Application of the Manakov-PMD Equation to Studies of Signal Propagation in Optical Fibers with Randomly Varying Birefringence

Dietrich Marcuse, *Fellow, IEEE,* C. R. Menyuk, *Senior Member, IEEE,* and P. K. A. Wai, *Senior Member, IEEE*

***Abstract*— We report here on our investigations of the Manakov-polarization mode dispersion (PMD) equation which can be used to model both nonreturn-to-zero (NRZ) and soliton signal propagation in optical fibers with randomly varying birefringence. We review the derivation of the Manakov-PMD equation from the coupled nonlinear Schrödinger equation, and we discuss the physical meaning of its terms. We discuss our numerical approach for solving this equation, and we apply this approach to both NRZ and soliton propagation. We show by comparison with the coupled nonlinear Schrödinger equation, integrated with steps that are short enough to follow the detailed polarization evolution, that our approach is orders of magnitude faster with no loss of accuracy. Finally, we compare our approach to the widely used coarse-step method and demonstrate that the coarse-step method is both efficient and valid.**

***Index Terms*—Birefringence, optical fibers, polarizations, random bifringence, solitons.**

I. Introduction

BECAUSE of their low loss, optical fibers have become an important medium for long-distance communications. Even though the glass fiber is a remarkably linear transmission medium, its slight nonlinearity, due primarily to the Kerr effect, became important with the advent of the erbium-doped optical amplifier that permitted transmission of light signals through optical fibers over transoceanic distances without regeneration, allowing the effect of the nonlinearity to accumulate over the entire distance.

When light travels in an optical fiber for thousands of kilometers, the effect of the nonlinearity can no longer be ignored. On the one hand, it can lead to significant signal distortion—particularly when coupled with the spontaneous emission noise due to the erbium-doped fiber amplifiers (EDFA's); on the other hand, it can be used to counteract dispersion, as in solitons. The development of communications technology has made it necessary to develop mathematical tools for describing light transmission in nonlinear optical fibers. Hasegawa and Tappert [1] showed that when the slowly varying envelope approximation applies, which is the case in communication signals to date [2], and fiber birefringence can be ignored, then the transmission is well-described by the nonlinear Schrödinger equation. Unfortunately, it is often impossible to ignore the fiber birefringence, particularly in modern-day NRZ (nonreturn-to-zero) communications, in which polarization scrambling is consistently used. Menyuk [3] showed that the CNLS (coupled nonlinear Schrödinger equation) applies to birefringent optical fibers, but early work focused on fibers with constant birefringence [4]. In the parameter regime in which present-day communication systems operate, the birefringence is large in the sense that the beat lengths are typically 10–100 m, while the dispersive and nonlinear scale lengths are hundreds of kilometers—but is rapidly varying in the sense that the length scale on which the birefringence orientation varies is 0.3 to 300 m. When nonlinearity can be ignored, Poole and Wagner [5]–[7] showed that the large but rapidly and randomly varying birefringence leads to PMD (polarization mode dispersion), which is a very important effect in communication fibers. An important advance in dealing with both nonlinearity and the rapidly varying birefringence was made when it was recognized that if the polarization states of the light is averaged over the Poincaré sphere, the Manakov equation can be derived from the CNLS [8], [9]. However, the assumption that the polarization states are mixed on the Poincaré sphere was not justified, and the nature of the correction terms was not completely elucidated, although it was recognized that they are important [8]; it is these terms that are responsible for PMD. Since then, in a series of works that carefully investigated the statistics on the Poincaré sphere using physically reasonable models [10], [11], we were able to justify the assumption in [8] and [9] and to elucidate the nature of the corrections to the Manakov equation. This work culminated in the derivation of the Manakov-PMD equation [11] which it is the object of this paper to study.

It is our view that the derivation of the Manakov-PMD equation completely resolves the outstanding theoretical issues concerning appropriate models for modeling birefringence in optical fibers. Moreover, this equation leads to highly rapid numerical schemes and can be used to justify the widely used coarse-step method. From a physical standpoint, the mathematical transformation that allows us to derive the

Manuscript received November 8, 1996. This work was supported by ARPA through the AFOSR, NSF, and DOE.

D. Marcuse and C. R. Menyuk are with the Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21228 USA.

P. K. A. Wai is with the Department of Electronic Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

Publisher Item Identifier S 0733-8724(97)06657-7.

Manakov-PMD equation from the CNLS equation amounts to freezing the linear motion of the signal's central frequency on the Poincaré sphere. Since the motion of the other frequencies in the signal and the changes due to nonlinearity are slow relative to this point's motion on the sphere, we can take long computational steps. The terms that appear in the Manakov-PMD equation have a transparent physical meaning. There are two terms that account for the Kerr nonlinearity averaged over the Poincaré sphere and chromatic dispersion; these are the terms that appear in the Manakov equation. There is another term that corresponds to the usual linear PMD and when both chromatic dispersion and nonlinearity can be neglected, this term yields all the effects that Poole and others have described. Finally, there are additional nonlinear terms that have not been previously discussed in detail (but see [11]) and that we refer to as nonlinear PMD terms. We will show that these terms, which are due physically to incomplete mixing on the Poincaré sphere, are vanishingly small in cases of present-day, practical importance in optical communications and can safely be ignored in these cases.

At the present time, most simulations of polarization effects in long-distance optical fiber transmission are done using a somewhat *ad hoc* procedure that is discussed in [8] and [9] and that we refer to here as the coarse-step method. The starting point for this approach is the CNLS without the rapidly varying terms that is described in [3] and [4]. One then integrates these equations using step sizes that are large compared to the fiber correlation length—typically several kilometers—and then randomly rotates the data on the Poincaré sphere. This approach leads to an artificially increased fiber correlation length, but one also artificially reduces the strength of the birefringence. By doing so, one can obtain the same magnitude for the linear PMD as one would in simulations which use realistic fiber correlation lengths and birefringence strengths, and one realizes an enormous saving in computation time. We shall show in this paper that this approach, as it has been used to date, both exaggerates the nonlinear PMD and deals with it inaccurately, but we will also show that the nonlinear PMD is so small in cases of present-day practical importance in long-distance communication systems that this exaggeration does not matter. Moreover, we will show how to eliminate this difficulty. Thus, our work provides a theoretical justification for the use of the coarse-step method.

The reminder of this paper is organized as follows: In Section II, we review the derivation of the Manakov-PMD equation, and in Section III, we discuss the numerical algorithms for solving it. In Section IV, we apply this equation to examples of NRZ and soliton communication signals, and we discuss the nonlinear PMD. In Section V, we compare the Manakov-PMD equation to the coarse-step method. Section VI contains the conclusions.

II. Derivation of the Manakov-PMD Equation

In this section, we will review the derivation of the Manakov-PMD equation that was earlier discussed in [11], presenting it from a more physical perspective. Optical fibers used in telecommunications are designed to support only a single guided mode which, however, can exist in two mutually orthogonal polarization [12]. In a perfectly isotropic, circularly symmetric fiber the two polarization components of the mode would travel with identical phase and group velocities. Unavoidable fiber imperfections break this polarization degeneracy so that at each point along the fiber there exists a preferred direction perpendicular to its axis. A light field polarized orthogonal to this direction has the smallest phase and group velocity. We designate the propagation constant of the wave polarized in the preferred direction as β_1 and the propagation constant of the orthogonally polarized wave as β_2. The difference of these two propagation constants is defined as the birefringence parameter $b = (\beta_1 - \beta_2)/2$. The specific group delay per unit length is represented by $b' = (\beta_1' - \beta_2')/2$ where the primes indicate differentiation with respect to the angular frequency. In this paper, in contrast to our earlier paper [11], we do not use normalized units so that β and b have the dimension of the inverse length, while β' and b' have dimensions of time divided by length. In this paper, we also use α to designate the preferred direction, rather than $\theta = 2\alpha$ as in [11] since α is the geometric rotation angle. The angle α rotates randomly; b and b' may also vary randomly.

