Perturbation theory for the Manakov soliton and its applications to pulse propagation in randomly birefringent fibers

T. I. Lakoba
Department of Mathematics and Computer Science, Clarkson University, Potsdam, New York 13699-5817

D. J. Kaup
Department of Mathematics and Computer Science, Clarkson University, Potsdam, New York 13699-5815
and Department of Physics, Clarkson University, Potsdam, New York 13699-5815

(Received 2 May 1997)

We develop perturbation theory for the soliton of the Manakov equation, and apply it to the problem of soliton propagation in randomly birefringent fibers. We calculate both the slow evolution of the soliton parameters (through second order for two of them) as well as the radiation emitted by the soliton. Our analytical results agree well with the corresponding numerical results of earlier studies. We also relate results obtained with perturbation theory with those obtained from the intuitive picture of the dynamics of the soliton components as interacting quasiparticles. [S1063-651X(97)12611-3]

PACS number(s): 42.65.Tg

I. INTRODUCTION

In this work, we develop perturbation theory for the one-soliton solution of the exactly integrable, two-component vector nonlinear Schrödinger (NLS) model, and apply the results to a particular problem in the optical fiber communications. The model which we consider has the form

$$
\begin{aligned}
&iu_z+u_{\tau\tau}+2u(|u|^2+|v|^2)=\epsilon R_u[u,v,z,\tau],\\
&iv_z+v_{\tau\tau}+2v(|u|^2+|v|^2)=\epsilon R_v[u,v,z,\tau],
\end{aligned}
\tag{1.1}
$$

where, in the context of nonlinear optics, u and v are the envelopes of the electric field in the fiber; z and τ are, respectively, the distance along the fiber and the time in the pulse's reference frame (thus z being the evolution coordinate); ϵR_u and ϵR_v are perturbations of arbitrary form; and ϵ is a small parameter characterizing the magnitude of these perturbations. Manakov [1] was the first to show that Eqs. (1.1) with $\epsilon=0$ are integrable by the method of the inverse scattering transform (IST); therefore, we will refer to the left-hand side of Eqs. (1.1) as the Manakov equation. It should be noted that the perturbation theory for a rather general class of exactly integrable (by the IST) evolution equations, which includes the Manakov equation as a special case, was already worked out in Ref. [2]. The results of Ref. [2] allow one, in principle, to find both the evolution of the scattering data associated with the pulse, and the form of the so-called squared eigenfunctions, which give the basis functions over which an arbitrary perturbation of the pulse's profile can be expanded (see, e.g., Refs. [3,4]). However, as those results refer to a general class of equations, one would have to extract from them the explicit formulas necessary for application to a concrete problem. In fact, in two very recent works [5,6], the IST has been used to develop the perturbation theory for the Manakov equation with the background solution of a general (in particular, N-soliton) form.

The perturbation theory that we present below does *not* make use of the IST formalism for the Manakov equation. Instead, we take a more direct approach and exploit the similarity between the one-soliton solution of the Manakov equation and that of a single NLS equation. The soliton of the Manakov equation has the form:

$$
\begin{aligned}
&u_0=A\cos\beta\,\mathrm{sech}A\tau\, e^{iA^2z},\\
&v_0=A\sin\beta\,\mathrm{sech}A\tau\, e^{iA^2z},
\end{aligned}
\tag{1.2}
$$

where we assumed for the moment that the velocity of the soliton, the position of its center, and the constant phases of both u and v components are all zero. In Eq. (1.2), the constant variable β, which is the polarization angle of the soliton, can take on arbitrary (real) values. Also, the Manakov equation, i.e., the system of Eqs. (1.1) with $\epsilon=0$, is invariant with respect to the unitary transformation

$$
\begin{pmatrix} u' \\ v' \end{pmatrix}=\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix},
\tag{1.3}
$$

