Periodic compensation of polarization mode dispersion

M. Chertkov and I. Gabitov

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

I. Kolokolov

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and Budker Institute, Novovsibirsk 630090, Russia

T. Schäfer

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

Received December 18, 2002; revised manuscript received October 6, 2003; accepted October 14, 2003

Polarization mode dispersion is the effect of signal broadening in a fiber with birefringent disorder. The disorder, frozen into the fiber, is characterized by the so-called vector of birefringence (VB). In a linear medium a pulse broadens as the two principal states of polarization split. It is well-known that, under the action of short-correlated disorder, naturally present in fibers, the dispersion vector (DV), characterizing the split, performs a Brownian random walk. We discuss a strategy of passive (i.e., pulse-independent) control of the DV broadening. The suggestion is to pin (compensate) periodically or quasi-periodically the integral of VB to zero. As a result of the influence of pinning, the probability distribution function of the DV becomes statistically steady in the linear case. Moreover, pinning improves confinement of the pulse in the weakly nonlinear case. The theoretical findings are confirmed by numerical analysis.

OCIS codes: 060.5530, 190.0190, 000.5490.

1. INTRODUCTION

To design high-bit information transmission systems in optical communications, it is necessary to understand various effects that influence the propagation of a signal in an optical fiber. One prominent effect is polarization mode dispersion (PMD), which has been studied in the linear case in a seminal series of papers by Poole and co-authors,[1–4] and, to the present day, PMD has remained one of the major limitations of a transmission system.[5] The effects of PMD on pulse propagation have been studied experimentally,[6,7] and measurement techniques have been analyzed.[8,9] Inspired by experimental observations,[10] Poole and Wagner introduced the concept of principal states,[11] which has been used widely in theory[12,13] and experiments.[14] Extensions of their concept account for polarization-dependent loss[15–17] and nonlinearity.[18] One of the most important problems, however, is how to compensate for PMD. The scope of this paper is to show that periodic compensation (pinning) can help to improve a system's performance. This method works (however, to different extents) in both the linear and the nonlinear regime.

The paper is organized as follows. First, in Section 2, we will introduce the basic equation and the general idea of pinning. Then we review the theory of principal states and its extension to pulses in Section 3. In Subsections 3.A–3.C we will introduce a methodology to characterize the effect of PMD on pulse propagation that is closely linked to such quantities arising naturally in the linear theory, such as the polarization dispersion vector $\mathbf{\Omega}$. We show that the quantities are meaningful in the nonlinear theory as well. Multifrequency statistics of the polarization dispersion vector for natural noise (no pinning) are described in Subsection 3.D. Then, in Section 4, we turn to the analytical description of the pulse statistics in both the linear and the nonlinear case for pinned short-correlated noise. Multifrequency statistics of $\mathbf{\Omega}$ are described in Subsection 4.A, where it is shown, in particular, that the compensation due to pinning in the linear case is ideal, in the sense that the pinning compensates for the growth of the pulse width completely. Subsection 4.B discusses the effect of pinning in the weakly nonlinear case. Section 5 is devoted to numerical analysis of PMD and comparison of the effects for natural and pinned cases. We demonstrate numerically the effectiveness of the pinning method. Subsections 5.A and 5.B are devoted to analysis of the linear and nonlinear cases, respectively. Finally, Section 6 is reserved for brief conclusions.

2. BASIC MODEL OF POLARIZATION MODE DISPERSION

An optical fiber is not circular in cross section but rather elliptic. As a result of this[19] and other influences such as twist and stress,[20,21] light changes its polarization while propagating through the fiber. This phenomenon is at the core of our description, and we start the paper with an introduction into the linear and nonlinear physics of the one-dimensional trapped propagation of light through a monomode optical fiber with birefringent, but isotropic (on average), disorder.

The electric field $\mathbf{E}$ is tangential to the major propagation direction. It decomposes into two complex components as $\mathbf{E} = \mathcal{E}\exp(ik_0 z - i\omega_0 t) + \mathcal{E}^* \exp(-ik_0 z + i\omega_0 t)$, where z is the coordinate along the fiber. ω_0 and k_0 are the carrier frequency and the wave vector of the electric signal, respectively. In the monomode regime the coarse-grained (envelope) description of the signal propagation is given by the complex, two-component field Ψ, $\mathcal{E} \equiv \Sigma_{\alpha=1,2}\Psi_\alpha \mathbf{e}_\alpha$. Here $\mathbf{e}_{1,2}$ are unit vectors, orthogonal to each other and to the waveguide direction. The envelope approximation equation (derived directly from Maxwell's equations; see Refs. 22 and 23) is as follows:

$$i\partial_z \Psi_\alpha + \Delta^{\alpha\beta}\Psi_\beta + im^{\alpha\beta}\partial_t \Psi_\beta + d(z)\partial_t^2 \Psi_\alpha = -\frac{\delta\langle W\rangle}{\delta\Psi_\alpha^*}. \tag{1}$$

This is a partial differential equation of second order in t, where t is the retarded time measured from the reference frame moving along the fiber with the mean group velocity of the signal. The matrices $\hat{m}$ and $\hat{\Delta}$ characterize the birefringence and will be discussed in more detail below. Superscripts and subscripts are used synonymously in order to make use of Einstein's summation convention. The dispersion $d(z)$ is a function of z that can usually be assumed to be piecewise constant. The right-hand side (rhs) of Eq. (1) is reserved for the nonlinear Kerr term. We consider both linear and nonlinear problems in the paper. The former means zero for W. In the latter case, W is the Kerr energy. In the isotropic medium, $W = \mathbf{E}^4/6$. (The dimensional coefficient is rescaled to 1/6 by a proper choice of the z units.) $\langle W\rangle$ stands for the average of the energy over the fast carrier frequency and wave-vector oscillations (see Ref. 24 for a similar derivation in a three-dimensional isotropic medium):

$$\begin{aligned}6\langle W\rangle &= \langle[\mathcal{E}^2\exp(2i\kappa_0 z) + 2|\mathcal{E}|^2 + \mathcal{E}^{*2}\exp(-2i\kappa_0 z)]^2\rangle \\ &= 2\mathcal{E}^2\mathcal{E}^{*2} + 4|\mathcal{E}|^4 \\ &= 2(\Psi_1^2 + \Psi_2^2)(\Psi_1^{*2} + \Psi_2^{*2}) + 4(|\Psi_1|^2 + |\Psi_2|^2)^2,\end{aligned} \tag{2}$$

$$\frac{\delta W}{\delta\Psi_\alpha^*} = \frac{2}{3}\Psi_\alpha^*(\Psi_1^2 + \Psi_2^2) + \frac{4}{3}\Psi_\alpha(|\Psi_1|^2 + |\Psi_2|^2). \tag{3}$$

It is easy to check that expressions (2) and (3) are invariant under any rotation of the polarization axes:

$$\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix} \to \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix}. \tag{4}$$

Let us discuss various terms on the left-hand side (lhs) of Eq. (1). The last term describes combined (material and waveguide) dispersion of the fiber. One assumes here that the dispersion is constant along the fiber (see Refs. 25 and 26 for a discussion of the effects associated with variations of the fiber dispersion) and rescales the dispersion coefficient to unity by the proper choice of the temporal units. The matrix of birefringence, $\hat{m}$, describes anisotropy in the group velocity for the two states of polarization. The matrix $\hat{m}$ is traceless as a result of the proper choice of the retarded-time reference frame. One assumes that there is no energy loss, which means that $\hat{m}$ is self-adjoint. The $\hat{m}$ term on the lhs of Eq. (1) is the only source of medium anisotropy that we consider in this paper. Anisotropic (birefringent) corrections to the terms in Eq. (1) other than the $\hat{\Delta}$ and $\hat{m}$ ones are usually less important, i.e., they have a weaker effect on propagation, if the fiber is isotropic on average. As far as the other anisotropic term in Eq. (1), $\hat{\Delta}$, is concerned, we can actually exclude it from consideration. The reason for this simplification is twofold. In some fibers $\hat{\Delta}$ is just small. In the general case (e.g., in the opposite and more realistic limit of large $\hat{\Delta}$, where $|\Delta| \gg |m|/b$ and b is the pulse width), one can change Ψ_α in Eq. (1) to the slow field $\tilde{\Psi}_\alpha = T\exp\{[\int_0^z \hat{\Delta}(z')\mathrm{d}z']^{\alpha\beta}\}\Psi_\beta$, where $T\exp$ signifies a z-ordered exponential, and then one can average over fluctuations of $\hat{\Delta}(z)$. The linear part of the equation for the averaged $\tilde{\Psi}$ is identical to the lhs of Eq. (1), while the nonlinear term depends on details of the statistics of the matrix $\hat{\Delta}$. Some discussion of the Δ-averaging procedure can be found in the literature,[27–29] where the special case of averaging over an almost circular fiber, so that the bit length and the correlation length of the Δ disorder are comparable, was considered. In the part of this paper dealing with nonlinear and weakly nonlinear cases, we adopt another possible scenario of the nonlinear term averaging[30] (different from the one described in Refs. 27–29), where the structure of the averaged nonlinear term remains the same as that given by the rhs of Eq. (1). Note that both the theoretical and the numerical analysis of this paper can be easily extended to account for other formulations of nonlinear terms. This generalization/modification does not change the major statement of the paper that pinning of the $\hat{m}$ term in Eq. (1) suppresses pulse degradation. Thus, in this paper, we concentrate on analyzing Eq. (1) with the $\hat{\Delta}$ term neglected.

The matrix $\hat{m}$, which is parameterized through the Pauli matrices as

$$\hat{m} = \sum_j h^j(z)\hat{\sigma}_j, \tag{5}$$

(where $j = 1, 2, 3$), is zero on average. We call $\mathbf{h}$ the vector of birefringence (VB). The correlation scale of the random $h^j(z)$, naturally occurring because of imperfections in fiber production, is short. Therefore, according to the central limit theorem, $h_j(z)$ at the greater scales (corresponding to the long-haul transmission that we study here) can be treated as a homogeneous Gaussian random process with zero mean. The noise intensity is described by $D_{jk} = \int \mathrm{d}z\langle h^j(z)h^k(z')\rangle$. One assumes that the isotropy of the noise is restored on average, i.e., $D^{jk} = D\delta^{jk}$. In standard fiber optics jargon, D corresponds to the so-called PMD coefficient.