Optical signals in fibers can be distorted due to a combination of several effects. First, if $b' \neq 0$, the signals undergo PMD [5]–[7]. Second, if $\beta_1'' \approx \beta_2'' \neq 0$, the signals undergo chromatic dispersion. Finally, optical fibers are weakly nonlinear which means that the refractive index of the fiber material changes as a function of the light intensity. The propagation of light pulses, taking all these effects into account, can be described by the coupled nonlinear Schrödinger equation [3], [4]

$$i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma\mathbf{A} + ib'\Sigma\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{A}}{\partial t^2} + n_2 k_0\left[\frac{5}{6}|\mathbf{A}|^2\mathbf{A} + \frac{1}{6}(\mathbf{A}^\dagger\sigma_3\mathbf{A})\sigma_3\mathbf{A} + \frac{1}{3}\mathbf{B}\right] = 0 \quad (1)$$

or equivalently

$$i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma\mathbf{A} + ib'\Sigma\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{A}}{\partial t^2} + n_2 k_0\left[|\mathbf{A}|^2\mathbf{A} - \frac{1}{3}(\mathbf{A}^\dagger\sigma_2\mathbf{A})\sigma_2\mathbf{A}\right] = 0 \quad (2)$$

which are written in a form that is appropriate for a fiber with a randomly varying birefringence orientation. Here, $\mathbf{A} = (A_1, A_2)^t$ is a column vector with elements A_1 and A_2 which are the complex envelopes of the two polarization components. The vector $\mathbf{B}$ is defined as $(A_1^* A_2^2, A_1^2 A_2^*)^t$, where $*$ designates complex conjugation. The z-coordinate measures distance along the fiber axis, while the t-coordinate represents retarded time—the time relative to the moving center of the signal. The matrix

$$\Sigma = \sigma_3\cos(2\alpha) + \sigma_1\sin(2\alpha) \quad (3)$$

is defined in terms of the Pauli's matrices

$$\mathsf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (4)$$

and by using the same matrix Σ in the second and third terms of (1), we are effectively assuming that the orientation of the axes of birefringence is only a weak function of frequency. The nonlinear Kerr coefficient is represented by n_2 and the wavenumber by $k_0 = 2\pi/\lambda$, where λ is the vacuum wavelength of light. The parameter β'' in (1) and (2) does not have a subscript since we are assuming that $\beta_1'' \approx \beta_2'' = \beta''$.

A fundamental difficulty in solving (1) or equivalently (2) is that α changes on a length scale of 0.3–100 m, while the length scales for PMD, chromatic dispersion, and the Kerr nonlinearity are hundreds or even thousands of kilometers. If one wishes to solve (1) accurately, taking into account rapid and random variations of the birefringence, one must do so on a very short length which is computationally expensive. Moreover, the different physical effects that contribute to the signal evolution in (1) are mixed together in a complicated way. It is known that when the nonlinearity and chromatic dispersion can be ignored, the signal undergoes polarization mode dispersion [5]–[7]. Conversely, when the PMD is small but the birefringence is rapidly and randomly varying, it is known that the Manakov equation applies. It is not immediately apparent where in (1) these effects are embedded.

The Manakov-PMD equation resolves both these difficulties. It not only allows us to develop a rapid numerical scheme for determining the signal evolution with no loss of accuracy, but the different physical effects are isolated in separate terms of the equation that have a transparent physical meaning. Thus, this equation is the natural starting point for analysis of signal propagation in optical fibers. In particular, we will use it to validate the widely used coarse-step method in the parameter regime that is relevant to long-distance communication systems.

To obtain the Manakov-PMD, we proceed in two successive steps [11]. First, the field $\mathbf{A}$ is transformed from the fixed coordinate system to a coordinate system that rotates with the principal axes of the fiber. The new field is designated by the vector $\mathbf{\Psi}$ that is related to $\mathbf{A}$ as follows:

$$\mathbf{\Psi} = \begin{pmatrix} U \\ V \end{pmatrix} = \mathsf{R}^{-1}\mathbf{A} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{5}$$

If we consider a continuous wave (CW) signal at a single frequency and neglect the fiber nonlinearity, then (1) reduces to

$$i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma\mathbf{A} = 0 \tag{6}$$

and the transformation (5) implies [13]

$$i\frac{\partial \Psi}{\partial z} + \tilde{\Sigma}\mathbf{\Psi} = 0 \tag{7}$$

where, letting $\alpha_z = d\alpha/dz$, we find

$$\tilde{\Sigma} = \begin{pmatrix} b & -i\alpha_z \\ i\alpha_z & -b \end{pmatrix}. \tag{8}$$

Second, we introduce a new field vector $\bar{\mathbf{\Psi}}(z,t)$ via the unitary matrix T

$$\mathbf{\Psi}(z,t) = \mathsf{T}(z)\bar{\mathbf{\Psi}}(z,t) \tag{9}$$

where T is the solution to the ordinary differential equation

$$i\frac{\partial \mathsf{T}}{\partial z} + \tilde{\Sigma}\mathsf{T} = 0 \tag{10}$$

with the initial condition

$$\mathsf{T}(z=0) \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{11}$$

One then finds that (1) yields $\partial\bar{\mathbf{\Psi}}(z,t)/\partial t = 0$ in the linear, CW limit, so that $\bar{\mathbf{\Psi}}$ is frozen on the Poincaré sphere in this limit. In effect, the transformation T follows the rapid and random motion of a single state on the Poincaré sphere. This motion occurs on a length scale of 0.3–100 m. Of course, $\bar{\mathbf{\Psi}}(z,t)$ is not frozen on the Poincaré sphere at all t-points except in the special case of a linear CW field. However, in all cases of present-day interest in communication systems, its evolution is quite slowly varying—on a length scale of hundreds of kilometers, associated with the scale lengths of PMD, chromatic dispersion, and nonlinearity. This physical observation is at the heart of our approach.

After substitution of this transformation into (1), we obtain the Manakov-PMD equation [11]

$$i\frac{\partial \bar{\mathbf{\Psi}}}{\partial z} - \frac{1}{2}\beta''\frac{\partial^2 \bar{\mathbf{\Psi}}}{\partial t^2} + \frac{8}{9}n_0k_2|\bar{\mathbf{\Psi}}|^2\bar{\mathbf{\Psi}} = -ib'\bar{\sigma}\frac{\partial \bar{\mathbf{\Psi}}}{\partial t} - \frac{1}{3}n_0k_2(\hat{\mathbf{N}} - \langle\hat{\mathbf{N}}\rangle). \tag{12}$$

When the right-hand side is zero, (12) reduces to the Manakov equation. The second term on the left-hand side of (12) includes the effect of chromatic dispersion, and the third term on the left-hand side of (12) includes the effect of Kerr nonlinearity averaged over the Poincaré sphere with the well-known 8/9 factor [8], [9]. We will show that the first term on the right-hand side of (12) leads to the usual linear PMD that is discussed in [5]–[7]. Finally, the second term on the right-hand side of (12) is a new term that we refer to as nonlinear PMD. Physically, it is due to incomplete mixing on the Poincaré sphere. To our knowledge, this effect has not been previously discussed except briefly near the end of [11]. In this paper, we will show that its effect is negligible in the parameter regime in which communication fibers normally operate.

There are three quantities, $\bar{\sigma}$, $\hat{\mathbf{N}}$, and $\langle\hat{\mathbf{N}}\rangle$, that require special discussion. The matrix T defined in (9) and (10) is a special unitary matrix so that it has the form

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

with the additional constraint $|u_1|^2 + |u_2|^2 = 1$. We note that our notation differs from that of [11] in which the vector $\mathbf{\Psi}_{\mathrm{CW}} = (u_1, u_2)^t$ now corresponds to the continuous wave signal of (7). The other solution of (7) is given by $\mathbf{\Phi}_{\mathrm{CW}} =$

$(-u_2^*, u_1^*)^t$. We define the Stokes parameters of the CW wave as

$$\begin{aligned} x_1 &= \mathbf{\Psi}_{\mathrm{CW}}^\dagger \sigma_3 \mathbf{\Psi}_{\mathrm{CW}} = |u_1|^2 - |u_2|^2 \\ y_1 &= \mathbf{\Psi}_{\mathrm{CW}}^\dagger \sigma_1 \mathbf{\Psi}_{\mathrm{CW}} = (u_1 u_2^* + u_1^* u_2) \\ z_1 &= \mathbf{\Psi}_{\mathrm{CW}}^\dagger \sigma_2 \mathbf{\Psi}_{\mathrm{CW}} = i(u_1 u_2^* - u_1^* u_2). \end{aligned} \tag{14}$$

Using the solutions $\mathbf{\Psi}_{\mathrm{CW}}$ and $\mathbf{\Phi}_{\mathrm{CW}}$, we also define

$$\begin{aligned} x_2 + i x_3 &= \mathbf{\Phi}_{\mathrm{CW}}^\dagger \sigma_3 \mathbf{\Psi}_{\mathrm{CW}} = -2u_1 u_2, \\ y_2 + i y_3 &= \mathbf{\Phi}_{\mathrm{CW}}^\dagger \sigma_1 \mathbf{\Psi}_{\mathrm{CW}} = u_1^2 - u_2^2, \\ z_2 + i z_3 &= \mathbf{\Phi}_{\mathrm{CW}}^\dagger \sigma_2 \mathbf{\Psi}_{\mathrm{CW}} = i\big(u_1^2 + u_2^2\big). \end{aligned} \tag{15}$$

