{B}_1{}^\dagger \boldsymbol{\sigma}_1 \mathbf{B}_1$, and $Z_1 = \mathbf{B}_1{}^\dagger \boldsymbol{\sigma}_2 \mathbf{B}_1$, we obtain

$$
\begin{aligned}
X_1 &= [\cos^2(c_2 z) + \cos(2\phi_2)\sin^2(c_2 z)]X_0 \\
&\quad + \sin\phi_2 \sin(2c_2 z)Y_0 + \sin(2\phi_2)\sin^2(c_2 z)Z_0\,, \\
Y_1 &= -\sin\phi_2 \sin(2c_2 z)X_0 \\
&\quad + \cos(2c_2 z)Y_0 + \cos\phi_2 \sin(2c_2 z)Z_0\,, \\
Z_1 &= \sin(2\phi_2)\sin^2(c_2 z)X_0 \\
&\quad - \cos\phi_2 \sin(2c_2 z)Y_0 \\
&\quad + [\cos^2(c_2 z) - \cos(2\phi_2)\sin^2(c_2 z)]Z_0\,. \qquad (31)
\end{aligned}
$$

This result closely resembles Eqs. (27), but quantities on the Z axis are interchanged with quantities on the Y axis. In particular, if we let

$$
X_0 = \cos\phi_2 = \frac{b}{(b^2 + \Omega^2/4)^{1/2}}\,, \qquad Y_0 = 0\,,
$$
$$
Z_0 = \sin\phi_2 = \frac{\Omega/2}{(b^2 + \Omega^2/4)^{1/2}}\,, \qquad (32)
$$

we find that X_1, Y_1 and Z_1 do not evolve. More generally, the trajectory on the sphere (X_1, Y_1, Z_1) will be circular about the axis given in Eqs. (32). When we transform back to the original reference frame and calculate X, Y, Z, we find that the motion is a combination of the circular motion given by Eqs. (31) and circular motion about the Z axis, leading to nutatory motion that is like the motion of a spinning top.[45] We do not present the full equation of evolution, as it is lengthy without being physically revealing.

An important result that emerges from Eqs. (31) and (32) is that in the rotating frame given by $x = \cos(\Omega z)$, $y = \sin(\Omega z)$, where x and y refer to the transverse coordinates of the optical fiber, the fiber appears to be elliptically birefringent. That is the physical meaning of the nonzero Z_0 component in Eqs. (32). This occurs even though the underlying local modes of the fiber are linearly birefringent! This behavior contrasts with the situation in which the underlying local modes are elliptical, as would occur, for example, in a fiber whose core is elliptically shaped and fixed in direction but is helically strained. This sort of fiber has been discussed by Rashleigh[22] but is difficult to make.[5] It is easier to make the sort of fiber described here, in which the local eigenmodes are linearly polarized but in which twisting leads to elliptical eigenmodes in a rotated frame.[22]

Elliptical birefringence has implications for nonlinear behavior in optical fibers. In particular, Menyuk[46] has shown that the ratio of the cross-phase- to the self-phase-modulation coefficients is given by $B = (2 + 2\sin^2\alpha)/(2 + \cos^2\alpha)$, where α is the angle of ellipticity. When $B = 1$, i.e., when $\tan\alpha = 1/\sqrt{2}$, nonlinear polarization rotation is eliminated. Since this effect is a nuisance in fiber sensors, its elimination would be helpful. It is natural to wonder if this result still holds when a fiber is twisted so that the underlying local birefringence is still linear but there is an elliptical birefringence in a rotating frame. The answer turns out to be positive, as we have recently demonstrated.[47]

C. Sinusoidal Rocking ($\theta = \epsilon \sin \Omega z$)

Periodically twisted fibers play an important role in a number of fiber devices that do switching and filtering.[5,14,15,21,22]