The statistics of the VB are given by the functional measure

$$\mathcal{D}\mathbf{h}(z)\exp\left(-\frac{1}{2D}\int \mathbf{h}^2 \mathrm{d}z\right), \tag{6}$$

which is unambiguously characterized by the pair correlation function

$$\langle h^i(z)h^j(z')\rangle = D\delta^{ij}\delta(z - z'). \tag{7}$$

Here δ^{ij} is the Kronecker symbol, and $\delta(z)$ is the Dirac δ function. Note that h^i is t independent, as it describes anisotropy frozen into the fiber. Note also that averaging in Eqs. (6) and (7) corresponds to either collecting statistics over many realizations (many fibers or separated-in-time states of birefringence of the same fiber) or spatial z averaging over one particular realization. In the latter case the correlation scale of the birefringent disorder is assumed to be much shorter than the averaging scale. Following standard notations,[31] the path-integral formulation for the averaging is introduced into expression (6) for the convenience of notation and further manipulations and derivations. (See, e.g., Ref. 31 for a description and a detailed explanation of the path-integral approach, widely used in statistical physics.)

In parallel with the natural process (7), we will also consider a synthetic one. The synthetic model enforces the pinning to zero of the integral of the VB. It is described by

$$\mathcal{D}\mathbf{h}(z)\exp\left(-\frac{1}{2D}\int \mathbf{h}^2 \mathrm{d}z\right)\prod_n \delta\left(\int_{L(n)}^{L(n+1)} \mathrm{d}z\mathbf{h}\right), \tag{8}$$

where $L(n)$, $n = 1,...,$ is the sequence of the pinning points, $\delta(x)$ stands for the Dirac δ function, and we use standard notation for the path-integral measure. The process remains Gaussian with zero mean. It is unambiguously described by the pair correlation function, which is

$$\langle h^i(z_1)h^j(z_2)\rangle = D\delta^{ij}\left[\delta(z_1 - z_2) - \frac{1}{L(n+1) - L(n)}\right] \tag{9}$$

for z_1 and z_2 belonging to the same fiber span between two pinning points, i.e., $L(n) < z_{1,2} < L(n+1)$, and zero otherwise. The physical meaning of the model described by expressions (8) and (9) is to enforce (by some artificial means that we do not discuss in the paper, assuming simply that such means are experimentally available) return to zero (pinning) for the integral birefringence $\int_0^Z \mathrm{d}z\mathbf{h}(z)$ at the sequence of special points $Z = L(n)$. For the sake of simplicity, we consider the periodic pinning process $L(n+1) - L(n) = l$.

Practically speaking, the synthetic noise brings the integral of the noise back to zero periodically. Consider as an example the stochastic equation $\dot{\Xi} = \xi(z)$, with $\xi(z)$ being white noise and the dot denoting differentiation with respect to z. Then, after a length L, we will have $\langle \Xi(L)^2\rangle \propto L$, the standard result for Brownian motion. For a pinned process $\xi_p(t)$ with $L_{n+1} - L_n = l$, one derives $\dot{\Xi}_p = \xi_p(z)$, and further, directly from Eq. (9), one derives that $\langle \Xi(L)^2\rangle$ does not grow with L. It is intuitively clear that this suppression of the birefringence strength achieved by pinning should lead to reduction of the effect of the PMD on the signal, similarly to how pinning of random dispersion, discussed in Ref. 26, improves the transmission characteristics of a fiber line with a random dispersion coefficient.[26] In this paper we aim to quantify this intuitive statement.

3. LINEAR POLARIZATION MODE DISPERSION

A. Linear Propagation Equation

Here we discuss the linear version of Eq. (1), i.e., with zero rhs; hence

$$i\partial_z\Psi + [i\hat{m}(z)\partial_t + d(z)\partial_t^2]\Psi = 0. \tag{10}$$

The formal solution of the equation is

$$\Psi(t;\, z) = \int_{-\infty}^{+\infty}\mathrm{d}\omega\int_{-\infty}^{+\infty}\frac{\mathrm{d}t'}{2\pi}\exp[-i\omega(t - t') - i\omega^2 R(z)]\hat{W}(z|0;\, \omega)\Psi(t';\, 0), \tag{11}$$

$$\hat{W}(z|0;\, \omega) \equiv T\exp\left[i\omega\int_0^z \mathrm{d}z'\hat{m}(z')\right],$$

$$R(z) = \int_0^z d(z')\mathrm{d}z', \tag{12}$$

where $\hat{W}$ is the z-ordered exponential (i.e., it is an operator that satisfies $\mathrm{d}\hat{W}/\mathrm{d}z = i\omega\hat{m}\hat{W}$). The problem of describing the statistics of Ψ is thus reduced to averaging of some functionals of $\hat{W}$.

B. Polarization and Principal States

In this subsection we consider the dispersion-free case $d(z) = 0$. There are two equivalent possibilities to describe the effects of linear PMD. One way is to look at $\hat{W}$ and to describe properties of this time-ordered exponential in the space of unitary 2×2 matrices, SU(2). The second possibility is to use the fact that there is a homomorphism between SU(2) and SO(3) and to describe PMD on the Poincaré sphere. In this subsection we review the main results of this theory.

Following the first description, we note that $\hat{W}$ is unitary and can be written as

$$\hat{W} = \begin{bmatrix} u_1 & u_2 \\ -u_2^* & u_1^* \end{bmatrix}, \qquad |u_1(z;\, \omega)|^2 + |u_2(z;\, \omega)|^2 = 1. \tag{13}$$

The concept of principal states was introduced by looking at a *constant* input signal in $\Psi(\omega;\, 0) = \Psi_0$.[11] Then the output signal depends on the frequency, and one finds

$$\Psi(z;\, \omega) = \hat{W}(z|0;\, \omega)\Psi_0, \tag{14}$$

$$\frac{\partial\Psi}{\partial z} - i\omega\hat{m}\Psi = 0. \tag{15}$$

This allows derivation of a relation between the frequency derivative of the output state and the output state itself:

$$\frac{\partial}{\partial\omega}\Psi(z;\, \omega) = \left[\frac{\partial}{\partial\omega}\hat{W}(z|0;\, \omega)\right]\hat{W}^{-1}(z|0;\, \omega)\Psi(z;\, \omega). \tag{16}$$

Therefore one finds that the matrix

$$\hat{J} \equiv \left[\frac{\partial}{\partial\omega}\hat{W}(z|0;\ \omega)\right]\hat{W}^{-1}(z|0;\ \omega) \tag{17}$$

is traceless:

$$\mathrm{Tr}(\hat{J}) = \frac{\partial}{\partial\omega}(|u_1|^2 + |u_2|^2) = 0. \tag{18}$$

To find the evolution equation for $\hat{J}$, one uses that

$$\partial_z(\hat{W}^{-1}) = -i\omega\hat{W}^{-1}\hat{m} \tag{19}$$

and then obtains

$$\partial_z\hat{J} = i\hat{m} + i\omega[\hat{m},\ \hat{J}]. \tag{20}$$

Here $[\cdot,\ \cdot]$ denotes the commutator of the two matrices. Motivated by experiments,[10] the coordinate system, called the principal states of polarization,[11] is used for the description of linear PMD. The idea is to write the output signal $\Psi(l;\ \omega)$ (l marks the end of the line) in the form

$$\Psi(l;\ \omega) = \begin{pmatrix}\Psi_1(l;\ \omega)\\ \Psi_2(l;\ \omega)\end{pmatrix} = \psi_p(l,\ \omega)e_p(l,\ \omega), \tag{21}$$

where $\psi_p(l,\ \omega)$ is complex and $e_p(l,\ \omega)$ is a complex unit vector that satisfies

$$\frac{\mathrm{d}}{\mathrm{d}\omega}e_p(l,\ \omega) = 0, \tag{22}$$

as a mathematical formulation of the fact that the principal polarization state of the output signal shall be independent of the frequency to first order. The validity and the extension of the concept of these principal states, from a theoretical point of view and in relation to experiments as well, have been the subject of various publications; we refer, e.g., to Refs. 12, 17, and 18. The principal states are related to the eigenvectors of $\hat{J}$, as the frequency dependence of $\Psi(l;\ \omega)$ can be expressed by

$$\begin{aligned}\frac{\mathrm{d}\Psi(l,\ \omega)}{\mathrm{d}\omega} &= \frac{\mathrm{d}\psi_p(l,\ \omega)}{\mathrm{d}\omega}e_p(l,\ \omega)\\ &= \hat{J}(l,\ \omega)\Psi(l,\ \omega) = \hat{J}(l,\ \omega)\psi_p(l,\ \omega)e_p(l,\ \omega);\end{aligned} \tag{23}$$

hence

$$\hat{J}(l,\ \omega)e_p(l,\ \omega) = \frac{1}{\psi_p(l,\ \omega)}\frac{\mathrm{d}\psi_p(l,\ \omega)}{\mathrm{d}\omega}e_p(l,\ \omega). \tag{24}$$

Therefore $\lambda = [1/\psi_p(l,\ \omega)][\mathrm{d}\psi_p(l,\ \omega)/\mathrm{d}\omega]$ is an eigenvalue of $\hat{J}$, and $e_p(l,\ \omega)$ is the corresponding eigenvector. The eigenvalues can be computed directly from Eq. (13) (Ref. 11) and are given by

$$\lambda = \pm i[|u_1'(l;\ \omega)|^2 + |u_2'(l;\ \omega)|^2]^{1/2}, \tag{25}$$

where the prime denotes differentiation with respect to ω. In experiments one can measure the time delay Δt between the two polarization states. Δt is related to the output signal in the following way: One writes $\psi_p(l,\ \omega)$ in polar coordinates (with real-valued amplitude a_p and real-valued phase ϕ_p), and hence

$$\psi_p(l,\ \omega) = a_p(l,\ \omega)\exp[i\phi_p(l,\ \omega)]. \tag{26}$$

Then one finds

$$\begin{aligned}\lambda &= \frac{1}{\psi_p(l,\ \omega)}\frac{\mathrm{d}\psi_p(l,\ \omega)}{\mathrm{d}\omega}\\ &= \frac{1}{a_p(l,\ \omega)}\frac{\mathrm{d}a_p(l,\ \omega)}{\mathrm{d}\omega} + i\frac{\mathrm{d}\phi_p(l,\ \omega)}{\mathrm{d}\omega}.\end{aligned} \tag{27}$$