by means of which the angle β in Eq. (1.2) can always be made equal to zero. In the latter case, the first component of the soliton (1.2) reduces to the soliton of the scalar NLS equation, and then the first of Eqs. (1.1), with $\epsilon=0$, reduces to the NLS equation. The second of those equations becomes satisfied identically, because for $\beta=0$, we have $v\equiv0$. When one slightly perturbs the resulting equations by allowing ϵ to be small but nonzero, the solution will acquire the form

$$
u=u_0|_{\beta=0}+\epsilon u_1+\cdots,\quad v=\epsilon v_1+\cdots.
$$

Then for u_1, one obtains exactly the linearized NLS equation with a background of a pure soliton. The solution to this equation has been known for a long time [4,7–10]; it constitutes the perturbation theory for the NLS soliton. The equation for v_1 now becomes simply the *linear* Schrödinger equation with a sech^2 potential, whose solution can be found in almost any textbook on quantum mechanics. Thus the perturbation theory for the single soliton of the Manakov equation can be reduced to two well-known problems. The details of this approach will be given in Sec. II.

It should be mentioned that, earlier, Malomed [11] found evolution equations for the parameters of the soliton of the perturbed Manakov equation by using the averaged Lagrangian density. This allowed him to find the adiabatic evolution of the perturbed soliton. On the other hand, our results provide full information about the evolution of both adiabatic and nonadiabatic effects, with the latter describing the radiation emitted by the perturbed soliton.

In Refs. [12–14], it was shown that the Manakov equation governs the average evolution of a pulse in long optical communication fibers. It is known that such fibers are birefringent, with the strength, as well as the direction of the axis of birefringence, varying over a distance which is on the order of 10–100 m (Ref. [15]; see also Ref. [14]). However, the autocorrelation spectrum of these random variations may have components from a *few centimeters* to tens or hundreds of meters [16]. This causes depolarization of the initial pulse [14,16] over tens or hundreds of meters, which, in turn, over even longer distances, leads to the equalization, on the average, of the nonlinear self- and cross-phase modulation coefficients in the equations for the pulse components. Thus the Manakov equation results.

One should note that studies of pulse propagation in optical fibers based on the single NLS equation have proven themselves so successful (in the sense that their results agree quite well with the results of real physical experiments) for two main reasons. The first reason is the aforementioned similarity between the one-soliton solutions of the NLS and Manakov equations. The second reason is that most of the perturbations which were considered for the NLS, if they were rewritten for Eqs. (1.1), would not destroy the invariance of those equations with respect to transformation (1.3). However, in Ref. [14] it was shown that random variations of the birefringence parameters in the fiber produce two perturbation terms, which in Ref. [14] were called linear and nonlinear polarization mode dispersions (PMD's), which *do* destroy that invariance. Therefore, to determine their effect on the soliton, the perturbation theory for the NLS is no longer adequate, and the corresponding theory for the Manakov equation is required.

In this work, we will also apply the perturbation theory, that we have developed for the Manakov equation, to analytically study the effects of both linear and nonlinear PMD's on the soliton. These effects have been studied numerically in Refs. [12,17]. Also, the deterministic analog (see below) of the linear PMD alone, without its nonlinear counterpart, has been thoroughly studied by both analytical [18–21] and numerical [20,22] means for two nonlinearly coupled NLS equations, with the cross-coupling coefficient being not necessarily unity [as it is in Eq. (1.1)]. This deterministic birefringence can be modeled by adding the terms $(+i\delta u_\tau)$ and $(-i\delta v_\tau)$ to the left-hand sides of the first and second of the corresponding analogs of Eqs. (1.1), respectively. Here 2δ is the group velocity difference between the u and v components. When the parameter δ is sufficiently small, then, for a given amplitude of the soliton, the pulses in the individual components propagate with the same velocity, because in this case, the cross-nonlinearity is strong enough to hold them together in a bound state. When δ exceeds a certain threshold value $\delta_{\mathrm{thr}} \approx 5.0$ [19] [here and below we use the units of Eqs. (1.1)], then a splitting of that bound state occurs, and the soliton components escape from each other, with each of the resulting ''principal'' solitons being possibly accompanied by a ''shadow'' in the other component. The speed with which the escaping solitons separate is approximately proportional to $(\delta-\delta_{\mathrm{thr}})$. On the other hand, when the birefringence is random, numerical simulations [12] have shown that its only effects are ''soliton delay, deformation of the soliton, and creation of shadows.'' [Note that, in Ref. [12], $\delta(z)$ was taken as randomly assuming only two values, $\pm\delta$.] No splitting of the soliton was observed, even for a rather high value of the birefringence parameter, $\delta=15.0$, at which the simulations in Ref. [12] were run.