If we directly apply the formalism of Section 2, we find the following hierarchy of equations:

$$
\begin{aligned}
\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}z} &= i[b\cos(\epsilon\sin\Omega z)\boldsymbol{\sigma}_3 + b\sin(\epsilon\sin\Omega z)\boldsymbol{\sigma}_1]\mathbf{A}, \\
\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} &= i[b\boldsymbol{\sigma}_3 + \frac{1}{2}\epsilon\Omega\cos(\Omega z)\boldsymbol{\sigma}_2]\mathbf{B}_1, \\
\frac{\mathrm{d}\mathbf{B}_2}{\mathrm{d}z} &= i\left[c_2\boldsymbol{\sigma}_3 + \frac{\epsilon\Omega^2}{4b}\frac{\sin\Omega z}{1 + (\epsilon^2\Omega^2/4b^2)\cos^2\Omega z}\boldsymbol{\sigma}_1\right]\mathbf{B}_2\,, \\
&\cdots \qquad (33)
\end{aligned}
$$

where $c_2 = [b^2 + (\epsilon^2\Omega^2/4)\cos^2\Omega z]^{1/2}$. The equations for $\mathbf{B}_n$ become progressively more complicated, but the qualitative behavior is already apparent from Eqs. (33). As n increases, the off-diagonal term is multiplied by successively higher powers of Ω/b, and successively higher powers of Ω/b appear in $c_n{}^2$. While it is not apparent that this sequence will be convergent even when $\Omega/b \ll 1$, it is apparent that at least the first few terms are rapidly decreasing, so this expansion will yield a useful approximation. This result is analogous to the theory of the driven harmonic oscillator.[45]

It is useful to point out that the trajectory on the Poincaré sphere is quasi-periodic, a result that is analogous to Bloch's theorem in solid-state physics.[48] To obtain this result, we note that

$$
\mathbf{B}_1(z) = \mathsf{T}_1(z)\mathbf{B}_{1,0} = \mathsf{T}_1(s)\mathsf{T}_1{}^m(2\pi/\Omega)\mathbf{B}_{1,0}\,, \qquad (34)
$$

where m is the integer part of $\Omega z/2\pi$ and s is the remainder, so that $z = 2\pi m/\Omega + s$. Since $\mathsf{T}_1(2\pi/\Omega)$ is a unitary matrix, we may write

$$
\mathsf{T}_1(2\pi/\Omega) = \mathsf{R}^{-1}\mathsf{D}\,\mathsf{R}\,, \qquad (35)
$$

where $\mathsf{D} \equiv \cos(2\pi\alpha/\Omega)\mathsf{I} + i\sin(2\pi\alpha/\Omega)\boldsymbol{\sigma}_3$ is diagonal and R is the diagonalizing transformation, which must be unitary. Defining $\mathsf{D}(z) \equiv \cos(\alpha z)\mathsf{I} + i\sin(\alpha z)\boldsymbol{\sigma}_3$, we obtain

$$
\begin{aligned}
\mathbf{B}_1(z) &= \mathsf{T}_1(s)\mathsf{R}^{-1}\mathsf{D}^{-1}(s)\mathsf{D}(z)\mathsf{R}\mathbf{B}_{1,0}\,, \\
&= \mathsf{R}^{-1}\mathsf{S}(z)\mathsf{D}(z)\mathsf{R}\mathbf{B}_{1,0}\,, \qquad (36)
\end{aligned}
$$

where $\mathsf{S}(z) = \mathsf{R}\mathsf{T}_1(s)\mathsf{R}^{-1}\mathsf{D}^{-1}(s)$ is evidently periodic, with a period $2\pi/\Omega$. Thus the motion combines the frequencies α and Ω. When these frequencies are incommensurable, the trajectory will ergodically fill some region on the Poincaré sphere; when they are commensurable, the trajectory will execute a Lissajous figure on the sphere. From Eqs. (33) and the definition of c_2, we find that $\alpha = b + \epsilon^2\Omega^2/16b$ through second order in Ω/b.

4. APPLICATION TO RANDOM VARIATION

We now turn to calculations of the polarization mode dispersion. We begin by considering once again Eq. (13),

$$\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}z} = i\boldsymbol{\Psi}_1\mathbf{B}_1, \tag{37}$$

where

$$\boldsymbol{\Psi}_1 = b\boldsymbol{\sigma}_3 + \frac{\theta_z}{2}\boldsymbol{\sigma}_2 = \begin{bmatrix} b & -i\theta_z/2 \\ i\theta_z/2 & -b \end{bmatrix}. \tag{38}$$