Let us denote the two different eigenvalues by λ_+ and λ_- and use this notation for all associated quantities as well. The time delay is related to the phase of the output signal by

$$\begin{aligned}\Delta t &= \frac{\mathrm{d}\phi_{p+}(l,\ \omega)}{\mathrm{d}\omega} - \frac{\mathrm{d}\phi_{p-}(l,\ \omega)}{\mathrm{d}\omega}\\ &= 2[|u_1'(l;\ \omega)|^2 + |u_2'(l;\ \omega)|^2]^{1/2}.\end{aligned} \tag{28}$$

Therefore Δt is directly related to $\hat{J}$ as

$$\det(\hat{J}) = |u_1'(l;\ \omega)|^2 + |u_2'(l;\ \omega)|^2 = \frac{\Delta t^2}{4}. \tag{29}$$

Let us now turn to the second possibility to describe the effects of PMD on the Poincaré sphere. To define the corresponding quantities on the Poincaré sphere, we use the Pauli matrices with the following signs:

$$\hat{\sigma}_1 = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \qquad \hat{\sigma}_2 = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}, \qquad \hat{\sigma}_3 = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}. \tag{30}$$

Following Foschini and Poole,[12] the object associated with the two-component field $\Psi(z;\ \omega)$ is the three-component vector of polarization $\mathbf{\Pi}(z;\ \omega)$ defined by

$$\Pi_i = (\Psi_1^*\Psi_2^*)\hat{\sigma}_i\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix}. \tag{31}$$

The evolution of the signal Ψ according to Eq. (14) corresponds to

$$\mathbf{\Pi}(z;\ \omega) = \mathcal{R}(z;\ \omega)\mathbf{\Pi}_0, \tag{32}$$

where $\mathcal{R}$ is a 3×3 matrix that can be related to $\hat{W}$ by

$$2\mathcal{R}(z;\ \omega) = \begin{bmatrix} u_1^{*2} - u_2^2 + u_1^2 - u_2^{*2} & u_1^{*2} - u_2^2 - u_1^2 + u_2^{*2} & -2(u_1u_2 + u_1^*u_2^*)\\ i(u_1^2 + u_2^{*2} - u_1^{*2} - u_2^2) & u_1^2 + u_2^{*2} + u_1^{*2} + u_2^2 & 2i(u_1^*u_2^* - u_1u_2)\\ 2(u_1u_2^* + u_1^*u_2) & 2i(u_1^*u_2 - u_1u_2^*) & 2(|u_1|^2 - |u_2|^2)\end{bmatrix}. \tag{33}$$

To find the quantity corresponding to $\hat{J}$, one recalls that $\hat{J}$ is traceless. Thus one can define a three-component vector $\mathbf{\Omega}$ by

$$\hat{J} = i\sum_i \Omega_i \sigma_i = i\mathbf{\Omega}\cdot\boldsymbol{\sigma}, \tag{34}$$

and one derives

$$2i\Omega_1 = u_1 u_2' - u_1' u_2 + u_1^{*'} u_2^* - u_1^* u_2^{*'}, \tag{35}$$

$$2\Omega_2 = -u_1' u_2 + u_1 u_2' + u_1^* u_2^{*'} - u_1^{*'} u_2^*, \tag{36}$$

$$i\Omega_3 = u_1' u_1^* + u_2' u_2^* = -u_1 u_1^{*'} - u_2 u_2^{*'}, \tag{37}$$

which also shows that the vector $\mathbf{\Omega}$ is real. $\mathbf{\Omega}$ is usually called the polarization dispersion vector.[2,4,12,14,20,21,32] This new object allows the presentation of the frequency dependence of $\mathbf{\Pi}$ as a rotation on the Poincaré sphere. One takes the definition (31) of $\mathbf{\Pi}$ and uses

$$\frac{\partial}{\partial\omega}\Psi(z;\ \omega) = \begin{bmatrix} i\Omega_3 & i\Omega_1 + \Omega_2 \\ i\Omega_1 - \Omega_2 & -i\Omega_3 \end{bmatrix}\Psi(z;\ \omega)$$

to express the frequency derivatives of Ψ through Ψ itself. The result is

$$\frac{\partial}{\partial\omega}\mathbf{\Pi}(z;\ \omega) = 2[\mathbf{\Pi}(z;\ \omega) \times \mathbf{\Omega}(z;\ \omega)]. \tag{38}$$

From Eqs. (15), (31), and (38), one finds

$$\frac{\partial}{\partial z}\mathbf{\Pi}(z;\ \omega) = 2\omega[\mathbf{\Pi}(z;\ \omega) \times \mathbf{h}(z)]. \tag{39}$$

To describe the evolution in z of the polarization dispersion vector, we go back to Eq. (20) and the definition of $\mathbf{\Omega}$ and find

$$[\hat{m},\ \hat{J}] = i\sum_{k,j} h_k \Omega_j [\hat{\sigma}_k,\ \hat{\sigma}_j] = -2[(h_1\Omega_2 - h_2\Omega_1)\hat{\sigma}_3 - (h_1\Omega_3 - h_3\Omega_1)\hat{\sigma}_2 + (h_2\Omega_3 - h_3\Omega_2)\hat{\sigma}_1], \tag{40}$$

which leads to

$$\frac{\partial}{\partial z}\mathbf{\Omega}(z;\ \omega) = \mathbf{h}(z) - 2\omega[\mathbf{h}(z) \times \mathbf{\Omega}(z;\ \omega)]. \tag{41}$$

This equation can be used to describe the statistics of $\mathbf{\Omega}$ for random birefringence.[12,33,34] Therefore the polarization dispersion vector $\mathbf{\Omega}$ can be used in two ways. First, it can be measured easily by using Eq. (38) (Ref. 3), and second, the length of $\mathbf{\Omega}$ is related to $\hat{J}$ and therefore also to the transfer matrix (z-ordered exponential) $\hat{W}$ by

$$|\mathbf{\Omega}(l,\ \omega)| = \sqrt{\det(\hat{J})} = \frac{\Delta t}{2}, \tag{42}$$

as follows directly from the definition of $\mathbf{\Omega}$ and the above considerations.

C. Polarization-Mode-Dispersion-Induced Growth of the Pulse Width

We show now how $\mathbf{\Omega}$ can be related directly to measurements and numerical simulations. In both cases pulses are used—in telecommunications they represent the input signal, and in the numerical modeling, the initial conditions of Eq. (10).

Let us therefore consider a fixed input pulse and at first put $d = 0$ in Eq. (10):

$$\Psi(0;\ \omega) = \begin{pmatrix} f(\omega) \\ 0 \end{pmatrix} \tag{43}$$

with a real-valued function $f(\omega)$. Then, from the above discussion, a natural object to look at is

$$\begin{pmatrix} \phi_1(\omega) \\ \phi_2(\omega) \end{pmatrix} = \hat{W}(z,\ \omega)\Psi(0;\ \omega), \tag{44}$$

and, in analogy with Eq. (29), one shall look at

$$\tilde{T}^2 = \frac{1}{2\pi}\int [|\phi_1'(\omega)|^2 + \phi_2'(\omega)|^2]\mathrm{d}\omega = \int t^2[|\phi_1(t)|^2 + |\phi_2(t)|^2]\mathrm{d}t, \tag{45}$$

where the prime denotes differentiation with respect to ω and $\phi_1(t)$ is the inverse Fourier transform of $\phi_1(\omega)$. Using the properties of the transfer matrix $\hat{W}$, one finds

$$|\phi_1'|^2 + |\phi_2'|^2 = (|u_1'|^2 + |u_2'|^2)|f|^2 + |f'|^2, \tag{46}$$

and it is therefore suggestive to look at

$$\tilde{T}^2_{\mathrm{PMD}} = \frac{1}{2\pi}\int (|\phi_1'|^2 + |\phi_2'|^2 - |f'|^2)\mathrm{d}\omega \tag{47}$$

to extract the PMD-induced pulse broadening. In a similar way we can include the dispersive case $d \neq 0$ in our considerations. In the linear regime two effects are responsible for pulse broadening: dispersion and PMD. Therefore our objective here is to separate the regular dispersive broadening from the one associated with PMD. From the solution of Eq. (10), it is clear that dispersion is a phase rotation in Fourier space gained with accumulated dispersion $R(z)$. Therefore we have to look at

$$\begin{pmatrix} \psi_1(\omega,\ z) \\ \psi_2(\omega,\ z) \end{pmatrix} = \exp[-i\omega^2 R(z)]\begin{pmatrix} u_1(\omega,\ z)f(\omega) \\ -u_2^*(\omega,\ z)f(\omega) \end{pmatrix}, \tag{48}$$

taking into account

$$|\psi_1'|^2 + |\psi_2'|^2 = (|u_1'|^2 + |u_2'|^2)|f|^2 + |f'|^2 + 4\omega^2 R^2|f|^2 + 2i\omega R(u_1^* u_1' - u_1 u_1'^* + u_2^* u_2' - u_2 u_2'^*)|f|^2, \tag{49}$$

where primes denote again differentiation with respect to frequency. From these formulas it is clear that the growth of the pulse width, measured by T^2_{PMD}, is directly related to the vectors $(u_1,\ u_2)$ and $(u_1',\ u_2')$ and therefore to $\mathbf{\Omega}$.

For instance, for $\langle u_j^* u_j' \rangle = \langle u_j^{*'} u_j \rangle$, one arrives at

$$\langle T^2_{\mathrm{PMD}} \rangle = \frac{1}{2\pi}\left\langle \int (|\psi_1'|^2 + |\psi_2'|^2)\mathrm{d}\omega - \int (4\omega^2 R^2|f|^2 + |f'|^2)\mathrm{d}\omega \right\rangle. \tag{50}$$

This formula shows again that the pulse width is influenced by different effects. If we subtract the initial pulse width and the pulse broadening induced by dispersion, we are left with the pulse broadening caused solely by PMD.

For the special case in which the statistics of $|\boldsymbol{\Omega}|$ do not depend on the frequency, we obtain the important relation

$$\langle T_{\mathrm{PMD}}^2\rangle = \langle|\boldsymbol{\Omega}|^2\rangle \int |f|^2 \mathrm{d}\omega. \tag{51}$$

Therefore we understand the PMD-induced pulse broadening if we understand the statistics of $\boldsymbol{\Omega}$.