Let us note, however, that the outcome of the pulse evolution for random birefringence will crucially depend (for a given initial amplitude of the soliton) on the relation between the birefringence strength and its correlation length l_{cor}. The latter is a typical distance over which the birefringence parameter(s) may vary considerably. It was shown in Ref. [17], by means of numerical simulations, that the random birefringence will almost certainly not split the composite soliton of a unit amplitude if

$$\delta^2 l_{\mathrm{cor}} < 0.05 . \tag{1.4}$$

Here the correspondence between the notations of this paper and those of Ref. [17] is:

$$(\delta, l_{\mathrm{cor}})_{\text{ this paper}} = [2\delta, 1/(2h)]_{\text{ Ref. [17]}} .$$

In Sec. V, we demonstrate how the numerical criterion (1.4) can be obtained, in the order-of-magnitude sense, by analytical means. It is very important to emphasize that the criterion of the soliton splitting *cannot* be obtained from the perturbation theory. Indeed, the latter is based on the fundamental assumption that the discrete eigenvalue that corresponds to the soliton in the associated scattering problem [1] is only slightly changed by the perturbation. Conversely, the splitting implies a significant change in the eigenvalue(s) of the scattering problem [19,20]. In Sec. III, we show that relation (1.4) is satisfied for some realistic set of values of the fiber and pulse parameters, and so the soliton will not split into two components. Thus we are indeed justified to use our perturbation theory for the optical soliton. That is, one can consider the soliton to be a single entity, i.e., a bound state of its two components, instead of having to consider the dynamics of the two components as of separate entities [18,23,24].

It should be mentioned that, from a practical viewpoint, there is a very important physical situation where the results of our work *cannot* be applied. This is either when the initial pulse amplitude is not large enough to give rise to soliton formation, or else when the distance of propagation is sufficiently short for the nonlinearity and dispersion to affect the pulse evolution significantly. In both these cases, the pulse propagation is linear in the lowest order, with the nonlinearity playing the role of a first-order correction. It is this linear limit of pulse propagation for which much of experimental measurements and theoretical calculations have been done [16,25,26]; see also Ref. [14]. Thus it is important to discuss the relation between some of our results, obtained in Sec. IV, and the corresponding results which were earlier obtained for the linear limit. We will do this in Sec. IV.

The perturbation theory, which we developed in Sec. II, can also be useful in other situations, where the Manakov equation arises as the main-order evolution equation. In particular, this equation was shown [27] to describe pulse evolution in *nonrandom* birefringent fibers with a certain ellipticity of the eigenmodes. It was also recently reported [28] that the Manakov equation describes spatial solitons in $Al_xGa_{1-x}As$ planar waveguides. Finally, we note that this equation can also be derived [29] when one averages the equations of the nonlinear directional coupler when taking into account random variations of the linear coupling coefficient and phase velocity mismatch between the two cores.