In a similar manner, $\mathbf{B}_1$ will transform as a function of frequency so that[30–39]

$$\frac{\mathrm{d}\mathbf{B}_1}{\mathrm{d}\omega} = i\mathsf{F}_1\mathbf{B}_1, \tag{39}$$

where ω is angular frequency and F_1 is a traceless, Hermitian matrix. The eigenvectors of F_1 are the principal states of polarization, i.e., the directions of $\mathbf{B}_1$ that do not change as a function of ω to first order.[31–39] Writing the eigenvalues of F_1 as λ_1 and $-\lambda_1$, we find that $\tau_d = 2\lambda_1$ is the polarization mode dispersion. Alternatively, we may write ${\tau_d}^2 = -4\det\mathsf{F}_1$. Taking the z derivative of Eq. (39) and the ω derivative of Eq. (37), we find

$$\frac{\partial\mathsf{F}_1}{\partial z} = \frac{\partial\boldsymbol{\Psi}_1}{\partial\omega} + i[\boldsymbol{\Psi}_1, \mathsf{F}_1]. \tag{40}$$

We now specify physical models for b and θ_z that allow us to determine $\boldsymbol{\Psi}_1$ and $\partial\boldsymbol{\Psi}_1/\partial\omega$. The quantity b, which measures the birefringence strength, equals half the magnitude of the phase difference per unit distance between the two eigenmodes. To the extent that the phase and the group velocities are the same, which is ~10% in optical fibers,[49] then $b \propto \omega$. It is certainly legitimate to treat the variation of b as a linear function of ω over a physically relevant bandwidth, setting $b = c + b'\omega$, where c and b' are independent of ω and depend only on z. At the same time θ_z, which is due to the intrinsic core ellipticity or doping anisotropy, is not expected to depend significantly on ω, and we will treat it as depending only on z. We conclude that

$$\frac{\partial\boldsymbol{\Psi}_1}{\partial\omega} = b'\boldsymbol{\sigma}_3. \tag{41}$$

To complete our physical description, we must specify the variations of b and θ_z as functions of z. We assume that b' is proportional to b with a fixed constant of proportionality, since the phase and the group velocities in optical fibers are nearly the same.[49] There is little that is known experimentally about the variation of b and θ_z, and we consider here two physically plausible models. In the first model we assume that θ varies randomly, while b and hence b' are constant. In the second model we assume that $r = b\cos\theta$ and $s = b\sin\theta$ vary randomly, are Gaussian distributed, and are uncorrelated with each other. As discussed in Section 2, we also assume in both models that $\theta_z/2b$ is small in magnitude, while $(b\theta_{zz} - b_z\theta_z)/2b^3$ is large in magnitude on average. We briefly discuss at the end of this section what happens if the assumption that $|\theta_z/2b| \ll 1$ on average is relaxed.

The physical models that we have described here closely resemble the phenomenological models of Curti *et al.*[37] and Poole *et al.*[34] The principal distinction with the model of Poole *et al.* is that Poole *et al.* assume that there are three independent driving fields in $\boldsymbol{\Psi}_1$, whereas in our models there are only one or two driving fields. This distinction has no significant effect on the behavior of the polarization mode dispersion. The reason is that in either of the physical models that we are considering, the variable $\mathbf{B}_1$ is completely randomized over a sufficiently long length, a result that has been amply verified in simulations.[28,50] The same holds for the phenomenological model of Poole *et al.* Equivalently, one can say that the vector (X_1, Y_1, Z_1) corresponding to $\mathbf{B}_1$ is randomized on the Poincaré sphere. Writing

$$\mathsf{F}_1 = f_X\boldsymbol{\sigma}_3 + f_Y\boldsymbol{\sigma}_1 + f_Z\boldsymbol{\sigma}_2, \tag{42}$$

we find that the effect of the second term in Eq. (40) is to randomize the direction of the vector (f_X, f_Y, f_Z) without changing its magnitude, which equals $\tau_d/2$. The first term in Eq. (40) does affect τ_d, but, since this term is always oriented in the $+f_X$ direction, its effect on τ_d depends on the orientation of the vector (f_X, f_Y, f_Z). We thus find that

$$\frac{\mathrm{d}\tau_d}{\mathrm{d}z} = 2b'\cos\xi, \tag{43}$$

where ξ is the angle that the vector (f_X, f_Y, f_Z) makes with respect to the $+f_X$ direction. If ξ is completely randomized over the Poincaré sphere, then Eq. (43) is precisely analogous to the equation that determines the magnitude of a particle's excursion from the origin in Brownian motion.[38] As a consequence, τ_d will become Maxwellian distributed over a length scale that is long compared with that over which the direction of (f_X, f_Y, f_Z) is randomized on the Poincaré sphere, even though, as was pointed out earlier by Poole *et al.*,[34] the physical process that we are considering is quite different from Brownian motion.