D. Multifrequency Statistics of $\boldsymbol{\Omega}$ and Natural Noise

In this subsection we will calculate multifrequency statistics of $\boldsymbol{\Omega}$, all taken at the same z. For this purpose we use Eq. (41) and average over VB statistics described by expression (6). The path-integral representation for the joint probability distribution function of $\Omega_1,\ldots,\Omega_n,\ldots,\Omega_N$, where the index $n = 1,\ldots,N$ marks the set of frequencies, is

$$\begin{aligned}\mathcal{P}(z|\{\boldsymbol{\Omega}_n\}) &= \int_{\{Q_n(z)=\boldsymbol{\Omega}_n\}}\left(\prod_{n=1}^{N} \mathcal{D}Q_n\mathcal{D}\mathbf{p}_n\right)\\ &\quad\times \exp\left[\int_0^z \mathrm{d}z'\left(-i\sum_{n=1}^{N}\mathbf{p}_n\frac{\mathrm{d}}{\mathrm{d}z'}Q_n - \frac{D}{2}\right.\right.\\ &\quad\left.\left.\times\left\{\sum_{n=1}^{N}[\mathbf{p}_n - 2\omega_n(Q_n\times\mathbf{p}_n)]\right\}^2\right)\right]\\ &\quad\times \mathcal{P}(0|\{Q_n(0)\})\\ &= \int_{\{Q_n(z)=\boldsymbol{\Omega}_n\}}\left[\prod_{n=1}^{N}\mathcal{D}Q_n(z)\mathcal{D}\mathbf{p}_n(z)\right]\\ &\quad\times\exp\left(-\int_0^z \mathrm{d}z'\mathcal{S}\right)\mathcal{P}(0|\{Q_n(0)\}),\end{aligned} \tag{52}$$

$$\begin{aligned}\mathcal{S} = i\sum_{n=1}^{N}\mathbf{p}_n\frac{\mathrm{d}}{\mathrm{d}z'}Q_n + \frac{D}{2}\sum_{k,n=1}^{N}\{&\mathbf{p}_k\mathbf{p}_n + 2(\omega_n Q_n\\ &- \omega_k Q_k)(\mathbf{p}_k\times\mathbf{p}_n)\\ &+ 4\omega_n\omega_k[(\mathbf{p}_k\mathbf{p}_n)(Q_kQ_n) - (Q_k\mathbf{p}_n)(Q_n\mathbf{p}_k)]\}.\end{aligned} \tag{53}$$

The Fokker–Planck equation for the probability distribution function is then

$$\begin{aligned}&\left\{\partial_z - \frac{D}{2}\sum_{k,n=1}^{N}[\partial_{\Omega;k}^i\partial_{\Omega;n}^i + 2(\partial_{\Omega;k}\times\partial_{\Omega;n})(\omega_n\boldsymbol{\Omega}_n - \omega_k\boldsymbol{\Omega}_k)\right.\\ &\left.\quad + 4\omega_n\omega_k(\partial_{\Omega;k}^i\Omega_k^j\partial_{\Omega;n}^i\Omega_n^j - \partial_{\Omega;k}^i\Omega_k^j\partial_{\Omega;n}^j\Omega_n^i)]\right\}\mathcal{P} = 0.\end{aligned} \tag{54}$$

One can straightforwardly derive equations for correlation functions of Ω from Eq. (54), first by multiplying both the lhs and the rhs of the equation on the same $\boldsymbol{\Omega}$-dependent object, second by integrating the resulting expression with respect to Ω, and finally by using integration by parts to get a relation (differential equation with respect to z) between different correlation functions of $\boldsymbol{\Omega}$. For example, the sequence of equations for the single-frequency moments of $\boldsymbol{\Omega}$ given by

$$\partial_z\langle\Omega^{2n}\rangle = n(2n+1)D\langle\Omega^{2n-2}\rangle \tag{55}$$

is an immediate consequence of Eq. (54). The single-point probability distribution function of $|\Omega|$ can be reconstructed from Eq. (55):

$$P(|\Omega|) = \frac{4|\Omega|^2}{\Delta^3\sqrt{\pi}}\exp\left(-\frac{\Omega^2}{\Delta^2}\right), \tag{56}$$

$$\Delta^2 \equiv \tfrac{2}{3}\Omega_0^2 + 2Dz. \tag{57}$$

The Maxwellian statistics were found by Poole and coauthors in Refs. 1–4.

The next step is to deduce from Eq. (54) expressions for two-frequency objects. For the correlation functions of $(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)$ and $\Omega_2^2\Omega_1^2$, one gets

$$\begin{aligned}\partial_z\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n\rangle &= \frac{D}{2}\langle(\partial_{\Omega;2}+\partial_{\Omega;1})^2(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n\rangle + 2D(\omega_1\boldsymbol{\Omega}_1\\ &\quad - \omega_2\boldsymbol{\Omega}_2)\langle(\partial_{\Omega;2}\times\partial_{\Omega;1})(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n\rangle\\ &\quad + 2D\left\langle\left\{2\omega_1\omega_2[(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n(\partial_{\Omega;1}\partial_{\Omega;2})\right.\right.\\ &\quad - \Omega_1^i\partial_{\Omega;1}^j\Omega_2^j\partial_{\Omega;2}^i]\\ &\quad + \sum_{k=1}^{2}\omega_k^2(\Omega_k^j\partial_{\Omega;k}^i\Omega_k^j\partial_{\Omega;k}^i\\ &\quad\left.\left. - \Omega_k^j\partial_{\Omega;k}^i\Omega_k^i\partial_{\Omega;k}^j)\right\}(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n\right\rangle\\ &= \frac{n(n-1)}{2}D\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^{n-2}(\Omega_2^2+\Omega_1^2)\rangle\\ &\quad + Dn(n+2)\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^{n-1}\rangle\\ &\quad + 2Dn(n-1)\\ &\quad\times(\omega_1-\omega_2)^2\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^{n-2}\Omega_2^2\Omega_1^2\rangle\\ &\quad - 2Dn(n+1)(\omega_1-\omega_2)^2\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)^n\rangle,\end{aligned} \tag{58}$$

$$\begin{aligned}\partial_z\langle\Omega_2^{2n}\Omega_1^{2m}\rangle &= \frac{D}{2}\langle(\partial_{\Omega;2}+\partial_{\Omega;1})^2\Omega_2^{2n}\Omega_1^{2m}\rangle\\ &= D[n(2n+1)\langle\Omega_2^{2n-2}\Omega_1^{2m}\rangle\\ &\quad + m(2m+1)\langle\Omega_2^{2n}\Omega_1^{2m-2}\rangle]\\ &\quad + 4Dnm\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)\Omega_2^{2n-2}\Omega_1^{2m-2}\rangle.\end{aligned} \tag{59}$$

Equation (58) has a simple solution for the pair correlation function

$$\begin{aligned}\langle(\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)\rangle(z) &= \exp[-4D(\omega_1-\omega_2)^2z](\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2)(0)\\ &\quad + 3\frac{1-\exp[-4D(\omega_1-\omega_2)^2z]}{4(\omega_1-\omega_2)^2},\end{aligned} \tag{60}$$

which was found recently in Ref. 34 by means of direct summation of the discretized set of equations. The result shows that the PMD is actually a single-frequency effect; i.e., at any fixed nonzero value of the frequency shift $\omega_{12} \equiv \omega_1 - \omega_2$, the correlation function goes to a constant at $z \to \infty$. (This is contrary to the linear growth of the single-frequency object.)

Note that the statistics of $(\mathbf{\Omega}_1\mathbf{\Omega}_2)$ are not Gaussian (even though Gaussianity, as we will see below, is restored at $z \to \infty$). This is seen from expressions for the fourth-order correlation function. From Eq. (55) one derives

$$\partial_z\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)^2\rangle = D\langle\Omega_1^2 + \Omega_2^2\rangle + 8D\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle + 4D(\omega_1 - \omega_2)^2\langle\Omega_2^2\Omega_1^2\rangle - 12D(\omega_1 - \omega_2)^2\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)^2\rangle. \quad (61)$$

Finally, from Eqs. (59) and (61), one finds

$$\begin{aligned}\langle\Omega_2^2\Omega_1^2\rangle &= \int_0^z \mathrm{d}z'[3D\langle\Omega_1^2 + \Omega_2^2\rangle + 4D\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle] \\ &= 9D^2z^2 + 3D(\Omega_1^2 + \Omega_2^2)_0 z + (\mathbf{\Omega}_1\mathbf{\Omega}_2)_0\frac{1 - \exp[-4D(\omega_1 - \omega_2)^2 z]}{(\omega_1 - \omega_2)^2} \\ &\quad + 3\,\frac{4D(\omega_1 - \omega_2)^2 z - \{1 - \exp[-4D(\omega_1 - \omega_2)^2 z]\}}{4(\omega_1 - \omega_2)^4},\end{aligned} \quad (62)$$

$$\begin{aligned}\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)^2\rangle &= \exp[-12D(\omega_1 - \omega_2)^2 z](\mathbf{\Omega}_1\mathbf{\Omega}_2)_0^2 + \int_0^z \mathrm{d}z' \exp[-12D(\omega_1 - \omega_2)^2(z - z')][D\langle\Omega_1^2 + \Omega_2^2\rangle \\ &\quad + 8D\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle + 4D(\omega_1 - \omega_2)^2\langle\Omega_2^2\Omega_1^2\rangle]_{z'} \\ &= 3D^2z^2 + \frac{1}{6\omega_{12}^4} + \frac{Dz}{\omega_{12}^2} + Dz(\Omega_1^2 + \Omega_2^2)_0 + \frac{(\mathbf{\Omega}_1\mathbf{\Omega}_2)_0}{3\omega_{12}^2} \\ &\quad + \exp(-4Dz\omega_{12}^2)\left[\frac{(\mathbf{\Omega}_1\mathbf{\Omega}_2)_0}{2\omega_{12}^2} - \frac{3}{8\omega_{12}^4}\right] \\ &\quad + \exp(-12Dz\omega_{12}^2)\left[(\mathbf{\Omega}_1\mathbf{\Omega}_2)_0^2 + \frac{5(\mathbf{\Omega}_1\mathbf{\Omega}_2)_0}{6\omega_{12}^2} + \frac{5}{24\omega_{12}^4}\right].\end{aligned} \quad (63)$$

Let us summarize the results of these computations. In addition to the known results on the statistics of $|\mathbf{\Omega}|$ [Eq. (56)] and the pair correlation function of $\mathbf{\Omega}$ at different frequencies [Eq. (60)], we have explained here the general way to calculate higher-order and multifrequency correlation functions of $\mathbf{\Omega}$. We have applied this general technique to calculate fourth-order two-frequency objects [Eqs. (62) and (63)]. It is found that as $z \to \infty$, the asymptotic statistics of $\mathbf{\Omega}$ are Gaussian and frequency independent. [Higher-order corrections (in terms of $1/z$), accounting for angular correlations, do show, however, some frequency dependence.] These results [in combination with Eq. (51)] let us conclude that the major effect is that the pulse width squared will grow linearly in z because of the effect of PMD. This well-known effect presents a serious limit to telecommunication capacities, and it is therefore of high practical importance to find a way to reduce this growth of the width. Section 4 is devoted to checking if the pinning method, introduced by Eq. (9), is able to achieve the goal.