The remainder of this paper is organized as follows. In Sec. II, we work out the details of the perturbation theory for the single soliton of the Manakov equation. The reader who is not interested in the mathematical details of the perturbation theory may just browse the beginning of Sec. II to familiarize himself or herself with the notations, and then go directly to the next section. In Sec. III, we give a brief derivation of Eqs. (1.1) in randomly birefringent fibers and specify the form of the perturbations R_u and R_v, as well as estimate the magnitude of the small parameter ϵ. In Sec. IV we apply the first-order perturbation theory developed in Sec. II to study the effects of the linear and nonlinear PMD's on the soliton. In this way, both the slow evolution of the soliton parameters and the continuous-wave radiation emitted by the soliton are calculated and compared, whenever possible, with previous findings. In Sec. V, we first calculate the second-order changes in the two most important soliton parameters: its width (which is related to the amplitude) and velocity (which is related to the mean frequency). Then, in the same section, we relate the results for the soliton timing jitter and the radiation generated due to the linear PMD, as obtained by the perturbation theory in Sec. IV, to the corresponding results obtained from the intuitive picture of the dynamics of the soliton components as interacting quasiparticles. We also show, by using some semiqualitative arguments, that the root-mean-square (rms) distance between the centers of the soliton's two components is proportional to the fourth root of the birefringence strength. In Sec. VI, we summarize the results obtained in this work.

II. PERTURBATION THEORY FOR THE MANAKOV EQUATION

The general form of the one-soliton solution of Eqs. (1.1) with $\epsilon=0$ is

$$\vec{u}_0 \equiv \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = 2\eta\, \mathrm{sech}\theta\, e^{i\Psi} e^{i\sigma_3\varphi} \hat{B} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{2.1}$$

where

$$\Psi = \alpha(z) - \frac{\xi}{\eta}\theta, \quad \theta = 2\eta[\tau - \tau_c(z)], \tag{2.2a}$$

$$\frac{d\alpha}{dz} = 4(\xi^2+\eta^2), \quad \frac{d\tau_c}{dz} = -4\xi, \tag{2.2b}$$

$$\hat{B} = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}, \tag{2.2c}$$

the Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and ξ, η, φ, and β are constants. When Eqs. (1.1) are perturbed with $\epsilon \neq 0$, we will take the solution to be of the form

$$\vec{u} = e^{i\Psi} e^{i\sigma_3\varphi} \hat{B}(\vec{u}_{00} + \epsilon \vec{u}_{10} + \cdots), \tag{2.3a}$$

where

$$\vec{u}_{00} = 2\eta\, \mathrm{sech}\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \begin{pmatrix} u_{00} \\ 0 \end{pmatrix}, \quad \vec{u}_{10} = \begin{pmatrix} u_{10} \\ v_{10} \end{pmatrix}. \tag{2.3b}$$

In expansion (2.3), one also needs to assume the parameters, which were constant for soliton (2.1), to be slowly varying functions of z:

$$\frac{d}{dz}\eta = \epsilon\dot{\eta}_1 + \epsilon^2\dot{\eta}_2 + \cdots, \quad \frac{d}{dz}\xi = \epsilon\dot{\xi}_1 + \epsilon^2\dot{\xi}_2 + \cdots,$$

$$\frac{d}{dz}\beta = \epsilon\dot{\beta}_1 + \epsilon^2\dot{\beta}_2 + \cdots, \quad \frac{d}{dz}\varphi = \epsilon\dot{\varphi}_1 + \epsilon^2\dot{\varphi}_2 + \cdots, \tag{2.4}$$

$$\frac{d}{dz}\alpha = \dot{\alpha}_0 + \epsilon\dot{\alpha}_1 + \cdots, \quad \frac{d}{dz}\tau_c = \dot{\tau}_{c0} + \epsilon\dot{\tau}_{c1} + \cdots,$$

where $\dot{\alpha}_0$ and $\dot{\tau}_{c0}$ are given by Eq. (2.2b), and $\dot{\eta}_1$, $\dot{\eta}_2$, $\dot{\xi}_1$, etc. are to be determined.