We next determine the rate at which the expected value of τ_d increases and find the constraints that θ_z and b must satisfy so that we can carry out an appropriate spatial average. This issue is significant, because many experimental measurements of τ_d have been performed with averaging over frequency and even temperature,[33,34] while in communication systems it is the spatial average that is of practical importance.

The approach that we follow is closely related to that of Poole.[32] We begin by recalling the relationship $\mathbf{B}_1(z, \omega) = \mathsf{T}_1(z, \omega)\mathbf{B}_{1,0}$, from which it follows that

$$\mathsf{F}_1\mathsf{T}_1 = -i\frac{\partial\mathsf{T}_1}{\partial\omega}, \tag{44}$$

$${\tau_d}^2 = -4\det\mathsf{F}_1 = 4\det(\partial\mathsf{T}_1/\partial\omega), \tag{45}$$

where we note that $\det\mathsf{T}_1 = 1$. It is convenient to use T_1 rather than F_1 to calculate the expectation of ${\tau_d}^2$. Assuming that $|\theta_z/2b| \ll 1$ when averaged over the length of the fiber, the evolution of T_1 contains a rapidly varying portion that is due to the birefringence, and it is useful to remove this rapidly varying portion so that we can examine the slower variation that is due to θ_z. We may do so by defining

$$\mathbf{S} = \exp[-i\varphi(z)\boldsymbol{\sigma}_3]\mathbf{T}_1 = \{\cos[\varphi(z)]\mathbf{I} - i\sin[\varphi(z)]\boldsymbol{\sigma}_3\}\mathbf{T}_1, \tag{46}$$

where

$$\varphi(z) = \int_0^z b(z')\mathrm{d}z'. \tag{47}$$

We then find that

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}z} = \begin{bmatrix} 0 & \frac{1}{2}\theta_z\exp(-2i\varphi) \\ -\frac{1}{2}\theta_z\exp(2i\varphi) & 0 \end{bmatrix}\mathbf{S}. \tag{48}$$

From Eq. (45), we now obtain

$$\tau_d^{\,2} = 4\det(\mathbf{S}' + i\boldsymbol{\sigma}_3\varphi'\mathbf{S}), \tag{49}$$

where the primes indicate derivatives with respect to ω. Writing the components of $\mathbf{S}$ explicitly so that

$$\mathbf{S} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \tag{50}$$

and taking the first derivative of $\tau_d^{\,2}$, we obtain

$$\frac{\mathrm{d}\tau_d^{\,2}}{\mathrm{d}z} = 8b'\varphi' + 8ib'(s_1s_1'^* + s_2s_2'^*), \tag{51}$$

where we have neglected the z derivatives of s_1, s_1', s_2, and s_2', since we assume that they are slowly varying. Taking the z derivative of Eq. (51), we obtain

$$\begin{aligned}\frac{\mathrm{d}^2\tau_d^{\,2}}{\mathrm{d}z^2} = {}& 8(b')^2 + 8\varphi'\frac{\mathrm{d}b'}{\mathrm{d}z} - 4ib'\theta_z[s_1s_2'\exp(2i\varphi) \\ & - s_1^*s_2'^*\exp(-2i\varphi) - s_1's_2\exp(2i\varphi) \\ & + s_1'^*s_2^*\exp(-2i\varphi)].\end{aligned} \tag{52}$$