Defining the real vectors $\mathbf{R}_j = (x_j, y_j, z_j)$ and recalling that at $z = 0$, $u_1 = u_1^* = 1$ and $u_2 = u_2^* = 0$, it can be shown that the $\mathbf{R}_j$ are vectors on the Poincaré sphere [11] and that $\mathbf{R}_1$ is initially pointed in the x-direction, $\mathbf{R}_2$ is initially pointed in the y-direction, and $\mathbf{R}_3$ is initially pointed in the z-direction. As T evolves in accordance with (10), the $\mathbf{R}_j$ will wander on the surface of the Poincaré sphere but will remain orthogonal to each other. We have changed notation from that of [11] to emphasize the geometric connection to vectors on the Poincaré sphere. We now find that

$$\bar{\sigma} = \mathsf{T}^\dagger \sigma_3 \mathsf{T} = \begin{pmatrix} x_1 & x_2 - i x_3 \\ x_2 + i x_3 & -x_1 \end{pmatrix}. \tag{16}$$

We also find that $\hat{\mathbf{N}} = (\hat{\mathbf{N}}_1, \hat{\mathbf{N}}_2)^t$ where

$$\begin{aligned} \hat{\mathbf{N}}_1 &= z_1^2(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} - z_1(z_2 - i z_3)(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} \\ &\quad - z_1(z_2 + i z_3)\bar{U}^2\bar{V}^* - (z_2 - i z_3)^2\bar{V}^2\bar{U}^* \\ \hat{\mathbf{N}}_2 &= z_1^2(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} + z_1(z_2 + i z_3)(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} \\ &\quad + z_1(z_2 - i z_3)\bar{V}^2\bar{U}^* - (z_2 + i z_3)^2\bar{U}^2\bar{V}^* \end{aligned} \tag{17}$$

and

$$\begin{aligned} \langle \hat{\mathbf{N}}_1 \rangle &= \frac{1}{3}(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} \\ \langle \hat{\mathbf{N}}_2 \rangle &= \frac{1}{3}(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V}. \end{aligned} \tag{18}$$

The Manakov-PMD equation is useful from a conceptual standpoint because each of the principal effects that contribute to the signal evolution, the Kerr effect, chromatic dispersion, and polarization mode dispersion are clearly separated. The Manakov-PMD equation is useful from a numerical standpoint because $\bar{\mathbf{\Psi}}(z,t)$ is slowly varying. The x_j and z_j that appear in this equation are rapidly varying, but they are determined by solving (10) which is merely a second-order ordinary differential equation and can be rapidly solved. When solving Eq. (12) on the slowly varying scale, one must properly average over the rapidly varying terms, i.e., one must replace the x_j by $(1/l)\int_z^{z+l} dz' x_j(z')$, and we will discuss this issue in the next section. In principle, it is possible to take advantage of the central limit theorem to replace these averages with appropriate random variables [11]. This issue is currently being studied.

III. Numerical Approach

The numerical approach that we follow in solving the Manakov-PMD equation is closely related to the widely used split-step operator technique that is used to solved the nonlinear Schrödinger equation and the coupled nonlinear Schrödinger equation and is discussed, for example, in [2] and [4]. In this approach, time derivatives are evaluated in the frequency domain while functional multiplications are evaluated in the time domain. For the usual nonlinear Schrödinger equation and the coupled nonlinear Schrödinger equation, which do not take into account such complications as the Raman effect, this prescription implies that the linear terms are evaluated in the frequency domain while the nonlinear terms are evaluated in the time domain. By splitting the spatial step, so that the time domain evaluations are done in the middle of the frequency domain evaluations, it is possible to maintain second-order accuracy with one function evaluation per step for each term at every node point. We also note, because it is important in practice, that certain linear effects like chromatic dispersion do not change the spectral intensity and certain nonlinear effects like the Kerr nonlinearity do not change the temporal intensity. In order to avoid numerical instabilities, it is important that the numerical approximations of these effects retain these properties to far higher than second order, even though the overall scheme is only second-order accurate.

The vectors $\mathbf{A}$ and $\bar{\mathbf{\Psi}}$ are related to each other by a simple unitary transformation which can be found by integrating (10) and is rapidly determined, since it is only a second-order ordinary differential equation, even though it must be integrated on the short length scale given by the beat length and fiber correlation length. In most practical cases, it is not even necessary to transform from one to the other. Since the electrical output current from a realistic detection system is given by $|\mathbf{A}|^2 = |\bar{\mathbf{\Psi}}|^2$, it does not matter which we use. We stress this point because in devising numerical algorithms for solving (12) efficiently, we have found it somewhat advantageous to work with (1) rather than (12).

The linear part of (1) may be written in the form

$$i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma(z)\mathbf{A} + ib'\Sigma(z)\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{A}}{\partial t^2} = 0 \tag{19}$$

which in the Fourier domain becomes

$$i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma(z)\mathbf{A} + b'\Sigma(z)\omega\mathbf{A} + \frac{1}{2}\beta''\omega^2\mathbf{A} = 0. \tag{20}$$

The solution to (20) may be written

$$\mathbf{A}(z + \Delta z, \omega) = \exp[(i/2)\omega^2\beta''\Delta z]\mathsf{M}(\omega)\mathbf{A}(z,\omega) \tag{21}$$

where the transfer matrix M [14] satisfies the ordinary differential equation

$$i\frac{\partial \mathsf{M}}{\partial z} + (b + b'\omega)\Sigma(z)\mathsf{M} = 0 \tag{22}$$

with the initial condition $\mathsf{M}(z = 0, \omega) = \mathsf{I}$, where I is the

2×2 identity matrix. To calculate $\mathsf{M}(\omega)$, we subdivided each interval Δz into smaller subintervals δz, and in each subinterval we set

$$\mathsf{m}_j(\omega) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \tag{23}$$

where

$$m_{11} = m_{22}^* = \cos[(b + b'\omega)\delta z] + i\cos(2\alpha_j)\sin[(b + b'\omega)\delta z] \tag{24}$$

$$m_{12} = m_{21} = i\sin(2\alpha_j)\sin[(b + b'\omega)\delta z] \tag{25}$$

and the α_j are constant in each subinterval. The total transfer matrix is now expressed at

$$\mathsf{M}(\omega) = \prod_{j=0}^{N} \mathsf{m}_j(\omega). \tag{26}$$

The transfer matrix includes the rapid motion over the Poincaré sphere that occurs at the central frequency, but the variation as a function of frequency will be quite small as long as the interval Δz is small compared to the PMD length scale. One way to reach this conclusion is to note, using (5), (9), and (21), that

$$\bar{\mathbf{\Psi}}(z + \Delta z, \omega) = \mathsf{T}^{-1}(z + \Delta z)\mathsf{R}^{-1}(z + \Delta z) \times \{\mathsf{M}(\omega)\exp[(i/2)\omega^2\beta''\Delta z]\}\mathsf{R}(z)\mathsf{T}(z)\bar{\mathbf{\Psi}}(z, \omega). \tag{27}$$

Since T and R are independent of ω, and the change in $\bar{\mathbf{\Psi}}$ is small over the interval Δz, $\mathsf{M}(\omega)$ must vary slowly as a function of ω. Of course, we could directly calculate $\bar{\mathsf{M}}(\omega) = \mathsf{T}^{-1}\mathsf{R}^{-1}\mathsf{M}\mathsf{R}\mathsf{T}$, but this calculation is, in fact, slightly more elaborate because $\bar{\sigma}$ depends on α_z which has δ-function jumps.

The transfer matrix $\mathsf{M}(\omega)$ can be used to determine the evolution of $\mathbf{A}(z, \omega)$ according to (21), or equivalently, we can use $\bar{\mathsf{M}}(\omega)$ to determine $\bar{\mathbf{\Psi}}(z, \omega)$. If $\delta z \approx \Delta z/500$, as was typically the case in the examples that we chose, and we calculated $\mathsf{M}(\omega)$ at each of the frequencies that we used in the fast Fourier transform, then our procedure would still be slow, although faster than direct integration of (1) using the split step method by more than an order of magnitude. However, the variation of $\mathsf{M}(\omega)$ as a function of frequency is very slow on the length scale given by Δz, and it is only necessary to evaluate $\mathsf{M}(\omega)$ at a few frequencies (typically 16) and to interpolate between them with a second-order Lagrange interpolation polynomial.