In what follows, it will be convenient for us to introduce the following notations:

$$\vec{\mathbf{w}}_{10} = \begin{pmatrix} u_{10} \\ u_{10}^* \\ v_{10} \\ v_{10}^* \end{pmatrix},$$

$$\vec{R} = \begin{pmatrix} R_u e^{-i\varphi}\cos\beta + R_v e^{i\varphi}\sin\beta \\ -(R_u e^{-i\varphi}\cos\beta + R_v e^{i\varphi}\sin\beta)^* \\ -R_u e^{-i\varphi}\sin\beta + R_v e^{i\varphi}\cos\beta \\ -(-R_u e^{-i\varphi}\sin\beta + R_v e^{i\varphi}\cos\beta)^* \end{pmatrix}, \tag{2.5a}$$

$$\hat{\sigma}_3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}, \quad \hat{\sigma}_1 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{pmatrix}, \tag{2.5b}$$

and

$$\hat{\phi}_1 = \frac{1}{2\eta}\begin{pmatrix} iu_{00} \\ -iu_{00} \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\phi}_2 = \frac{1}{2\eta}\begin{pmatrix} \partial_\theta u_{00} \\ \partial_\theta u_{00} \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\phi}_1^D = \frac{1}{2}\begin{pmatrix} \partial_\eta u_{00} \\ \partial_\eta u_{00} \\ 0 \\ 0 \end{pmatrix}, \tag{2.6}$$

$$\hat{\phi}_2^D=\frac{1}{2\eta}\begin{pmatrix}-i\theta u_{00}\\ i\theta u_{00}\\ 0\\ 0\end{pmatrix},\quad \hat{\phi}_3=\frac{1}{2\eta}\begin{pmatrix}0\\ 0\\ iu_{00}\\ -iu_{00}\end{pmatrix},\quad \hat{\phi}_4=\frac{1}{2\eta}\begin{pmatrix}0\\ 0\\ u_{00}\\ u_{00}\end{pmatrix}.$$

Then, substituting Eqs. (2.1)–(2.4) into Eqs. (1.1), and retaining only the first-order (in ϵ) terms, one obtains the following evolution equation for the vector $\vec{\mathbf{w}}_{10}$:

$$i\partial_t\vec{\mathbf{w}}_{10}+4\eta^2L_M\vec{\mathbf{w}}_{10}=e^{-i\hat{\sigma}_3\Psi}\vec{R}-2i\eta\Big\{(\dot{\alpha}_1-2\xi\dot{\tau}_{c1})\hat{\phi}_1+(-2\eta\dot{\tau}_{c1})\hat{\phi}_2+\dot{\varphi}_1(\hat{\phi}_1\cos2\beta-\hat{\phi}_3\sin2\beta)+\dot{\beta}_1\hat{\phi}_4+\frac{\dot{\eta}_1}{\eta}\hat{\phi}_1^D+\dot{\xi}_1\hat{\phi}_2^D\Big\}=0. \tag{2.7}$$

Above, L_M is the block-diagonal operator:

$$L_M=\begin{pmatrix}L_{\rm NLS} & 0\\ 0 & L_\perp\end{pmatrix}, \tag{2.8a}$$

where

$$L_{\rm NLS}=\sigma_3(\partial_{\theta\theta}-1)+2(2\sigma_3+i\sigma_2)\mathrm{sech}^2\theta, \tag{2.8b}$$

$$L_\perp=\sigma_3(\partial_{\theta\theta}-1+2\ \mathrm{sech}^2\theta). \tag{2.8c}$$

Since Eq. (2.7) is a linear equation for $\vec{\mathbf{w}}_{10}$, one may solve it by using separation of variables: $\vec{\mathbf{w}}_{10}(\theta,z)=\vec{\mathbf{w}}_{10}(\theta)e^{i\lambda z}$. This yields an eigenvalue problem for the operator L_M:

$$L_M\psi=\lambda\psi. \tag{2.9}$$

Due to the block-diagonal structure of L_M, we can independently solve each of the two matrix blocks in Eq. (2.9).