Although $\mathbf{S}$ is slowly varying, its z derivative varies rapidly, and we can use this fact to replace s_1 and s_2 in Eq. (52) with the integral terms:

$$\begin{aligned}s_1(z) &= s_1(\zeta) - \int_\zeta^z \frac{\theta_z(z')}{2}\exp[-2i\varphi(z')]s_2^*(z')\mathrm{d}z' \\ &\simeq s_1(\zeta) - s_2^*(\zeta)\int_\zeta^z \frac{\theta_z(z')}{2}\exp[-2i\varphi(z')]\mathrm{d}z',\end{aligned} \tag{53}$$

and similarly

$$s_2(z) \simeq s_2(\zeta) + s_1^*(\zeta)\int_\zeta^z \frac{\theta_z(z')}{2}\exp[-2i\varphi(z')]\mathrm{d}z', \tag{54}$$

where $z - \zeta$ is a length that is long compared with the variation of θ_z and short compared with the variations of s_1 and s_2. The assumed existence of this intermediate-scale length, of the order of the birefringent beat length, plays a critical role in our argument. We also find

$$\begin{aligned}s_1'(z) \simeq {}& s_1'(\zeta) - [s_2'^*(\zeta) - 2i\varphi'(\zeta)s_2^*(\zeta)] \\ & \times\int_\zeta^z \frac{\theta_z(z')}{2}\exp[-2i\varphi(z')]\mathrm{d}z', \\ s_2'(z) \simeq {}& s_2'(\zeta) + [s_1'^*(\zeta) - 2i\varphi'(\zeta)s_1^*(\zeta)] \\ & \times\int_\zeta^z \frac{\theta_z(z')}{2}\exp[-2i\varphi(z')]\mathrm{d}z'.\end{aligned} \tag{55}$$

We now define

$$\begin{aligned}H(z) &= \mathrm{Re}\left[\int_\zeta^z \frac{\theta_z(z)\theta_z(z')}{4}\exp[2i\varphi(z) - 2i\varphi(z')]\,\mathrm{d}z'\right] \\ &\simeq \frac{1}{2}\int_{-\infty}^\infty \frac{\theta_z(z)\theta_z(z')}{4}\exp[2i\varphi(z) - 2i\varphi(z')]\mathrm{d}z',\end{aligned} \tag{56}$$

where we have used the assumed rapid variation of θ_z to extend the integral to $-\infty$. The quantity $H(z)$ is closely related to the spectral density of the autocorrelation function $R(\theta_z/2)$. Physically the autocorrelation length must be small compared with the birefringent beat length for $H(z)$ to differ significantly from zero. Using relations (53)–(56) in Eq. (52), we obtain

$$\begin{aligned}\frac{\mathrm{d}^2\tau_d^{\,2}}{\mathrm{d}z^2} = {}& 8(b')^2 + 8\varphi'\frac{\mathrm{d}b'}{\mathrm{d}z} - 32b'\varphi'H(z) \\ & - 32ib'H(z)(s_1s_1'^* + s_2s_2'^*),\end{aligned} \tag{57}$$

where we have dropped both small terms and rapidly varying terms that will not contribute on the length scale $z - \zeta$. The quantity $\mathrm{d}b'/\mathrm{d}z$ is zero in the first physical model, and it is rapidly varying in the second physical model on the length scale $z - \zeta$, so one might be tempted simply to drop the second term on the right-hand side of Eq. (57). However, the mean of $\mathrm{d}b'/\mathrm{d}z$ varies on a length scale that is comparable with the length scale on which b' itself varies. Hence we keep this contribution for now.

In our first physical model of an optical fiber the birefringence strength is fixed, and we can assume that $H(z)$ is a constant independent of z. Additionally, we obtain $\varphi = bz$ and $\varphi' = b'z$. In this case Eqs. (51) and (57) are the same as the equations found by Poole[32] and can be dealt with in a similar fashion, although some thought must be given to the appropriate way of replacing the ensemble average over "a collection of statistically equivalent fibers" with a spatial ensemble average. By contrast, in the second physical model, in which b and θ_z both vary, they are evidently correlated. When the birefringence strength is larger, the value of $\theta_z^{\,2}$ will be smaller on average. Since by assumption both b and θ change slowly, we may replace $\varphi(z) - \varphi(z')$ in relation (56) with $b(z - z')$, where b is the local value. As a consequence of the correlation just mentioned, $H(z)$ will have a slow variation on the same length scale on which b changes, and we cannot assume that b and H are uncorrelated. We can, however, assume that φ and the combination $s_1s_1'^* + s_2s_2'^*$ are uncorrelated with b and H over a sufficiently long length, since both φ and the combination $s_1s_1'^* + s_2s_2'^*$ are obtained by an integration from the origin to the local position. After a long length the local contribution of b and H to these integrals will be negligible. From our earlier discussion at the beginning of this section concerning the evolution of the vector (f_X, f_Y, f_Z), it follows that this length scale is the scale on which this vector samples the entire Poincaré sphere. To carry out a spatial average of Eqs. (51) and (57), we must average on a length scale L that is long compared with this randomization distance on the Poincaré sphere. We must also assume that τ_d is already large, so that it does not change significantly with respect to its own value over the length scale of the average. Under these circumstances we may define the ensemble average