To demonstrate the slow variation of $\mathsf{M}(\omega)$ with ω, we consider an example in which $\Delta z = 1$ km, $L_{\rm corr} = 100$ m, $\Lambda_{\rm beat} = 50$ m, $D_{\rm PMD} = 3$ ps/(km)$^{1/2}$, and $\delta z = \Lambda_{\rm beat}/30 = 1.67$ m, where $L_{\rm corr}$ is the fiber correlation length [10], [11], $\Lambda_{\rm beat}$ is the beat length, and $D_{\rm PMD}$ is the polarization mode dispersion coefficient. We recall that the parameters b, b', and $\langle(\Delta\alpha)^2\rangle^{1/2}$ where $\langle(\Delta\alpha)^2\rangle^{1/2}$ is the standard deviation of the change in α that occurs at every intervals δz_j, are given by

$$b = \frac{\pi}{\Lambda_{\rm beat}}, \quad b' = \frac{D_{\rm PMD}}{2(2L_{\rm corr})^{1/2}}$$
$$\langle(\Delta\alpha)^2\rangle^{1/2} = \left(\frac{\delta z}{2L_{\rm corr}}\right)^{1/2}. \tag{28}$$

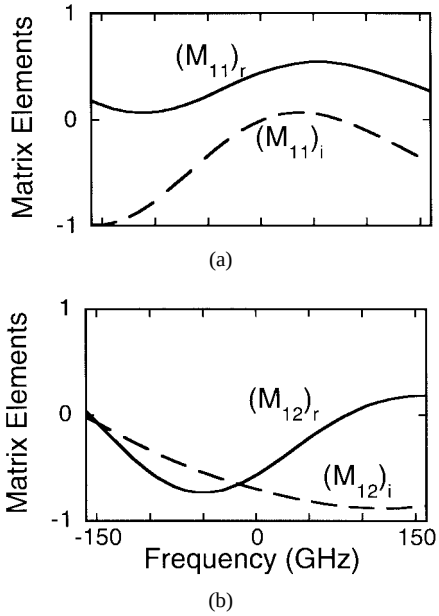

Fig. 1. Frequency dependence of the elements of the M-matrix for the CNLS; (a) real and imaginary parts of M_{11} and (b) real and imaginary parts of M_{12}.

In Fig. 1(a) and (b), we show $[\mathsf{M}_{11}(\omega)]_r$ and $[\mathsf{M}_{11}(\omega)]_i$, the real and imaginary parts of $\mathsf{M}_{11}(\omega)$, and $[\mathsf{M}_{12}(\omega)]_r$ and $[\mathsf{M}_{12}(\omega)]_i$, the real and imaginary parts of $\mathsf{M}_{12}(\omega)$. The smooth, slow oscillations might seem surprising at first, considering that the α_j fluctuate wildly over the 600 short fiber segments that makes up Δz. However, we note that $b'\omega\Delta z \approx 2\pi$ when $f = \omega/2\pi = 320$ GHz, which is larger than the frequency range of interest even though the PMD is large compared to standard fiber parameters.

Dealing with the nonlinear terms in the time domain was quite straightforward. The Manakov term is easily integrable, and we used the expression

$$\bar{\mathbf{\Psi}}(z + \Delta z, t) = \exp\left(in_2k_0\frac{8}{9}|\bar{\mathbf{\Psi}}|^2\right)\bar{\mathbf{\Psi}}(z, t) \tag{29}$$

to calculate its effect. To calculate the effect of the nonlinear PMD terms, we integrated over the rapidly varying coefficients and then did a one-step Euler iteration of $\bar{\mathbf{\Psi}}(z, t)$.

IV. Applications

Here, we will describe the application of the Manakov-PMD equation to studies of both NRZ and soliton signal propagation. We emphasize that since the Manakov-PMD is the basic equation that describes pulse propagation in optical fibers, taking into account the Kerr effect, chromatic dispersion, and polarization mode dispersion, it is equally applicable to both propagation formats. We will show that solving the Manakov-

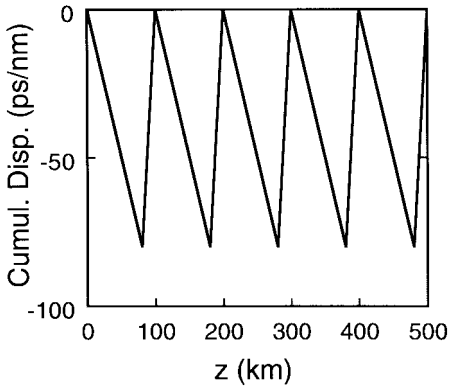

Fig. 2. Dispersion map showing the cumulative dispersion as a function of the length coordinate z.

PMD equation using the numerical scheme described in the last section leads to precisely the same result as is obtained by solving the coupled nonlinear Schrödinger equation on a short length scale, although it is computationally about three orders of magnitude faster in the cases that we considered. Finally, we will discuss the nonlinear PMD term, and we will show that it is negligible in communication fibers using present-day parameters.

A. NRZ Signals

In our studies of NRZ signals we selected the following parameters in addition to those that we listed in the previous section: the wavelength, $\lambda = 1.55\ \mu$m, the effective mode area, $A_{\text{eff}} = 52\ \mu\text{m}^2$, the Kerr coefficient, $n_2 = 2.6 \times 10^{-20}\ \text{m}^2/\text{W}$, an FFT with $2^{10} = 1024$ grid points, and a total fiber length, $L = 500$ km.

A very important point is the choice of a dispersion map. Since at zero dispersion the frequency spectrum of the signal broadens uncontrollably [15], it is essential to have some dispersion in the system. Using a fixed amount of dispersion is undesirable since it causes linear pulse distortion which, on the one hand, is of little theoretical interest and on the other is undesirable in any realistic communication systems. For this reason, we use here a dispersion map consisting of a periodic sequence of normal, $D < 0$, and anomalous, $D > 0$, fiber sections. The period length is 100 km, consisting of 80 km sections with normal dispersion, $D = -1$ ps/(nm $\cdot$ km), and of 20 km compensation sections with $D = 4$ ps/(nm $\cdot$ km). Thus, at the end of each period and of course also at the end of the fiber, the cumulative linear dispersion is zero. The shape of the cumulative dispersion map is shown in Fig. 2. In all simulations reported here, the third-order frequency derivative of β (and of course all derivatives of higher orders) were zero.

As the input pulse pattern we used a 16 bit pseudorandom word in NRZ (nonreturn to zero) format with a 0.2 ns pulse width corresponding to a signaling rate of 5 Gb/s. These pulses are generated by passing rectangularly shaped pulses through a low-pass Bessel filter of $B = 8.75$ GHz width, which gives the pulses slightly rounded edges. These optical input pulses are shown in Fig. 3(a) in units of mW. The peak

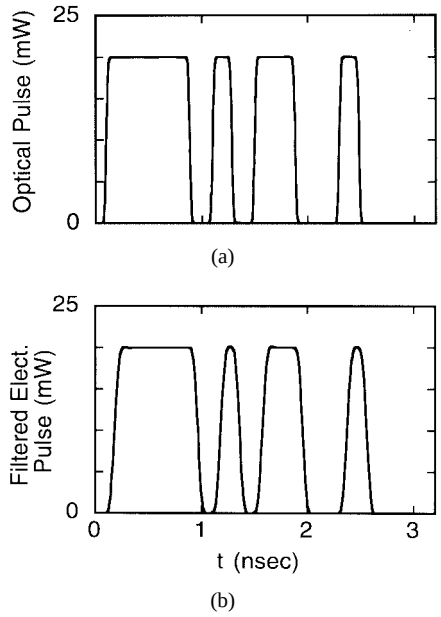

Fig. 3. The train of pseudorandom input pulses consisting of 16 bits. (a) Power of the optical pulses in mW and (b) detected and filtered current in mW units.

pulse power is $P_{\text{peak}} = 20$ mW. This power is somewhat high for communications purposes but was chosen here to produce a sizable nonlinear effect in 500 km of fiber. Subsequently, we present the shape of the received pulses as they appear after square-law detection and filtering with a $B_{\text{elect}} = 5$ GHz "electrical" Bessel filter. This process distorts the pulses very slightly as can be seen in Fig. 3(b) which represents the square-law-detected, filtered input pulse. The quantity plotted in Fig. 3(b) corresponds to a current so that it can go negative, but it is here labeled in units of mW for comparison with the optical pulse, since detected current corresponds to optical power.

All simulations reported here were run without fiber losses, implying that there are also no optical amplifiers in the system and, consequently, no amplifier (ASE) noise.

To compute the behavior of the system, two approaches were used. The first solved the coupled nonlinear Schrödinger equation on the short length scale, carrying out a full split-step calculation after traversing each short dz_j interval. This program is extremely slow and requires 1.5 days of running time for the job at hand on a Sun SPARCstation 10. The second program utilized the ideas explained in the preceding section and solved the Manakov-PMD equation in two minutes, and we ran it both with and without the nonlinear PMD correction. The result of the simulation is shown in Fig. 4(a), which represents the "electrical" output pulses. Both approaches produced exactly the same output. Moreover, there is no discernible difference whether the Manakov routine is run with or without nonlinear PMD.