4. PINNING METHOD

A. Multifrequency Statistics of Ω in the Linear Case

We consider first the statistics of $\mathbf{\Omega}$ within a single pinning leg ($0 < z < L$). The analogs of Eqs. (52) and (53) are

$$\begin{aligned}\mathcal{P}(z|\{\mathbf{\Theta}_n\}) &= \int \mathrm{d}\mathbf{q}\left(\prod_{n=1}^N \mathrm{d}\mathbf{\Phi}_n\right)\exp\left(-\frac{\mathbf{q}^2LD}{2}\right) \\ &\quad \times \left\langle \prod_{n=1}^N \delta(\mathbf{\Omega}_n - \mathbf{\Phi}_n)\right| \\ &\quad \times \exp[-\hat{H}_q(L - z)] \\ &\quad \times \left|\prod_{n=1}^N \delta(\mathbf{\Omega}_n - \mathbf{\Theta}_n)\right\rangle\left\langle \prod_{n=1}^N \delta(\mathbf{\Omega}_n - \mathbf{\Theta}_n)\right| \\ &\quad \times \exp(-\hat{H}_q z)|\mathcal{P}(0|\{\mathbf{\Omega}_n\})\rangle \\ &= \int \mathrm{d}\mathbf{q}\exp\left(-\frac{\mathbf{q}^2LD}{2}\right)\mathcal{P}_q(z|\{\mathbf{\Theta}_n\}),\end{aligned} \quad (64)$$

$$\mathcal{P}_q(z|\{\mathbf{\Omega}_n\}) \equiv \exp(-\hat{H}_q z)|\mathcal{P}(0|\{\mathbf{\Omega}_n\})\rangle,$$
$$\hat{H}_q \equiv \hat{H}_0 + iD\sum_{n=1}^N \partial_{\Omega;n}[\mathbf{q} - 2\omega_n(\mathbf{q}\times\mathbf{\Omega}_n)], \quad (65)$$

$$\begin{aligned}\hat{H}_0 \equiv -\frac{D}{2}\sum_{k,n=1}^N [&\partial_{\Omega;k}^i\partial_{\Omega;n}^i \\ &+ 2(\partial_{\Omega;k} \times \partial_{\Omega;n})(\omega_n\mathbf{\Omega}_n - \omega_k\mathbf{\Omega}_k) \\ &- 4\omega_n\omega_k(\partial_{\Omega;k}^i\Omega_k^j\partial_{\Omega;n}^i\Omega_n^j \\ &- \partial_{\Omega;k}^i\Omega_k^j\partial_{\Omega;n}^j\Omega_n^i)].\end{aligned} \quad (66)$$

One may therefore interpret the transition from the original statistics to the restricted ones as a change from $\mathbf{h}(z)$ to $\mathbf{h}(z) - D\mathbf{q}$, where $\mathbf{q}$ is a static (z-independent) Gaussian field with zero mean and the following pair correlation function:

$$\langle q_i q_j\rangle = \frac{\delta_{ij}}{DL}. \quad (67)$$

The rules of calculations are a little bit more complex here. One first needs to calculate objects conditioned to q and only afterward average the result over statistics of $\mathbf{q}$, described by Eq. (67). Thus, for the first two moments of $\mathbf{\Omega}$, one gets the following set of equations:

$$\partial_z\langle\mathbf{\Omega}\rangle_q = iD\mathbf{q}, \qquad \partial_z\langle\mathbf{\Omega}^2\rangle_q = 2iD\mathbf{q}\langle\mathbf{\Omega}\rangle_q + 3D, \tag{68}$$

$$\begin{aligned}\partial_z\langle\mathbf{\Omega}_1\mathbf{\Omega}_2\rangle_q &= iD\mathbf{q}\langle\mathbf{\Omega}_1 + \mathbf{\Omega}_2\rangle_q \\ &\quad - 2iD\omega_{12}\mathbf{q}\langle\mathbf{\Omega}_1\times\mathbf{\Omega}_2\rangle_q + 3D \\ &\quad - 4D\omega_{12}^2\langle\mathbf{\Omega}_1\mathbf{\Omega}_2\rangle_q,\end{aligned} \tag{69}$$

where $\langle\ \cdot\ \rangle_q$ stands for the average conditioned by q. Integration of Eqs. (68) gives

$$\langle\mathbf{\Omega}_i\rangle_q = \mathbf{\Omega}_i(0) + iD\mathbf{q}z, \tag{70}$$

$$\langle\mathbf{\Omega}_i^2\rangle_q = \mathbf{\Omega}_i^2(0) + 3Dz - D^2q^2z^2 + iDz\mathbf{q}\mathbf{\Omega}_i(0). \tag{71}$$

Let us drop the dependence on the initial conditions at $z = 0$. Then the problem is isotropic, and

$$\langle\Omega_1^\alpha\Omega_2^\beta\rangle_q = \delta^{\alpha\beta}\frac{A}{3} + \left(\frac{\delta^{\alpha\beta}}{3} - \frac{q^\alpha q^\beta}{q^2}\right)B, \tag{72}$$

the cross-product term of Eq. (69), disappears from the problem completely and one gets

$$\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle_{q;\text{isot}} = D\int_0^z \mathrm{d}z'\exp[-4D\omega_{12}^2(z - z')] \times (3 - 2q^2Dz'), \tag{73}$$

where the subscript isot stands for the isotropic part. It is easy to check that the general (anisotropic) solution is

$$\begin{aligned}\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle_q &= \langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle_{q;\text{isot}} + (\mathbf{\Omega}_1(0)\mathbf{\Omega}_2(0)) \\ &\quad\times\exp(-4D\omega_{12}^2z) + iD\mathbf{q}\{\mathbf{\Omega}_1(0) + \mathbf{\Omega}_2(0) \\ &\quad - 2\omega_{12}[\mathbf{\Omega}_1(0)\times\mathbf{\Omega}_2(0)]\} \\ &\quad\times\frac{1 - \exp(-4D\omega_{12}^2z)}{4D\omega_{12}^w},\end{aligned} \tag{74}$$

and after averaging with respect to $\mathbf{q}$ is done, we are left with

$$\begin{aligned}\langle(\mathbf{\Omega}_1\mathbf{\Omega}_2)\rangle &= \left(3 + \frac{\langle q^2\rangle}{2\omega_{12}^2}\right)\frac{1 - \exp(-4D\omega_{12}^2z)}{4\omega_{12}^2} - \frac{zD\langle q^2\rangle}{2\omega_{12}^2} \\ &\quad + (\mathbf{\Omega}_1(0)\mathbf{\Omega}_2(0))\exp(-4D\omega_{12}^2z).\end{aligned} \tag{75}$$

The expression contains, in particular, renormalization of Δ^2 to the Maxwellian distribution (57) as a result of pinning:

$$\Delta_{\text{pin}}^2 = \tfrac{2}{3}[\Omega_0^2 + 3Dz(1 - z/L)]. \tag{76}$$

The formulas derived above are valid within one compensation interval ($0 < z < L$). One may use all of them, however, with some modification for the next subsequent interval(s) as well, replacing $\mathbf{\Omega}_{1,2}(0)$, z and L by $\mathbf{\Omega}_{1,2}(L)$, $z - L$ and $2L$, respectively.

Finally, the conclusion of the linear consideration is that the pinning is ideal, in the sense that it compensates for the growth of Δ^2 [according to the extension of Eq. (76) to many pinning periods] completely. Figure 1 compares the unpinned growth of T_{PMD}^2 given by Eq. (57) with the pinned case (76) by the relation

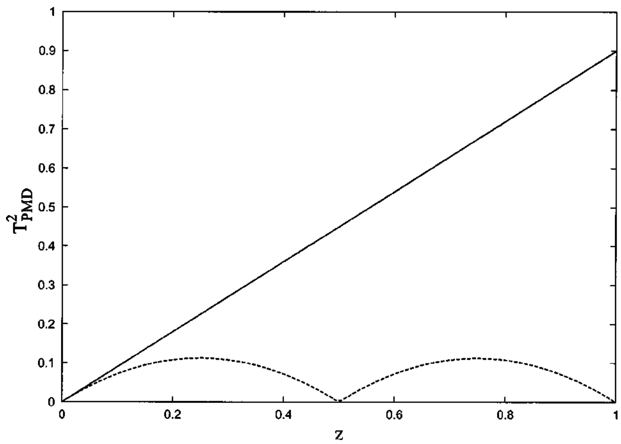

Fig. 1. Pinning in the case of δ-correlated noise. The amplitude of the noise is $D = 0.3$, corresponding to a PMD parameter of $\delta = 0.03\ \text{ps}/\sqrt{\text{km}}$,[23] if the propagation distance is normalized with a length $L = 100$ km. The growth of the pulse width is periodically suppressed. If the system is linear and the noise is δ correlated, this suppression is complete. The pinning period is here (and in the following numerical simulations) 0.5 (in dimensionless units), corresponding to 50 km.

$$\langle T_{\text{PMD}}^2\rangle = \langle|\mathbf{\Omega}|^2\rangle = \tfrac{3}{2}\Delta^2 \tag{77}$$

for a Gaussian input that is normalized by $\int|f(\omega)|^2\mathrm{d}\omega = 1$. From this we can see how pinning works: It is *not* a pointwise compensation, but the compensation is distributed in the whole span [0, 0.5] according to Eq. (9). Because of this modification of the correlation of the noise, the mean of T_{PMD}^2 will come back to zero on the pinning point. This is why the above curves are piecewise parabolic and do not follow a sawtooth.