$$\langle X(z)\rangle = \frac{1}{L}\int_{z-L}^{z} X(z')\,\mathrm{d}z', \tag{58}$$

and, averaging on both sides of Eqs. (51) and (57), we obtain

$$\frac{\mathrm{d}\langle {\tau_d}^2\rangle}{\mathrm{d}z} = 8\langle b'\rangle\langle\varphi'\rangle + 8i\langle b'\rangle\langle s_1{s_1}'^{*} + s_2{s_2}'^{*}\rangle,$$
$$\frac{\mathrm{d}^2\langle {\tau_d}^2\rangle}{\mathrm{d}z^2} = 8\langle (b')^2\rangle - 32\langle\varphi'\rangle\langle b'H\rangle - 32i\langle b'H\rangle\langle s_1{s_1}'^{*} + s_2{s_2}'^{*}\rangle, \tag{59}$$

where we note $\langle\varphi'\mathrm{d}b'/\mathrm{d}z\rangle = \langle\varphi'\rangle\langle\mathrm{d}b'/\mathrm{d}z\rangle = 0$. Writing now $\langle b'H\rangle = r\langle b'\rangle\langle H\rangle$, we obtain the ordinary differential equation

$$\frac{\mathrm{d}^2\langle {\tau_d}^2\rangle}{\mathrm{d}z^2} = -4r\langle H\rangle\frac{\mathrm{d}^2\langle {\tau_d}^2\rangle}{\mathrm{d}z} + 8\langle (b')^2\rangle, \tag{60}$$

where the quantity r is equal 1 in our first physical model but will not equal 1 in general in our second physical model. Equation (60) has the solution

$$\langle {\tau_d}^2\rangle = \frac{\langle (b')^2\rangle}{2r^2\langle H\rangle^2}[\exp(-4r\langle H\rangle z) - 1 + 4r\langle H\rangle z], \tag{61}$$

which equals 0 at $z = 0$.

The solution that appears in Eq. (61) is physically meaningful at any z when an ensemble average is carried out over different fibers or over frequency and temperature. However, Eq. (61) is physically meaningful only at large z when the spatial ensemble average defined in Eq. (58) is carried out. In this limit only the last term in Eq. (61) contributes significantly, and Eq. (61) becomes $\langle {\tau_d}^2\rangle = 2\langle (b')^2\rangle z/r\langle H\rangle$, exhibiting the linear dependence that is expected for a random walk. While we anticipate on physical grounds that ${\tau_d}^2 \propto z^2$ when z is small, it is evidently not possible to relate the coefficient to averaged fiber parameters, since the coefficient will depend on local fiber parameters at $z = 0$.

In the large-z limit we have determined the dependence of $\langle {\tau_d}^2\rangle$ on averaged fiber parameters $\langle (\beta')^2\rangle$, r, and $\langle H\rangle$ under restrictive assumptions that require that the autocorrelation distance for θ_z be short compared with the beat length and that the length scale on which the electric-field vector wanders over the entire Poincaré sphere be long compared with the beat length. If we relax these assumptions, then we still anticipate from the discussion at the beginning of this section that $\langle {\tau_d}^2\rangle \propto z$ when z becomes large. Determining the behavior of this coefficient as fiber parameters are varied is of considerable practical importance and will be the topic of future studies.

5. CONCLUSIONS

In this paper we have developed a procedure using repeated diagonalization, which allows us to solve the equations that govern field evolution in birefringent optical fibers. This procedure allows us to deal in a unified way with several practically important cases. These include twisted and sinusoidally rocked fibers. They also include fibers with randomly varying birefringence. In particular, we have shown that a phenomenological model developed by Poole and other workers applies to optical fibers whose axes of birefringence can take on any orientation.