The shape of the pulses, shown in Fig. 4(a), is strongly dependent on the random number sequence used for computing the random variations of the fiber orientation angle, α. The overshoot seen at the trailing edges of the pulses in Fig. 4(a) can be somewhat larger or smaller. They can even switch sides

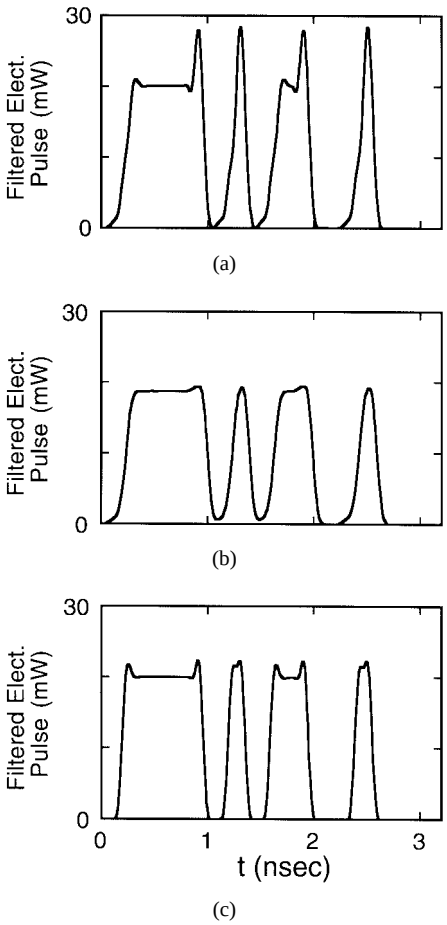

Fig. 4. Detected and filtered current of the output pulses for $L_{\rm corr} = 100$ m and $\lambda_{\rm beat} = 50$ m. (a) This figure shows the solution of the CNLS equation and the Manakov-PMD equation both with and without the nonlinear PMD term. All figures look alike. (b) Same as (a) but for a linear fiber with $n_2 = 0$. (c) Same as (a) but with PMD $= 0$.

and appear at the leading edges of the pulses. These differences correspond physically to the observation that different fibers with different random variations of the birefringence will have somewhat different behaviors. Nonetheless, our two approaches produce exactly the same result for the same assumed variation of the fiber birefringence.

To identify the source of the pulse distortion, we "turn off" the nonlinearity by setting $n_2 = 0$, obtaining Fig. 4(b). Since in a linear system the compensated dispersion map prevents dispersion from distorting the pulses, the remaining slight pulse distortion seen in Fig. 4(b) is caused entirely by linear PMD which is not affected by the dispersion map. However, not all the distortion in Fig. 4(a) is due to the interaction of PMD and the fiber nonlinearity, as can be seen in Fig. 4(c), which was computed with the regular value of n_2 but with $D_{\rm PMD} = 0$. This figure shows that the nonlinearity conspires with dispersion to distort the pulses even though the overall cumulative linear dispersion of the system is zero. The slight asymmetry of the pulses in Fig. 4(c) is attributable to the "electrical" filter which shows this tendency already in Fig. 3(b), but which is enhanced by the substantial broadening of the spectrum at the end of the fiber. A comparison of Fig. 4(a)–(c) makes it clear that the bulk of the pulse distortion is caused by the interaction of the linear PMD with the fiber nonlinearity. However, this interaction is almost entirely due to the interaction of the linear PMD with the Manakov nonlinearity, i.e., the Kerr nonlinearity averaged over the Poincaré sphere, since inclusion of nonlinear PMD leads to no "observable" effect.

B. Solitons

It is instructive to demonstrate by means of a numerical example how fiber birefringence affects the soliton solutions of the CNLS. Averaging over the fiber birefringence, which is explicitly incorporated into the Manakov-PMD equation, manifests itself by the appearance of the factor 8/9 on its left-hand side. The Manakov-PMD equation is exactly equivalent to the CNLS equation, but the averaged contribution of the spatially varying birefringence, which appears on the left-hand side of the equation, has been explicitly separated from the fluctuating contributions, which are on the right-hand side. The nonlinear contribution on the right-hand side may be neglected in most cases of practical interest. However, it is apparent that in the limit that the birefringence strength tends to zero, the right-hand side of the Manakov-PMD equation cannot be negligible and must contribute an amount appropriate to change the 8/9-factor to unity. Thus, fiber birefringence has a noticeable influence on soliton solutions of the Manakov equation where the 8/9-factor appears explicitly in the formula for the soliton pulse duration which, in soliton units, assumes the form

$$\tau_m = \left[\frac{9}{8(|U|^2 + |V|^2)_{\rm peak}}\right]^{1/2}. \tag{30}$$

To demonstrate the influence of fiber birefringence, the solid curve in Fig. 5(a) shows the soliton solution of the Manakov equation, computed without the negligible correction terms, while the dotted curve represents the soliton solution of the CNLS equation for a fiber without birefringence. We have confirmed by numerical solutions of the differential equations that both pulses maintain their shape as they travel along the fiber. Fig. 5(b) shows the soliton solution of the CNLS for a fiber with birefringence. Superimposed on this curve is the Manakov solution of Fig. 5(a). Both curves are almost indistinguishable. However, the height of the CNLS soliton fluctuates along the fiber due to the randomness of the birefringence. The maximum and minimum excursions of the calculated pulse heights differ by approximately 1% of the average pulse height, but Fig. 5(b) demonstrates clearly that the fiber birefringence does indeed modify the shape of the soliton pulse by the amount predicted by the 8/9-factor in the Manakov equation.

C. Nonlinear PMD

Mathematically, the terms that we refer to as nonlinear PMD corresponds to the difference between the Kerr effect terms and their average over the Poincaré sphere. Physically these terms correspond to the rapidly varying fluctuations in the strength of the Kerr effect as the polarization state

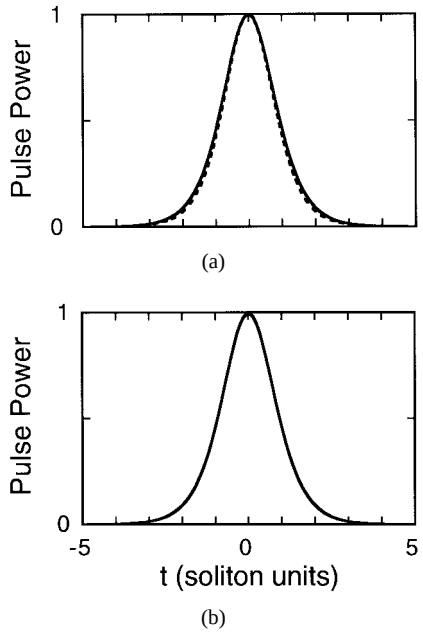

Fig. 5. (a) The solid curve is the solution of the Manakov equation, i.e., the Manakov-PMD equation without the PMD terms on its right-hand side, while the dotted curve is the soliton solution of the CNLS equation for a fiber without birefringence. (b) Soliton solution of the CNLS equation for a fiber with birefringence. Superimposed on this curve is the Manakov solution of (a). The height of the CNLS soliton fluctuates by about 1% due to the randomness of the fiber birefringence.

changes. The size of these fluctuations does not depend on the birefringence strength; they are about the same size as the Kerr effect itself. However, their duration depends on the rate at which the polarization state visits the different points on the Poincaré sphere. If this rate is rapid—as is the case in standard communication fibers with standard signal parameters—then the fluctuations rapidly change sign and their net effect on the signal evolution is nearly zero.

A significant contribution from nonlinear PMD can only occur when mixing on the Poincaré sphere is poor, i.e., this mixing occurs on the same length scale or even a longer length scale than is associated with the Kerr effect, chromatic dispersion, and polarization mode dispersion. From the studies in [10] and [11], one concludes that there are two limits in which this situation can occur. The first and most obvious is when $L_{\rm corr}$, the fiber correlation length, becomes comparable to the nonlinear and dispersive scale lengths. Since the mixing length on the Poincaré sphere must be longer than $L_{\rm corr}$, it is apparent that mixing will be poor. This limit, which applies to subpicosecond pulses, does not occur in present-day communication systems. In this limit, there is no advantage to solving the Manakov-PMD equation rather than directly solving the coupled nonlinear Schrödinger equation since there is no short length scale any more.

The second limit is a little less obvious. When $L_{\rm corr} \ll \Lambda_{\rm beat}$, the mixing length becomes proportional to $\Lambda_{\rm beat}^2/L_{\rm corr}$, and when the mixing length approaches the nonlinear and dispersive scale lengths, mixing will be poor [10], [11]. In this limit, which corresponds to decreasing the magnitude of the birefringence, the birefringence becomes sufficiently small that the polarization state is nearly unaffected by the randomly varying birefringence and, in a fixed frame, the electric field remains correlated over long lengths. This increased mixing length only affects the nonlinear PMD; the strength of the linear PMD is proportional to $L_{\rm corr}$. It has been proposed to significantly decrease the linear PMD by rapidly shifting the orientation of the birefringence axes in a fiber as it is drawn, which amounts to reducing $L_{\rm corr}$ while keeping $\Lambda_{\rm beat}$ fixed [11]. This approach raises the specter that nonlinear PMD might become important in these fibers, but that is not the case. An analytical calculation of the mixing length in this limit shows that it equals $\Lambda_{\rm beat}^2/(12\pi^2 L_{\rm corr})$ [11], and the factor of $1/(12\pi^2)$ is so small that it is difficult to achieve a significant mixing length even under extreme circumstances.