The operator $L_{\rm NLS}$ arises in the perturbation theory for the soliton of the NLS, and so its spectrum, as well as the eigenfunctions and their closure, are known (see, e.g., Ref. [8]). That is, one has

$$L_{\rm NLS}\psi^{\rm NLS}=(k^2+1)\psi^{\rm NLS}, \tag{2.10a}$$

$$\psi^{\rm NLS}=e^{ik\theta}\left[\left(1-\frac{2ik(1-\tanh\theta)}{(k+i)^2}\right)\begin{pmatrix}0\\1\end{pmatrix}+\frac{\mathrm{sech}^2\theta}{(k+i)^2}\begin{pmatrix}1\\1\end{pmatrix}\right], \tag{2.10b}$$

$$L_{\rm NLS}\overline{\psi}^{\rm NLS}=-(k^2+1)\overline{\psi}^{\rm NLS},\quad \overline{\psi}^{\rm NLS}=\sigma_1(\psi^{\rm NLS})^*, \tag{2.10c}$$

$$L_{\rm NLS}\hat{\phi}_{1,2}^{\rm NLS}=0,\quad L_{\rm NLS}(\hat{\phi}_{1,2}^D)^{\rm NLS}=-2i\hat{\phi}_{1,2}^{\rm NLS}, \tag{2.11}$$

where $\hat{\phi}_{1,2}^{\rm NLS}$ and $(\hat{\phi}_{1,2}^D)^{\rm NLS}$ are the two-component vectors whose components coincide with the first two components of the corresponding vectors defined in Eq. (2.6). Note that the vectors $\overline{\psi}^{\rm NLS}$, $\hat{\phi}_1^{\rm NLS}$, and $(\hat{\phi}_2^D)^{\rm NLS}$ are defined in a slightly different way than in Ref. [8]. It was also shown in Ref. [3] that the above eigenfunctions of $L_{\rm NLS}$ and the two associate functions $(\hat{\phi}_{1,2}^D)^{\rm NLS}$ form a basis for the expansion of an arbitrary two-component vector with sufficiently smooth and rapidly decaying components.

The second matrix block of the operator L_M, i.e., $L_\perp$, is a diagonal operator, and, moreover, each of its entries represents a well-known quantum-mechanical problem. Thus the solutions of

$$L_\perp\psi^\perp=(k^2+1)\psi^\perp\quad\text{and}\quad L_\perp\overline{\psi^\perp}=-(k^2+1)\overline{\psi^\perp} \tag{2.12a}$$

are the functions

$$\psi^\perp=\left(\frac{-ik+\tanh\theta}{-ik+1}\right)e^{ik\theta}\begin{pmatrix}0\\1\end{pmatrix},\quad \overline{\psi^\perp}=\sigma_1(\overline{\psi^\perp})^*. \tag{2.12b}$$

The operator $L_\perp$ also has two linearly independent bound states,

$$L_\perp\hat{\phi}_1^\perp=L_\perp\hat{\phi}_2^\perp=0, \tag{2.13}$$

where the components of $\hat{\phi}_{1,2}^\perp$ coincide, respectively, with the last two components of the vectors $\hat{\phi}_{3,4}$ in Eq. (2.6). Since the operator $L_\perp$ is Hermitian, then the eigenfunctions $\psi^\perp$, $\overline{\psi^\perp}$, and $\hat{\phi}_{1,2}^\perp$ form a basis in the space of two-component vectors.