ACKNOWLEDGMENTS

This work was supported by Department of Energy grant DE-FG05-89ER-14090 and National Science Foundation grant ECS-9113382.

REFERENCES AND NOTES

1. R. H. Stolen, V. Ramaswamy, P. Kaiser, and W. Pleibel, "Linear polarization in birefringent single-mode fibers," Appl. Phys. Lett. **33**, 699–701 (1978).
2. I. P. Kaminow, "Polarization in fibers," Laser Focus **16**(6), 80–84 (1980).
3. I. P. Kaminow, "Polarization in optical fibers," IEEE J. Quantum Electron. **QE-17**, 15-22 (1981).
4. L. Li, J. R. Qian, and D. N. Payne, "Current sensors using highly birefringent bow-tie fibers," Electron. Lett. **22**, 1142–1144 (1986).
5. R. I. Laming and D. N. Payne, "Electric current sensors employing spun highly birefringent optical fibers," J. Lightwave Technol. **7**, 2084–2094 (1989).
6. A. D. Kersey, A. Dandridge, and A. B. Tveten, "Dependence of visibility on input polarization in interferometric fiber-optic sensors," Opt. Lett. **13**, 288–290 (1988).
7. A. D. Kersey, M. J. Marone, and A. Dandridge, "Observation of input-polarization-induced phase noise in interferometric fiber-optics sensors," Opt. Lett. **13**, 847–849 (1988).
8. W. K. Burns, C.-L. Chen, and R. P. Moeller, "Fiber-optic gyroscopes with broad-band sources," J. Lightwave Technol. **1**, 98–105 (1983).
9. W. K. Burns, "Phase error bounds of fiber gyro with polarization holding fiber," J. Lightwave Technol. **4**, 8–14 (1986).
10. M. N. Islam, *Ultrafast Fiber Switching Devices and Systems* (Cambridge U. Press, Cambridge, 1992).
11. K. Kitayama, Y. Kimura, and S. Seikai, "Fiber-optic logic gate," Appl. Phys. Lett. **46**, 317–319 (1985).
12. H. G. Winful, "Polarization instabilities in birefringent nonlinear media: application to fiber-optic devices," Opt. Lett. **11**, 33–35 (1986).
13. T. Morioka, M. Saruwatari, and A. Takada, "Ultrafast optical multi/demultiplexer utilizing optical Kerr effect in polarisation-maintaining single-mode fibers," Electron. Lett. **23**, 453–454 (1987).
14. M. J. Marrone and C. A. Villaruel, "Fiber in-line polarization rotator and mode interchanger," Appl. Opt. **26**, 3194–3195 (1987).
15. R. H. Stolen, A. Ashkin, W. Pleibel, and J. M. Dziedzic, "In-line fiber-polarization-rocking rotator and filter," Opt. Lett. **9**, 300–302 (1984).
16. B. Daino, G. Gregori, and S. Wabnitz, "New all-optical devices based on third-order nonlinearity of birefringent fibers," Opt. Lett. **11**, 42–44 (1986).
17. S. C. Rashleigh, W. K. Burns, R. P. Moeller, and R. Ulrich, "Polarization holding in birefringent single-mode fibers," Opt. Lett. **7**, 40–42 (1982).
18. W. K. Burns and R. P. Moeller, "Measurement of polarization dispersion in high-birefringence fibers," Opt. Lett. **8**, 195–197 (1983).
19. A. Simon and R. Ulrich, "Evolution of polarization along a single-mode fiber," Appl. Phys. Lett. **31**, 517–520 (1977).