To illustrate this point, we carried out simulations with the same parameter set as in Section IV-A, except that we changed $L_{\rm corr}$ from 100 to 10 m and $\Lambda_{\rm beat}$ from 50 m to 10 km. Obviously, a beat length of 10 km is far larger than can be presently achieved in practice. Nonetheless, the mixing length is only 100 km in this case, which while approaching the nonlinear and dispersive scale lengths is not larger than them. Even under these fairly extreme conditions, we found no significant effect from the nonlinear PMD terms. To observe the effect of the nonlinear PMD terms, we had to further lower $L_{\rm corr}$ from 10 m to 1 m. In this case, the mixing length equals 1000 km which is larger than the nonlinear and dispersive scale lengths. (Note that in this case, we set $dz_j = L_{\rm corr}/30$ instead of $dz_j = \Lambda_{\rm beat}/30$.)

The pulse shape computed for these conditions with the CNLS as well as with the Manakov-PMD equation, including the nonlinear correction term, is shown in Fig. 6(a). The figures computed with the two programs are indistinguishable. Fig. 6(b) shows what happens when we neglect the nonlinear PMD. The pulse shape in Fig. 6(b) is distinctly different from that of Fig. 6(a). This example clearly shows that the nonlinear correction term in (12) does make a contribution if the inequality $L_{\rm corr} \ll \Lambda_{\rm beat}$ is sufficiently well satisfied. However, this example represents a highly extreme case that is unlikely to be found in practice so that the nonlinear correction term can safely be neglected for all communications fibers in use today.

V. Comparison with the Coarse-Step Method

From a practical standpoint, it is not feasible to integrate the coupled nonlinear Schrödinger equation on the small length scale that is required to mimic the rapidly changing birefringence orientation. In the absence of the approach based on the Manakov-PMD equation that we have described here, an *ad hoc* approach that was first introduced in [8] and [9] has become widely used. We refer to it here as the coarse-step method, and we will analyze it using the theoretical tools that were developed in Section II. We will show that the coarse-step method as it is presently used deals with the nonlinear PMD inaccurately but that this inaccuracy does not matter when the contribution of the nonlinear PMD is negligible as is the case in present-day communication systems. We will show that it deals with the linear PMD accurately but to do so the magnitude of the birefringence must be artificially lowered.

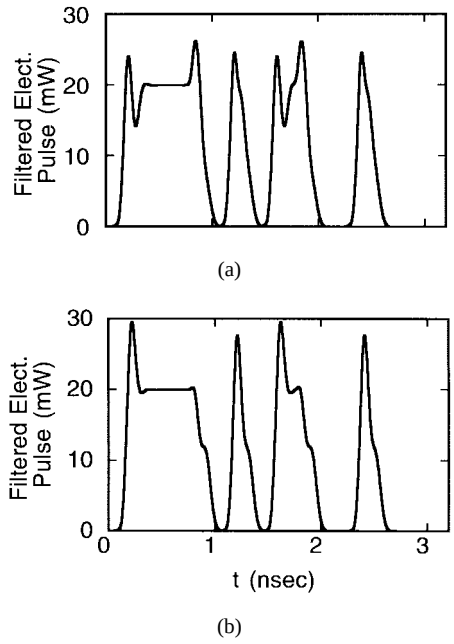

Fig. 6. Detected and filtered current of the output pulses for $L_{\rm corr} = 10$ m and $\lambda_{\rm beat} = 10\,000$ m. (a) This figure shows the solution of the CNLS equation and the Manakov-PMD equation with nonlinear PMD. (b) This figure shows the solution of the Manakov-PMD equation without nonlinear PMD.

We will also show how the method can in principle be modified to deal accurately with nonlinear PMD. Thus our results serve to validate the use of the coarse-step method in simulations of current communications systems.

Our starting point is the coupled nonlinear Schrödinger equation transformed into the frame of the local axes of birefringence. Substituting $\mathbf{\Psi}$, given by (5) into (1), we obtain

$$i\frac{\partial \mathbf{\Psi}}{\partial z} + \tilde{\Sigma}\mathbf{\Psi} + ib'\sigma_3\frac{\partial \mathbf{\Psi}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{\Psi}}{\partial t^2} + n_2 k_0\left[\frac{5}{6}|\mathbf{\Psi}|^2\mathbf{\Psi} + \frac{1}{6}(\mathbf{\Psi}^\dagger\sigma_3\mathbf{\Psi})\sigma_3\mathbf{\Psi} + \frac{1}{3}\mathbf{N}\right] = 0 \quad (31)$$

which may also be written

$$i\frac{\partial \mathbf{\Psi}}{\partial z} + \tilde{\Sigma}\mathbf{\Psi} + ib'\sigma_3\frac{\partial \mathbf{\Psi}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \mathbf{\Psi}}{\partial t^2} + n_2 k_0\left[|\mathbf{\Psi}|^2\mathbf{\Psi} - \frac{1}{3}(\mathbf{\Psi}^\dagger\sigma_2\mathbf{\Psi})\sigma_2\mathbf{\Psi}\right] = 0 \quad (32)$$

where $\mathbf{N} = (\mathbf{\Psi}_1^*\mathbf{\Psi}_2^2, \mathbf{\Psi}_1^2\mathbf{\Psi}_2^*)^t$. In the case where the birefringence axes are fixed, i.e., α_z in $\tilde{\Sigma}$ equals 0, it was long ago observed [3], [4] that the term $\mathbf{N}$ is rapidly varying and could be dropped from (31) with no loss of accuracy in the parameter regime that is used in communication fibers. Equation (31) then reduces to

$$i\frac{\partial \Psi}{\partial z} + ib'\sigma_3\frac{\partial \mathbf{\Psi}}{\partial t} - \frac{1}{2}\beta''\frac{\partial^2 \Psi}{\partial t^2} + n_2 k_0\left[\frac{5}{6}|\mathbf{\Psi}|^2\mathbf{\Psi} + \frac{1}{6}(\mathbf{\Psi}^\dagger\sigma_3\mathbf{\Psi})\sigma_3\mathbf{\Psi}\right] = 0 \quad (33)$$

where we have transformed from $\mathbf{\Psi}$ to $\mathbf{\Psi}' = \exp(-i\sigma_3 z)\mathbf{\Psi}$ and then removed the primes. In the coarse-step method used in [8] and [9], one simply solves (33) on a length scale that is somewhat longer than the fiber correlation length—100 m was used in [8] and 1 km in [9]—but that is still short compared to the nonlinear and dispersive scale lengths.[1] Since the random variations of the birefringence are no longer present in (33), the scattering of the field on the Poincaré sphere that they induce must be put back. In both [8] and [9], the authors do that by scattering the field on the Poincaré sphere at the end of each step, setting $\mathbf{\Psi}' = \mathsf{S}\mathbf{\Psi}$. In [8], the authors used the scattering matrix

$$\mathsf{S} = \begin{pmatrix} \cos\alpha & \sin\alpha\exp(i\phi) \\ -\sin\alpha\exp(-i\phi) & \cos\alpha \end{pmatrix} \quad (34)$$

while in [9], the authors used

$$\mathsf{S} = \begin{pmatrix} \cos\alpha\exp(i\phi) & \sin\alpha\exp(i\phi) \\ -\sin\alpha & \cos\alpha \end{pmatrix} \quad (35)$$

where $\cos 2\alpha$ and ϕ are randomly chosen from uniform distributions in the first case and α and ϕ are randomly chosen from uniform distributions in the second case. In neither case does this matrix introduce a uniform scattering on the Poincaré sphere. To achieve that, one must use, for example, the three Euler angles [16]. However, concatenating several of these matrices together does lead to rapid uniform mixing on the Poincaré sphere, as we will show. This point is significant because the Manakov-PMD equation is obtained by removing the rapid motion that a linear CW signal undergoes on the Poincaré sphere. In this case, a linear, CW signal is unaffected by (33) and the transfer matrix T just consists of a series of S-matrices multiplied together.