B. Weak Nonlinearity: Do We Have an Averaged Equation (Self-Averaging)?

In the linear case we can describe the effect of pinning analytically and compute the statistics of $\mathbf{\Omega}$. If we include nonlinearity, the system is much more complex and we have to apply a different method in order to study its properties. In this section we present an analytical analysis of the pinning effect in the case of weak nonlinearity, i.e., when the effect of PMD is strong. (Even though analytical consideration of the pinning effect in the opposite limit of weak PMD in the spirit of Refs. 35–37 is possible, we do not discuss it in this paper.) The more realistic case of moderate-strength PMD is addressed numerically in Section 5.

Weakness of the nonlinearity suggests transition to a "slower" variable $\mathbf{\Phi}$ related to Ψ through

$$\mathbf{\Psi}(t;\, z) \equiv \int_{-\infty}^{+\infty}\mathrm{d}\omega\exp(-i\omega t)\hat{W}(z;\,\omega)\mathbf{\Phi}_\omega(z). \tag{78}$$

The idea of the substitution is obvious. If the nonlinear term is not taken into account, Φ is subjected only to a pure dispersive spreading, as the change of polarization (the major fluctuating effect) is already accounted for through the W term in the integrand of Eq. (78). $\mathbf{\Phi}_\omega(z)$ satisfies the following integral equation:

$$(i\partial_z - \omega_1^2)\Phi_1^\alpha + 2\int \mathrm{d}\omega_{2,3,4}\delta(\omega_1 - \omega_2 + \omega_3 - \omega_4) \times \Xi^{\alpha\beta\nu\eta}(\omega_{1,2,3,4}|z)\mathbf{\Phi}_2^\beta\mathbf{\Phi}_3^{*\nu}\mathbf{\Phi}_4^\eta = 0, \quad (79)$$

$$\Xi^{\alpha\beta\nu\eta}(\omega,\ \omega_{1,2,3,4}|z) \equiv \tfrac{4}{3}[\hat{W}^{-1}(z;\ \omega_1)\hat{W}(z;\ \omega_2)]^{\alpha\beta} \times W^{*\mu\nu}(z;\ \omega_3)W^{\mu\eta}(z;\ \omega_4) + \tfrac{2}{3}[\hat{W}^{-1}(z;\ \omega_1)\hat{W}^*(z;\ \omega_3)]^{\alpha\nu} \times W^{\mu\beta}(z;\ \omega_2)W^{\mu\eta}(z;\ \omega_4). \quad (80)$$

The major question is the following: Is direct averaging of the equation allowed? A more formal way to pose this question is to ask if the quantity

$$\frac{1}{z}\int_0^z \Xi^{\alpha\beta\nu\eta}(\omega_{1,2,3,4}|z')\mathrm{d}z' \quad (81)$$

saturates to a nonfluctuating object at $z \to \infty$ or not.

A similar question has been addressed in Refs. 25 and 26, where the case of strong fluctuations in dispersion (and, respectively, weak nonlinearity) has been studied. The answer to the question was negative for the natural noise in dispersion: The analog of $\mathbf{\Phi}$ fluctuates strongly. The pinning of the integral dispersion, however, results in saturation of the integral kernel fluctuations. This guarantees the deterministic character of the envelope function (analog of Φ) in the $z \to \infty$ limit.

However, our problem here is more complex than the scalar one considered in Ref. 25. The major difference is due to the matrix character of the process: What was an exponential of an integral over a random process now turns into a z-ordered exponential $\hat{W}$. It is clear that the change from a scalar to a matrix process can only worsen the situation with the averaging: A process that is a self-averaged one in the scalar case may either remain self-averaged or lose this feature, but not the other way around (the matrix nature of the process adds fluctuations but cannot remove them).

Thus one concludes that the kernel (81) fluctuates strongly in the case of natural noise in birefringence. This fact is demonstrated in Appendix A, where the average of $\hat{V}(\omega_1|\omega_2) \equiv \hat{W}^{-1}(z;\ \omega_1)\hat{W}(\omega_2)$, which is a multiplicative part of the kernel in Eq. (79), and $\Sigma^{\alpha\beta;\mu\nu}(\omega_1,\ \omega_2;\ \omega_3,\ \omega_4)$, defined as the average of $V^{\alpha\beta}(\omega_1|\omega_2)V^{\mu\nu}(\omega_3|\omega_4)$ [the object is an additive part of the average of the kernel (79)], are calculated. Both $\langle\hat{V}\rangle$ and $\langle\Sigma\rangle$ decay exponentially for large z and a nondegenerate configuration of frequencies. This excludes the possibility of integral (81) being a self-averaged quantity. We have also considered the objects averaged over the pinned noise in Appendix A. Pinning results in reduction of the exponential decay rate of average kernel with z, i.e., the destruction of a pulse is getting weaker. However, the decay rate remains finite. We conclude that the matrix nature of the problem does not allow a complete saturation of the kernel in Eq. (79). In other words: We find that the pinning delays the effective decay of the kernel with z; however, it does not allow a complete compensation of the decay in the weakly nonlinear case.

5. NUMERICAL ANALYSIS OF POLARIZED MODE DISPERSION (COLORED NOISE WITH SHORT CORRELATIONS)

So far, we have considered δ-correlated noise, which is already a good approximation for physical systems. To show that the basic idea that pinning helps to reduce pulse spreading can be extended to the case of short-correlated noise, we performed numerical simulations.

A. Linear Case

Numerical simulations can be performed directly for the basic equation (10). The matrix $\hat{m}$ is a function of z, and, in the analytical calculations, its entries are assumed to obey white-noise statistics. In the numerical simulations the short-correlated noise is generated by an Ornstein–Uhlenbeck process. The equation for a step dz is

$$\hat{W}(z + \mathrm{d}z;\ \omega) = \exp[i(\omega\hat{m} + \omega^2\hat{I}_2)\mathrm{d}z], \quad (82)$$

where $\hat{I}_2 \equiv \mathrm{diag}(1, 1)$ is solved directly by using the explicit formula for a matrix exponential in the case of 2×2 matrices. In our simulations we take an initial Gaussian pulse

$$\psi_{\mathrm{in}}(t) = \pi^{-1/4}\exp(-t^2/2)\begin{pmatrix}1\\0\end{pmatrix} \quad (83)$$

and propagate it. From the output we can evaluate numerically T_{PMD}^2 by Eq. (50) and find its average by simulating different realizations.

Figure 2 shows that, in the unpinned case, $\langle T_{\mathrm{PMD}}^2\rangle$ grows linearly with z and, in the pinned case, its evolution is almost along parabolas, as predicted by Eq. (76). It is seen, however, that unlike the δ-correlated theory explained above, the compensation is not complete. We attribute this to the finite-correlated nature of the noise. We have checked that once the correlation length of the noise is decreased (while its strength is kept the same), the accumulated width (i.e., measure of the compensation incompleteness) decreases.

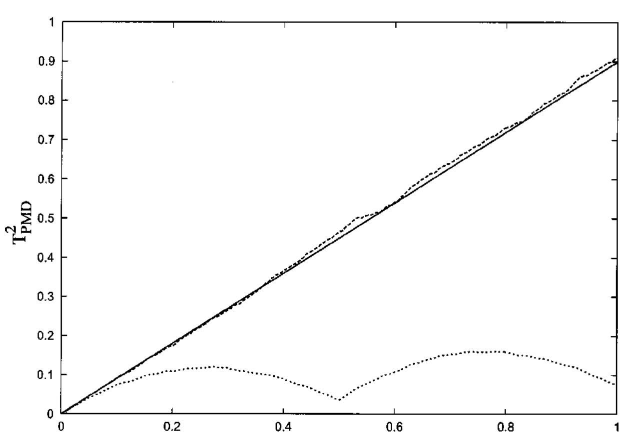

Fig. 2. Numerical results for the linear case for short-correlated noise. The amplitude of the noise is $D = 0.3$. The solid line represents the analytical result ($3Dz$) of the δ-correlated case, the dashes represent the numerical simulations in the unpinned case for short-correlated noise, and the dots represent the pinned case for short-correlated noise. T_{PMD}^2 is averaged over 500 realizations. The correlation length is 0.002. Again, the pinning period is 0.5, and the total length is normalized to unity.

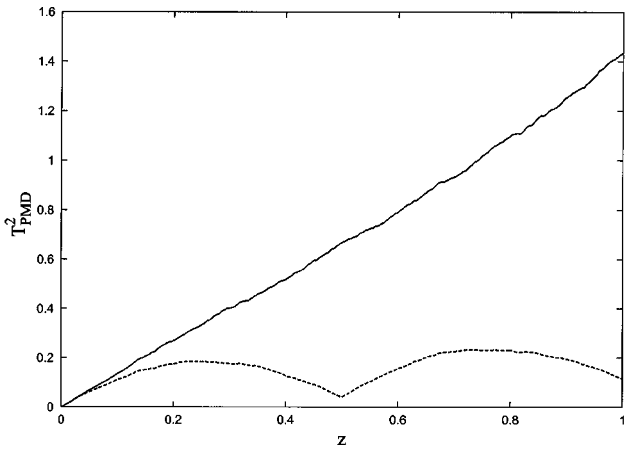

Fig. 3. Numerical results for the nonlinear case with the initial condition being a perfect soliton. The effect of PMD is considerably reduced by the application of pinning. The solid curve is for the unpinned case, and the dashed curve shows the pinned-case curve. The other parameters of the simulation are the same as those in the linear case.