20. V. Ramaswamy, R. D. Standley, D. Sze, and W. G. French, "Polarization effects in short length, single mode fibers," Bell Syst. Tech. J. **57,** 635–651 (1978).
21. A. J. Barlow, J. J. Ramskov-Hansen, and D. N. Payne, "Birefringence and polarization mode-dispersion in spun single-mode fibers," Appl. Opt. **20,** 2962–2968 (1981).
22. S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," J. Lightwave Technol. **1,** 312–331 (1983).
23. M. J. Marrone, C. A. Villaruel, N. J. Frigo, and A. Dandridge, "Internal rotation of the birefringence axes in polarization-holding fibers," Opt. Lett. **12,** 60–62 (1987).
24. R. Calvani, R. Caponi, and F. Cisternino, "Polarization measurements on single-mode fibers," J. Lightwave Technol. **7,** 1187–1196 (1989).
25. N. S. Bergano, "Undersea lightwave transmission systems using Er-doped fiber amplifiers," Opt. Photon. News **4**(1), 8–14 (1993).
26. L. F. Mollenauer, K. Smith, J. P. Gordon, and C. R. Menyuk, "Resistance of solitons to the effects of polarization dispersion in optical fibers," Opt. Lett. **14,** 1219–1221 (1989).
27. S. G. Evangelides, L. F. Mollenauer, and J. P. Gordon, "Polarization multiplexing with solitons," J. Lightwave Technol. **10,** 28–34 (1992).
28. 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).
29. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, Boston, 1991).
30. N. S. Bergano, C. D. Poole, and R. E. Wagner, "Investigation of polarization dispersion in long lengths of single-mode fiber using multilongitudinal mode lasers," J. Lightwave Technol. **5,** 1618–1622 (1987).
31. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," Electron. Lett. **22,** 1029–1030 (1986).
32. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," Opt. Lett. **13,** 687–689 (1988).
33. C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random coupling," Opt. Lett. **14,** 523–525 (1989).
34. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," Opt. Lett. **16,** 372–374 (1991).
35. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," J. Lightwave Technol. **9,** 1439–1456 (1991).
36. D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurement of the group-delay difference between principal states of polarization on a low-birefringence terrestrial fiber cable," Opt. Lett. **12,** 844–846 (1987).
37. F. Curti, D. Daino, Q. Mao, F. Matera, and C. G. Someda, "Concatenation of polarisation dispersion in single-mode fibres," Electron. Lett. **25,** 290–292 (1989).
38. F. Curti, B. Daino, G. De Marchis, and F. Maternola, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," J. Lightwave Tech. **8,** 1162–1166 (1990).
39. S. Betti, F. Curti, B. Daino, G. De Marchis, E. Iannone, and F. Matera, "Evolution of the bandwidth of the principal states of polarization in single-mode fibers," Opt. Lett. **16,** 467–469 (1991).
40. See, e.g., the discussion at the beginning of Chap. 1 of A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1984).
41. H. Poincaré, *Théorie Mathématique de la Lumière* (Georges Carré, Paris, 1892), Vol. 2, Chap. 12; see also M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light* (Pergamon, Oxford, 1984), Chap. 1.
42. N. J. Frigo, "A generalized geometrical representation of coupled mode theory," IEEE J. Quantum Electron. **QE-22,** 2131–2140 (1986).
43. D. David, D. D. Holm, and M. V. Tratnik, "Hamiltonian chaos in nonlinear optical polarization dynamics," Phys. Rep. **187,** 218–367 (1990).
44. G. B. Arfken, *Mathematical Methods for Physicists* (Academic, New York, 1985), Chap. 4.10.
45. See, e.g., H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1980).
46. C. R. Menyuk, "Pulse propagation in an elliptically birefringent Kerr medium," IEEE J. Quantum Electron. **25,** 2674–2682 (1989).
47. C. R. Menyuk and P. K. A. Wai, "Elimination of nonlinear polarization rotation in twisted fibers," J. Opt. Soc. Am. B **11,** 1305–1309 (1994).
48. C. Kittel, *Introduction to Solid State Phyics* (Wiley, New York, 1986), Chap. 7. In the mathematical community, this result is typically referred to as Floquet's theorem. See, e.g., G. Blanch, "Mathieu functions," in *Handbook of Mathematical Functions,* M. Abramowitz and I. A. Stegun, eds. (U.S. Government Printing Office, Washington, D.C., 1964). Note that in contrast to the usual result, the Bloch–Floquet exponent can only be imaginary, because the motion is confined to a sphere. Mathematically, $\mathbf{D}$ is unitary, so its eigenvalues can have only imaginary exponents.
49. G. Keiser, *Optical Fiber Communications* (McGraw-Hill, New York, 1991).
50. T. Ueda and W. Kath, "Dynamics of optical pulses in randomly birefringent fibers," Physica (The Hague) **D55,** 166–181 (1992).