To demonstrate the uniform mixing on the Poincaré sphere and to determine the mixing rate, we define the three Stokes parameters $X = \mathbf{\Psi}^\dagger\sigma_3\mathbf{\Psi}$, $Y = \mathbf{\Psi}^\dagger\sigma_1\mathbf{\Psi}$, and $Z = \mathbf{\Psi}^\dagger\sigma_2\mathbf{\Psi}$ which characterize the coordinates on the Poincaré sphere. For comparison, we note that X, Y, and Z correspond respectively to $\tilde{\mathbf{S}}_1$, $\tilde{\mathbf{S}}_2$, and $-\tilde{\mathbf{S}}_3$ in [11]. (The minus sign appears in the representation used in [11] for $\tilde{\mathbf{S}}_3$ is to be consistent with the definition used in [7] of the Stokes parameters. Here we return to the definition used in [9] and [13].) For the scattering matrix given in (34), one finds that the Stokes parameters of a linear, CW signal evolve according to the formula shown in (36) at the bottom of the next page and for (35), this result becomes

$$\begin{pmatrix} X_{n+1} \\ Y_{n+1} \\ Z_{n+1} \end{pmatrix} = \begin{pmatrix} \cos 2\alpha_n & \sin 2\alpha_n & 0 \\ -\sin 2\alpha_n\cos\phi_n & \cos 2\alpha_n\cos\phi_n & \sin\phi_n \\ \sin 2\alpha_n\sin\phi_n & -\cos 2\alpha_n\sin\phi_n & \cos\phi_n \end{pmatrix} \times \begin{pmatrix} X_n \\ Y_n \\ Z_n \end{pmatrix} \quad (37)$$

where the subscript n indicates the step number. Noting that the $\cos 2\alpha_n$ and ϕ_n are chosen randomly in each step for the

[1] In the description of the numerical algorithm in both [8] and [9], it was not explicitly stated that $\mathbf{N}$ is dropped. We are grateful to Dr. Evangelides for describing to us the numerical algorithm that was used in [9].

first case, one finds from (36) that

$$\begin{pmatrix}\langle {X_{n+1}}^2\rangle\\ \langle {Y_{n+1}}^2\rangle\\ \langle {Z_{n+1}}^2\rangle\end{pmatrix}=\begin{pmatrix}1/3 & 1/3 & 1/3\\ 1/3 & 1/2 & 1/6\\ 1/3 & 1/6 & 1/2\end{pmatrix}\begin{pmatrix}\langle X_n^2\rangle\\ \langle Y_n^2\rangle\\ \langle Z_n^2\rangle\end{pmatrix} \tag{38}$$

where the angular brackets indicate an average over all possible values of $\cos 2\alpha_n$ and ϕ_n. We finally conclude

$$\begin{pmatrix}\langle X_n^2\rangle\\ \langle Y_n^2\rangle\\ \langle Z_n^2\rangle\end{pmatrix}=\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}+a\frac{1}{3^n}\begin{pmatrix}0\\1\\-1\end{pmatrix}+b\delta_{n0}\begin{pmatrix}2\\-1\\-1\end{pmatrix} \tag{39}$$

where $\delta_{n0}=1$ if $n=0$ and 0 otherwise, and the values of a and b depends on the initial conditions. For example, if $X_0=1$, $Y_0=0$, and $Z_0=0$, we find $a=0$ and $b=1/3$. The convergence to a uniform distribution function is very rapid. Indeed, if we define the diffusion length as the length over which the initial conditions fall to within $1/e$ of their uniform value of $(1/3,1/3,1/3)^t$, then we find that this occurs within two steps, regardless of the initial conditions. A similar analysis for (37) yields

$$\begin{pmatrix}\langle X_n^2\rangle\\ \langle Y_n^2\rangle\\ \langle Z_n^2\rangle\end{pmatrix}=\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}+a\frac{1}{4^n}\begin{pmatrix}2\\-1\\-1\end{pmatrix}+b\delta_{n0}\begin{pmatrix}1\\-1\\0\end{pmatrix}. \tag{40}$$

The convergence in this case is even more rapid than in the previous case.

We now determine the polarization mode dispersion in the coarse-step method. Focusing on the linear part of (33), ignoring chromatic dispersion, and the Kerr effect, we find

$$\begin{aligned}\tilde{\mathbf{\Psi}}(z=N\Delta,\omega)&=\prod_{n=1}^{N}(\mathsf{S}_n\mathsf{S}_\omega)\tilde{\mathbf{\Psi}}(z=0,\omega)\\ &=\mathsf{S}_n\mathsf{S}_\omega\mathsf{S}_{n-1}\mathsf{S}_\omega\mathsf{S}_{n-2}\mathsf{S}_\omega\\ &\quad\cdots\mathsf{S}_1\mathsf{S}_\omega\tilde{\mathbf{\Psi}}(z=0,\omega)\end{aligned} \tag{41}$$

where $\tilde{\mathbf{\Psi}}(z,\omega)$ is the Fourier transform of $\mathbf{\Psi}(z,t)$, the S_n are the scattering matrices given by (34) or (35), and

$$\mathsf{S}_\omega=\exp(ib'\sigma_3\omega\Delta)=\begin{pmatrix}\exp(ib'\omega\Delta) & 0\\ 0 & \exp(-ib'\omega\Delta)\end{pmatrix}. \tag{42}$$

Taking the derivative of (41) with respect to ω and writing $\tilde{\mathbf{\Psi}}'(z=N\Delta,\omega=0)=\mathsf{F}\tilde{\mathbf{\Psi}}(z=N\Delta,\omega=0)$ we find

$$\mathsf{F}=ib'\Delta\sum_{m=1}^{N}\mathsf{V}_m\sigma_3\mathsf{V}_m^\dagger \tag{43}$$

where $\mathsf{V}_m=\prod_{n=m}^{N}(\mathsf{S}_n\mathsf{S}_\omega)=\mathsf{S}_N\mathsf{S}_\omega\mathsf{S}_{N-1}\mathsf{S}_\omega\mathsf{S}_{N-2}\mathsf{S}_\omega\cdots\mathsf{S}_m\mathsf{S}_\omega$. Physically, the matrix $\mathsf{V}_m^\dagger$ corresponds to the backward transformation from $z=N\Delta$ to $z=m\Delta$. By analogy with (16), we write

$$\mathsf{V}_m\sigma_3\mathsf{V}_m^\dagger=\begin{pmatrix}x_{1,m} & x_{2,m}-ix_{3,m}\\ x_{2,m}+ix_{3,m} & -x_{1,m}\end{pmatrix} \tag{44}$$

where letting $\mathbf{R}_j=(x_j,y_j,z_j)$, $\mathbf{R}_1=(1,0,0)$, $\mathbf{R}_2=(0,1,0)$, and $\mathbf{R}_3=(0,0,1)$ at $z=N\Delta$.

Noting that $\langle {\tau_D}^2\rangle=-4\langle\det(\mathsf{F})\rangle$ [13], we substitute (44) into (43) and use the results, which may be derived directly from (36) and (37), that the x_j are uncorrelated from step to step to conclude

$$\langle\tau_D^2\rangle=-4(b')^2\Delta^2N=4(b')^2\Delta Z. \tag{45}$$

By contrast, one finds in the large-z limit from [11, eq. (22)]

$$\langle\tau_D^2\rangle=8(b')^2L_{\rm corr}Z. \tag{46}$$

Thus, the polarization mode dispersion will be too large by a factor of $(\Delta/2L_{\rm corr})^{1/2}$.

However, if one lowers b' and hence the birefringence, by a factor of $(2L_{\rm corr}/\Delta)^{1/2}$, then the polarization mode dispersion returns to its original value in a statistical sense. If, for example, $L_{\rm corr}=30$ m and $\Delta=1$ km, one must lower the birefringence by a factor of four. The birefringence is often not even known within a factor of this magnitude!

We now consider the nonlinear terms. The first point to note is that $\mathbf{N}$, given in (31), does not contribute to the average over the Poincaré sphere which is $(8/9)n_2k_0|\mathbf{\Psi}|^2\mathbf{\Psi}$ regardless of whether the term is kept or dropped. By contrast, the effect on the nonlinear PMD is profound. We previously found that the nonlinear contribution involves the z_j, as can be seen by reference to (12)–(18). In this case, we find that the nonlinear contribution to $\mathbf{\Psi}$ is given by

$$i\frac{\partial\mathbf{\Psi}}{\partial z}+n_2k_0\left[\frac{5}{6}|\mathbf{\Psi}|^2\mathbf{\Psi}+\frac{1}{6}(\mathbf{\Psi}^\dagger\bar{\sigma}\mathbf{\Psi})\bar{\sigma}\mathbf{\Psi}\right]=0 \tag{47}$$

where $\bar{\sigma}$, given by (16) involves the x_j, rather than the z_j. We can write out an expression that is analogous to (17) but there is little point. The basic result is that the nonlinear PMD is inaccurately described by the coarse-step method. It was already noted in [11] that the coarse-step method tends to exaggerate the strength of the PMD coefficients by a factor of about $(\Delta/L_{\rm mix})$ where $L_{\rm mix}$ is the appropriate mixing length. Even more serious, however, is that the coefficients for nonlinear PMD are incorrect in the coarse-step method. In particular, because $L_{\rm mix}\propto L_{\rm corr}$ for the x_j regardless of $\Lambda_{\rm beat}$, while $L_{\rm mix}\propto {\Lambda_{\rm beat}}^2/L_{\rm corr}$ for the z_j when $\Lambda_{\rm beat}\gg L_{\rm corr}$, the coarse-step method will not show the expected rise in nonlinear PMD in this limit. However, the poor treatment of nonlinear PMD in the coarse-step method has little practical significance for simulations of communications fibers to date in which both the actual nonlinear PMD and the nonlinear PMD generated by the coarse-step method are both negligible.