B. Numerical Analysis of the Nonlinear Polarization Mode Dispersion Equations

1. New Numerical Scheme for Solving the Nonlinear Equations

To compare the analytical results derived above with numerical results, we performed numerical solutions of the full system (1). In these numerical simulations we used the usual split-step scheme. The linear step used the solution in Fourier space in the same way as it was presented in Section 3 on linear PMD. In the general case for a nonlinear term, presented in Eq. (3), the nonlinear step can be performed by using a Runge–Kutta integrator.[23] A different treatment is also possible. We consider the equations governing the nonlinear step in the form

$$\partial_z \Psi_1 = ic(3|\Psi_1|^2\Psi_1 + 2|\Psi_2|^2\Psi_1 + \Psi_1^*\Psi_2^2),$$

$$\partial_z \Psi_2 = ic(3|\Psi_2|^2\Psi_2 + 2|\Psi_1|^2\Psi_2 + \Psi_2^*\Psi_1^2), \quad (84)$$

with $c = 1/3$. From these equations we obtain directly conservation of energy, i.e.,

$$|\Psi_1|^2 + |\Psi_2|^2 = \text{const.} = E, \quad (85)$$

and, additionally,

$$\Psi_1\Psi_2^* - \Psi_2\Psi_1^* = \text{const.}, \quad (86)$$

$$(\Psi_1^2 + \Psi_2^2)(z) = \exp(6icEz)\zeta(0),$$

$$\zeta(0) = (\Psi_1^2 + \Psi_2^2)(0). \quad (87)$$

These equations can be used to decouple the system (84):

$$\partial_z \Psi_\alpha = ic[2E\Psi_\alpha + \exp(6icEz)\zeta(0)\Psi_\alpha^*]. \quad (88)$$

With the transformation $\Psi_\alpha = \eta_\alpha \exp(3icEz)$, one obtains

$$\partial_z \eta_\alpha = ic[-E\eta_\alpha + \zeta(0)\eta_\alpha^*]. \quad (89)$$

This system of linear equations can be solved easily by calculating the matrix exponential of the corresponding real equations. We compared this scheme, used in a split-step integrator, with a finite-difference scheme for a fixed value of the birefringence matrix. Both schemes yield the same result.

2. Numerical Results in the Nonlinear Case

Let us now consider the case of a soliton having the same amplitude as that of the linear pulse considered above for a dispersion $d(z) = 1.0$. Now the equations will be coupled in the linear and in the nonlinear part, and we want to check if pinning allows reduction of the effect of PMD-induced broadening. As the soliton is considered, one has to subtract the initial width; hence we are looking again at

$$\tilde{T}^2_{\text{PMD}} = \frac{1}{2\pi}\int (|\phi_1'|^2 + |\phi_2'|^2 - |f'|^2)\mathrm{d}\omega. \quad (90)$$

The numerical simulations presented in Fig. 3 show that even in the nonlinear case, pinning is an appropriate method to reduce the effect of PMD-induced pulse broadening. It is interesting to note that if we compare the results of the numerical simulations of the linear and nonlinear cases, T^2_{PMD} grows faster in the nonlinear case. This is not surprising, as the nonlinear terms in Eq. (1) introduce additional energy flow from the excited polarization state to the initial unexcited state.

6. CONCLUSIONS

We have shown that pinning is an effective method to reduce broadening of the pulse width induced by PMD. For linear, δ-correlated noise we proved this analytically. In the weakly nonlinear, δ-correlated case, we found it helpful to analyze the kernel of an associated integral equation. The calculations show that the effect of pinning is positive; however, it is not able to completely suppress the PMD-induced pulse broadening. For linear and nonlinear systems, the effectiveness of the pinning method was shown numerically for finite- but short-correlated noise in birefringence. For this purpose we suggested an effective numerical way to perform integration of the coupled nonlinear equations.

APPENDIX A: AVERAGE KERNEL

In this technical appendix we first calculate the auxiliary object $\langle V\rangle$, where

$$\hat{V}(\omega_1|\omega_2) \equiv \hat{W}^{-1}(z;\ \omega_1)\hat{W}(\omega_2). \quad (A1)$$

$\langle V\rangle$ is found in Appendix A.1 for both the unpinned and pinned cases. Evaluation of these simpler averages helps us in Appendix A.2 to find

$$\Sigma^{\alpha\beta;\mu\nu}(\omega_1,\ \omega_2;\ \omega_3,\ \omega_4) \quad (A2)$$

$$= \langle V^{\alpha\beta}(\omega_1|\omega_2)V^{\mu\nu}(\omega_3|\omega_4)\rangle, \quad (A3)$$

which is the average of an additive part of the kernel Ξ.

Let us first look at a discretization in z for the unpinned case. In this case the transition from $\hat{W}(z - \epsilon)$ to $\hat{W}(z)$, representing one infinitesimal step from $z - \epsilon$ to z, can be described by the following relation:

$$\hat{W}(z) = [1 + i\omega\epsilon h_i(z)\hat{\sigma}_i - \tfrac{1}{2}\omega^2\epsilon^2(h_i\hat{\sigma}_i)^2]\hat{W}(z - \epsilon), \tag{A4}$$

where h_i is a Gaussian zero mean field described by

$$\langle h_i(z)h_j(z)\rangle = \frac{D}{\epsilon}\delta_{ij}. \tag{A5}$$

The generalization of the key differential formula (A4) for the pinned case is

$$\hat{W}(z|q) = \{1 + i\omega\epsilon[h_i(z) - Dq_i]\hat{\sigma}_i - \tfrac{1}{2}\omega^2\epsilon^2(h_i\hat{\sigma}_i)^2\} \times \hat{W}(z - \epsilon|q), \tag{A6}$$

where averaging over q should be done according to Eq. (67).

1. Calculation of $\langle V\rangle$

a. Natural Disorder (Unpinned) Case

From Eqs. (A4) and (A5), one derives

$$\frac{\mathrm{d}}{\mathrm{d}z}\langle\hat{V}\rangle = -\frac{3D}{2}\omega_{12}^2\langle\hat{V}\rangle. \tag{A7}$$

The solution of Eq. (A7) is

$$\langle\hat{V}(\omega_1|\omega_2)\rangle = \hat{1}\exp(-3D\omega_{12}^2 z/2), \tag{A8}$$

where $\hat{1}$ stands for a unitary 2×2 matrix.

b. Pinned Case

Calculating the average $\hat{V}$, conditioned to the given q, one finds that the description is closed in terms of the two objects, i.e., one gets the following closed system of equations for the two auxiliary objects:

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\langle\hat{V}\rangle_q \\ q_i\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]\rangle_q\end{pmatrix} = D\begin{bmatrix} -\dfrac{3\omega_{12}^2}{2} & i\omega_{12} \\ iq^2\omega_{12} & -\dfrac{\omega_{12}^2}{2}\end{bmatrix}\begin{pmatrix}\langle\hat{V}\rangle_q \\ q_i\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]\rangle_q\end{pmatrix}. \tag{A9}$$

Equation (A9) is straightforward to solve. Averaging the result over statistics of q, and taking into account that $\langle q_i\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]\rangle\rangle_q$ at $z = 0$ is zero (because of isotropy), one derives

$$\langle\hat{V}(z)\rangle = \hat{1}\frac{(DL)^{3/2}\exp(-Dz\omega_{12}^2)}{\sqrt{2\pi}}\int_0^\infty q^2\mathrm{d}q \times \exp\left(-\frac{q^2DL}{2}\right)\left[\cosh\left(Dz\frac{\omega_{12}}{2}\sqrt{\omega_{12}^2 - 4q^2}\right) - \frac{\omega_{12}}{\sqrt{\omega_{12}^2 - 4q^2}}\sinh\left(Dz\frac{\omega_{12}}{2}\sqrt{\omega_{12}^2 - 4q^2}\right)\right]. \tag{A10}$$

Generalization of the object accounting for evolution through n pinning legs (each of length l) is $\langle\hat{V}(n \times l)\rangle = \langle\hat{V}(l)\rangle^n$. One concludes that at $z \gg L = l \times n$, $\langle\hat{V}\rangle$ decays exponentially with L, even though the rate of decay is reduced by the factor 1/3.

2. Calculation of Σ

a. Unpinned Case

In addition to Σ, it is also convenient to introduce an auxiliary object

$$\Phi^{\alpha\beta;\mu\eta}(z) = \langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]^{\alpha\beta}[\hat{W}_3^{-1}(z)\hat{\sigma}_i\hat{W}_4(z)]^{\mu\eta}\rangle. \tag{A11}$$

From Eqs. (A4) and (A5) and also taking into account the Pauli matrix relations

$$\hat{\sigma}_i\hat{\sigma}_j\hat{\sigma}_i = -\hat{\sigma}_j, \qquad \hat{\sigma}_i\hat{\sigma}_j = \delta_{ij} + i\varepsilon_{ijk}\hat{\sigma}_k, \tag{A12}$$

one gets the following system of linear equations for differential change of Σ and Φ in z:

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\Sigma^{\alpha\beta;\mu\eta} \\ \Phi^{\alpha\beta;\mu\eta}\end{pmatrix} = -\frac{3D}{2}\begin{bmatrix}\omega_{12}^2 + \omega_{34}^2 & \tfrac{2}{3}\omega_{12}\omega_{34} \\ 2\omega_{12}\omega_{34} & \omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2 + \tfrac{2}{3}\omega_1\omega_2 + \tfrac{2}{3}\omega_3\omega_4 - \tfrac{4}{3}(\omega_1 + \omega_2)(\omega_3 + \omega_4)\end{bmatrix}\begin{pmatrix}\Sigma^{\alpha\beta;\mu\eta} \\ \Phi^{\alpha\beta;\mu\eta}\end{pmatrix}. \tag{A13}$$

Integration of Eq. (A13) is straightforward but bulky, and we do not present it here so as to save space.

b. Pinned Case

The generalization of Eq. (A13) (thus conditioned to a certain value of q) becomes

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\Sigma_q^{\alpha\beta;\mu\eta} \\ \Phi_q^{\alpha\beta;\mu\eta}\end{pmatrix} = -\frac{3D}{2}\begin{bmatrix}\omega_{12}^2 + \omega_{34}^2 & \tfrac{2}{3}\omega_{12}\omega_{34} \\ 2\omega_{12}\omega_{34} & \omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2 + \tfrac{2}{3}\omega_1\omega_2 + \tfrac{2}{3}\omega_3\omega_4 - \tfrac{4}{3}(\omega_1 + \omega_2)(\omega_3 + \omega_4)\end{bmatrix}\begin{pmatrix}\Sigma_q^{\alpha\beta;\mu\eta} \\ \Phi_q^{\alpha\beta;\mu\eta}\end{pmatrix} + D\begin{pmatrix} i\omega_{12}X_{12;34}^{\alpha\beta;\mu\eta} + i\omega_{34}X_{34;12}^{\mu\eta;\alpha\beta} \\ i\omega_{34}X_{12;34}^{\alpha\beta;\mu\eta} + i\omega_{12}X_{34;12}^{\mu\eta;\alpha\beta} + (\omega_{12} - \omega_{34})\Upsilon_{12;34}^{\alpha\beta;\mu\eta}\end{pmatrix}, \tag{A14}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}\Sigma\\ \Phi\\ Y\end{pmatrix} = D\begin{bmatrix} -3\omega_{12}^2 & -\omega_{12}^2 & 2i\omega_{12}\\ -\frac{3}{2}\omega_{12}^2 & -\omega_1^2 + 2\omega_1\omega_2 - 5\omega_2^2 + 8\omega_2\omega_3 - 4\omega_3^2 & 2i\omega_{12}\\ iq^2\omega_{12} & iq^2\omega_{12} & -4(\omega_1^2 + \omega_2^2 - \omega_1\omega_2)\end{bmatrix}\begin{pmatrix}\Sigma\\ \Phi\\ Y\end{pmatrix}. \quad (A21)$$

where the following new objects have been introduced:

$$X_{12;34}^{\alpha\beta;\mu\eta}(z|q) \equiv q_i\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]^{\alpha\beta}[\hat{W}_3^{-1}(z)\hat{W}_4(z)]^{\mu\eta}\rangle, \quad (A15)$$

$$\Upsilon_{12;34}^{\alpha\beta;\mu\eta}(z|q) \equiv \varepsilon_{ijk}q_i\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_j\hat{W}_2(z)]^{\alpha\beta}[\hat{W}_3^{-1}(z)\hat{\sigma}_k\hat{W}_4(z)]^{\mu\eta}\rangle. \quad (A16)$$

The evolution equation for the derived object X is

$$\frac{\mathrm{d}}{\mathrm{d}z}X_{12;34}^{\alpha\beta;\mu\eta} = iq^2D\omega_{12}\Sigma_q^{\alpha\beta;\mu\eta} + iD\omega_{34}q^2G^{\alpha\beta;\mu\eta} - \omega_{12}\omega_{34}D(X_{34;12}^{\mu\eta;\alpha\beta} + i\Upsilon_{12;34}^{\alpha\beta;\mu\eta}) - \frac{3D}{2}\left(\omega_1^2 + \omega_2^2 + \frac{2}{3}\omega_1\omega_2 + \omega_{34}^2\right)X_{12;34}^{\alpha\beta;\mu\eta}, \quad (A17)$$

$$G^{\alpha\beta;\mu\eta} \equiv \frac{q_iq_j}{q^2}\langle[\hat{W}_1^{-1}(z)\hat{\sigma}_i\hat{W}_2(z)]^{\alpha\beta}[\hat{W}_3^{-1}(z)\hat{\sigma}_j\hat{W}_4(z)]^{\mu\eta}\rangle, \quad (A18)$$

$$\frac{\mathrm{d}}{\mathrm{d}z}G^{\alpha\beta;\mu\eta} = -\frac{3D}{2}\left[\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2 + \frac{2}{3}\omega_1\omega_2 + \frac{2}{3}\omega_3\omega_4 - \frac{4}{3}(\omega_1 + \omega_2)(\omega_3 + \omega_4)\right]G^{\alpha\beta;\mu\eta} - \omega_{12}\omega_{34}D(\Sigma_q^{\alpha\beta;\mu\eta} + G^{\alpha\beta;\mu\eta} - \Phi_q^{\alpha\beta;\mu\eta}) + iD(\omega_{34}X_{12;34}^{\alpha\beta;\mu\eta} + \omega_{12}X_{34;12}^{\mu\eta;\alpha\beta}). \quad (A19)$$

It is also straightforward to check that $\Upsilon_{12;34}^{\alpha\beta;\mu\eta}(z|q) = 0$ [since the respective equation takes the form $(\mathrm{d}/\mathrm{d}z)\Upsilon \sim \Upsilon$, while Υ is zero initially, at $z = 0$]. The resonance condition $\omega_{12} = \omega_{34}$ guarantees that the different X objects enter the problem only through

$$Y_{12;34}^{\alpha\beta;\mu\eta} \equiv X_{12;34}^{\alpha\beta;\mu\eta} + X_{34;12}^{\mu\eta;\alpha\beta}. \quad (A20)$$

Yet another simplification is due to the fact that while at $z = 0$, we have $G^{\alpha\beta;\mu\eta}(0) = \Phi_q^{\alpha\beta;\mu\eta}$, the equations for G and Φ are generally (at any z) symmetric under the $G \leftrightarrow \Phi$ transformation. This results in the equality $G = \Phi$ being valid at any z. The final system of linear equations for the three (as yet not averaged over q) objects is

The system of equations can be solved for any value of q (the result will be expressed in terms of the roots of a cubic equation), and subsequent averaging over q is (at least formally) straightforward.

ACKNOWLEDGMENTS

We thank A. Fouxon, V. Lebedev, E. Kuznetsov, R. Moore, and V. Zakharov for useful comments. The support of the Laboratory Directed Research and Development Exploratory Research Program Statistical Physics of Fiber Communications at Los Alamos National Laboratory, U.S. Department of Energy, through contract W-7-405-ENG-36 and its Program in AMS, KJ-01-01 (I. Gabitov), and of the Russian Foundation for Basic Research through grant 03-02-16147a and the Russian Foundation for Promotion of Science (I. Kolokolov) is greatly acknowledged. T. Schäfer's work was supported by the National Science Foundation under grant 0073923.

REFERENCES

1. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," Opt. Lett. **13**, 687–689 (1988).
2. C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," Opt. Lett. **14**, 523–525 (1989).
3. C. D. Poole, N. S. Bergano, R. E. Wagner, and H. J. Schulte, "Polarization dispersion and principal states in a 147-km undersea lightwave cable," J. Lightwave Technol. **6**, 1185–1190 (1988).
4. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," Opt. Lett. **16**, 372–374 (1991).
5. J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," Proc. Natl. Acad. Sci. U.S.A. **97**, 4541–4550 (2000).
6. S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single-mode fibers," Opt. Lett. **3**, 60–62 (1978).
7. 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. **LT-5**, 1618–1622 (1987).
8. N. Gisin, B. Gisin, J. P. Von der Weid, and R. Passy, "How accurately can one measure a statistical quantity like polarization-mode dispersion," IEEE Photon. Technol. Lett. **8**, 1671–1673 (1996).
9. L. E. Nelson, R. M. Jopson, H. Kogelnik, and J. P. Gordon, "Measurement of polarization mode dispersion vectors using the polarization-dependent signal delay method," Opt. Express **6**, 158–167 (2000).
10. S. Machida, I. Sakai, and T. Kimura, "Polarization conservation in single-mode fibres," Electron. Lett. **17**, 494–495 (1981).
11. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibres," Electron. Lett. **22**, 1029–1030 (1986).
12. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fiber," J. Lightwave Technol. **9**, 1439–1456 (1991).

13. F. C. Curti, B. Daino, G. de Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," J. Lightwave Technol. **8**, 1162–1165 (1990).
14. D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurements of the group velocity difference between the principal states of polarization on a low-birefringence terrestrial fiber cable," Opt. Lett. **12**, 844–846 (1987).
15. N. Gisin and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," Opt. Commun. **142**, 119–125 (1997).
16. L. Chen, J. Cameron, and X. Bao, "Statistics of polarization mode dispersion in presence of the polarization dependent loss in single mode fibers," Opt. Commun. **169**, 69–73 (1999).
17. Y. Li and A. Yariv, "Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses," J. Opt. Soc. Am. B **17**, 1821–1827 (2000).
18. M. Midrio, "Nonlinear principal states of polarization in optical fibers with randomly varying birefringence," J. Opt. Soc. Am. B **17**, 169–177 (2000).
19. I. P. Kaminow, "Polarization in optical fibers," IEEE J. Quantum Electron. **QE-17**, 15–22 (1981).
20. W. Eickhoff, Y. Yen, and R. Ulrich, "Wavelength dependence of birefringence in single-mode fiber," Appl. Opt. **20**, 3428–3435 (1981).
21. R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," Appl. Opt. **18**, 2241–2251 (1979).
22. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, New York, 1989).
23. A. Carena, V. Curri, R. Gausino, P. Poggiolini, and S. Benedetto, "A time-domain optical transmission system simulation package accounting for nonlinear and polarization-related effects in fiber," IEEE J. Sel. Areas Commun. **15**, 751–765 (1997).
24. A. L. Berkhoer and V. E. Zakharov, "Self excitation of waves with different polarizations in nonlinear media," Sov. Phys. JETP **31**, 486–490 (1970).
25. M. Chertkov, I. Gabitov, and J. Moeser, "Pulse confinement in optical fibers with random dispersion," Proc. Natl. Acad. Sci. U.S.A. **98**, 14208–14211 (2001).
26. M. Chertkov, I. Gabitov, P. Lushnikov, J. Moeser, and Z. Toroczkai, "Pinning method of pulse confinement in optical fiber with random dispersion," J. Opt. Soc. Am. B **19**, 2538–2550 (2002).
27. 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).
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. P. K. Wai, W. L. Kath, C. R. Menyuk, and J. W. Zhang, "Nonlinear polarization-mode dispersion in optical fibers with randomly varying birefringence," J. Opt. Soc. Am. B **14**, 2967–2979 (1997).
30. M. Chertkov, I. Gabitov, I. Kolokolov, V. Lebedev, and K. Turitsyn, "Solitons, birefringence and disorder," in preparation.
31. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
32. R. Ulrich, "Representation of codirectional coupled waves," Opt. Lett. **1**, 109–111 (1977).
33. N. Gisin and J. P. Pellaux, "Polarization mode dispersion: time versus frequency domains," Opt. Commun. **89**, 316–323 (1992).
34. M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," Opt. Lett. **24**, 939–941 (1999).
35. M. Chertkov, I. Gabitov, I. Kolokolov, and V. Lebedev, "Shedding and interaction of solitons in an imperfect medium," JETP Lett. **74**, 357–361 (2001).
36. M. Chertkov, I. Gabitov, I. Kolokolov, and V. Lebedev, "Solitons in a disordered anisotropic optical medium," JETP Lett. **74**, 535–538 (2001).
37. M. Chertkov, Y. Chung, A. Dyachenko, I. Gabitov, I. Kolokolov, and V. Lebedev, "Shedding and interaction of solitons in weakly disordered optical fibers," Phys. Rev. E **67**, 036615 (2003).