Our results suggest a modest improvement in the coarse-step method. One could use an Euler angle transformation at the

$$\begin{pmatrix}X_{n+1}\\ Y_{n+1}\\ Z_{n+1}\end{pmatrix}=\begin{pmatrix}\cos 2\alpha_n & \sin 2\alpha_n\cos\phi_n & -\sin 2\alpha_n\sin\phi_n\\ -\sin 2\alpha_n\cos\phi_n & \cos^2\alpha_n-\sin^2\alpha_n\cos 2\phi_n & \sin^2\alpha_n\sin 2\phi_n\\ \sin 2\alpha_n\sin\phi_n & \sin^2\alpha_n\sin 2\phi_n & \cos^2\alpha_n+\sin^2\alpha_n\cos 2\phi_n\end{pmatrix}\times\begin{pmatrix}X_n\\ Y_n\\ Z_n\end{pmatrix} \tag{36}$$

end of each step which would guarantee complete mixing on the Poincaré sphere. Should it ever be of practical importance, our results also suggest an approach that will yield correct results in a statistical sense for the nonlinear PMD. Keeping all the nonlinear PMD terms, i.e., not dropping $\mathbf{N}$, one multiplies the nonlinear PMD terms by a factor $(\Delta/L_{\rm mix})^{1/2}$, where $L_{\rm mix}$ is given in [11]. Just as lowering the coefficient of the linear PMD restores the correct statistical behavior of the linear term, lowering the coefficient of the nonlinear PMD terms restores the correct statistical behavior of the nonlinear terms. Of course, one must leave the coefficient of the Manakov terms $(8/9)n_0k_2|\bar{\mathbf{\Psi}}|^2\bar{\mathbf{\Psi}}$ unchanged.

In summary, our results validate the use of the coarse-step method as it is currently practised. Moreover, we have carried out simulations of the cases presented in Section IV-A using it, and we have verified that it appears in a statistical sense to yield similar results. Here, one of the principal practical difficulties with the coarse-step method asserts itself. We were able to compare the Manakov-PMD equation solution directly to solutions of the coupled nonlinear Schrödinger equation solved on a short length scale, which is not possible to do with the coarse-step method. We note that in our simulations the coarse-step method is somewhat faster than our simulations based on the Manakov-PMD method, but the two approaches have comparable speeds.

VI. Conclusion

In this paper, we reported our investigations of the Manakov-PMD equation which can be derived from the coupled nonlinear Schrödinger equation and is fully equivalent to it. Its use is appropriate when modeling optical fibers whose birefringence orientation changes rapidly by comparison with the typical length scales for chromatic dispersion, polarization mode dispersion, and nonlinearity. We first derived the Manakov-PMD equation from the coupled nonlinear Schrödinger equation, and we found that it consists of a set of terms that correspond physically to linear PMD, chromatic dispersion, the Kerr nonlinearity, and nonlinear PMD. We interpreted nonlinear PMD as the residue of incomplete mixing of the Kerr nonlinearity over the Poincaré sphere. We next developed a numerical algorithm for solving the Manakov-PMD equation and applied it to both NRZ and soliton signals. We also examined the nonlinear PMD terms and showed that they are negligible even under fairly extreme circumstances. Thus these terms can be safely dropped in almost all cases, simplifying the numerical algorithm. We next carried out an analysis of the widely used coarse-step method, and we showed that it will yield reasonable results in the parameter regime that is currently used in communication systems. At the same time, we note that the additional computation load required to solve the Manakov-PMD equation is slight, and it allows us to both accurately model particular representations of the birefringence variation along the fiber and to determine the importance of the different physical effects that contribute to the Manakov-PMD equation in a physically transparent way.

References

[1] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear pulses in dispersive dielectric fibers—I: Anomalous dispersion," *Appl. Phys. Lett.*, vol. 23, pp. 142–1446, 1973.

[2] G. P. Agrawal, *Nonlinear Fiber Optics*. New York: Academic, 1989.

[3] C. R. Menyuk, "Nonlinear pulse propagation in birefringence optical fibers," *IEEE J. Quantum Electron.*, vol. 23, pp. 174–176, 1987; see also, C. R. Menyuk, "Pulse propagation in an elliptically birefringent Kerr medium," *IEEE J. Quantum Electron.*, vol. 25, pp. 2674–2682, 1989.

[4] ———, "Stability of solitons in birefringent optical fibers. I: Equal propagation amplitudes," *Opt. Lett.*, vol. 12, pp. 614–616, 1987; see also, C. R. Menyuk, "Stability of solitons in birefringent optical fibers. II: Arbitrary amplitudes," *J. Opt. Soc. Amer. B.*, vol. 5, pp. 392–402, 1988.

[5] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, pp. 1029–1030, 1986.

[6] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 13, pp. 687–689, 1988.

[7] C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," *Opt. Lett.*, vol. 14, pp. 523–525, 1989.

[8] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Opt. Lett.*, vol. 16, pp. 1231–1233, 1991.

[9] S. G. Evangelides Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightwave Tech.*, vol. 10, pp. 28–35, 1992.

[10] P. K. A. Wai and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Opt. Lett.*, vol. 19, pp. 1517–1519, 1994; P. K. A. Wai and C. R. Menyuk, "Anisotropic diffusion of the state of polarization in optical fibers with randomly varying birefringence," *Opt. Lett.*, vol. 20, pp. 2490–2492, 1995.

[11] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 14, pp. 148–157, 1996.

[12] D. Gloge, "Weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2252–2258, 1971.

[13] C. R. Menyuk and P. K. A. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," *J. Opt. Soc. Amer. B*, vol. 11, pp. 1288–1296, 1994.

[14] W. A. Shurcliff and S. S. Ballard, *Polarized Light*. Princeton, NJ: Van Nostrand, 1964, chs. 7, 8.

[15] D. Marcuse, "Single-channel operation in very long nonlinear fibers with optical amplifiers at zero dispersion," *J. Lightwave Technol.*, vol. 9, pp. 356–361, 1991.

[16] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Wesley, 1980, pp. 143–158.

Dietrich Marcuse (M'58–F'73) was born in Koenigsberg, East Prussia, Germany, in 1929. He received the degree of Diplom Physiker from the Freie Universitaet Berlin, Germany, in 1954 and the degree of Doktor Ingenieur from the Technische Hochschule, Karlsruhe, Germany, in 1962.

From 1954 to 1957, he worked at the Central Laboratory of Siemens and Halske, Berlin, on transmission line problems and on the development of the circular electric millimeter waveguide. From 1957 to 1994, he was a Member of Technical Staff at AT&T Bell Laboratories, where he worked initially on problems of wave propagation in circular electric millimeter waveguides and on hydrogen cyanide maser. But throughout most of his career, he worked on theoretical aspects of various laser problems and on the theory of wave propagation in optical fibers. From 1966 to 1967, he spent one year on leave of absence from AT&T Bell Laboratories to teach at the Department of Electrical Engineering of the University of Utah, Salt Lake City. Two one-month leaves were spent at the Institute of Advanced Studies of the Australian National University, Canberra, in 1975 and 1990. After retiring from AT&T Bell Laboratories, he became a part-time Visiting Research Professor at the University of Maryland, Baltimore County. He has been a U.S. citizen since 1966. He is the author of four books and 200 technical papers and holds 13 patents.

Dr. Marcuse is a Fellow of the Optical Society of America. He received the Quantum Electronics Award of the Quantum Electronics and Applications Society of the IEEE in 1981 and the Max Born Award of the Optical Society of America in 1989.

C. R. Menyuk (SM'88) was born on March 26, 1954. He received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, in 1976 and the Ph.D. degree from the University of California, Los Angeles, in 1981.

He had worked as a Research Associate at the University of Maryland, College Park, and also worked at the Science Applications International Corporation, McLean, VA. In 1986, he became an Associate Professor in the Department of Electrical Engineering at the University of Maryland, Baltimore County, and was its founding member. In 1993, he became a Professor. For the last few years, his primary research area has been theoretical and computational studies of nonlinear and guided-wave optics. He has written computer codes to model the nonlinear propagation of light in optical fiber which have been used by industrial, governmental, and university research groups throughout the United States.

Dr. Menyuk is a member of the APS, OSA, and SIAM.

P. K. A. Wai (SM'96) was born in Hong Kong on January 22, 1960. He received the B.S. degree from the University of Hong Kong, Hong Kong, in 1981 and M.S. and Ph.D. degrees from the University of Maryland, College Park, in 1985 and 1988, respectively.

In 1988, he joined the Science Applications International Corporation, McLean, VA, as a Research Scientist, where he worked on the Tethered Satellite Project. In 1990, he became a Research Associate in the Department of Electrical Engineering at the University of Maryland, Baltimore County. In 1996, he became an Assistant Professor at the Department of Electronic Engineering of the Hong Kong Polytechnic University, Hong Kong. He was promoted to an Associate Professor in 1997. His research interests include theory of solitons, modeling of fiber lasers, simulations of integrated optical devices, long distance fiber optic communication, and neural networks.