Thus an arbitrary four-component vector $\vec{\mathbf{w}}_{10}$ can be expanded over the eigenfunctions and the associate functions of the operator L_M as follows:

$$\vec{\mathbf{w}}_{10}(\theta,z)=\int_{-\infty}^{\infty}[g_1(k,z)\psi_1+\overline{g_1}(k,z)\overline{\psi_1}+g_3(k,z)\psi_3+\overline{g_3}(k,z)\overline{\psi_3}]dk+\sum_{n=1}^{4}h_n(z)\hat{\phi}_n+[h_1^D(z)\hat{\phi}_1^D+h_2^D(z)\hat{\phi}_2^D]. \tag{2.14}$$

In this equation, the four-component eigenfunctions $\psi_{1,3}$ and $\overline{\psi}_{1,3}$ are

$$\psi_1=\begin{pmatrix}\psi^{\rm NLS}\\ 0\end{pmatrix},\quad \psi_3=\begin{pmatrix}0\\ \psi^\perp\end{pmatrix},\quad \overline{\psi_{1,3}}=\hat{\sigma}_1\psi_{1,3}^*, \tag{2.15}$$

where $\psi^{\rm NLS}$ and $\psi^\perp$ are given by Eqs. (2.10b) and (2.12b), and the exponentially localized functions $\hat{\phi}_1$ through $\hat{\phi}_2^D$ are given by Eq. (2.6). Also, $g_{1,3}$, $\overline{g}_{1,3}$, and h_1 through h_2^D are the expansion coefficients. In order to find these coefficients from Eq. (2.7), one needs to know the solution of the adjoint problem for a row vector ψ^A which is defined so as to satisfy the equation

$$\psi^A\overleftarrow{L}_M=\lambda\psi^A, \tag{2.16}$$

where the left-pointing arrow over L_M indicates the direction of the differentiation ∂_θ. Using the fact that the operator $\hat{\sigma}_3L_M$ is self-adjoint, one can show (see, e.g., Refs. [8,30]) that the vector $\psi^\dagger(k,\theta)\hat{\sigma}_3$, where ψ is a solution of Eq. (2.9) and the superscript ''†'' denotes the Hermitian conjugation, satisfies Eq. (2.16). Then, using the Wronskian relation for the operator L_M, one obtains, in a standard way, the following inner products:

$$\langle \psi_1^\dagger(k')|\hat{\sigma}_3|\psi_1(k)\rangle = \langle \psi_3^\dagger(k')|\hat{\sigma}_3|\psi_3(k)\rangle = -2\pi\delta(k-k'), \tag{2.17a}$$

$$\langle \overline{\psi}^\dagger(k')|\hat{\sigma}_3|\overline{\psi}(k)\rangle = \langle \overline{\psi_3^\dagger}(k')|\hat{\sigma}_3|\overline{\psi_3}(k)\rangle = 2\pi\delta(k-k'), \tag{2.17b}$$

where the inner product is defined by

$$\langle q|m|p\rangle = \int_{-\infty}^{\infty} d\theta\, q(\theta)m(\theta)p(\theta).$$

The inner products among the localized eigenfunctions and their adjoints can be found straightforwardly:

$$\langle(\hat{\phi}_1^D)^\dagger|\hat{\sigma}_3|\hat{\phi}_1\rangle = -\langle \hat{\phi}_1^\dagger|\hat{\sigma}_3|\hat{\phi}_1^D\rangle = 2i, \tag{2.18a}$$

$$\langle(\hat{\phi}_2^D)^\dagger|\hat{\sigma}_3|\hat{\phi}_2\rangle = -\langle \hat{\phi}_2^\dagger|\hat{\sigma}_3|\hat{\phi}_2^D\rangle = -2i, \tag{2.18b}$$

$$\langle \hat{\phi}_4^\dagger|\hat{\sigma}_3|\hat{\phi}_3\rangle = -\langle \hat{\phi}_3^\dagger|\hat{\sigma}_3|\hat{\phi}_4\rangle = 4i. \tag{2.18c}$$

All the other inner products are equal to zero.

Now, using the identity

$$\hat{\sigma}_1\vec{\mathbf{w}}_{10}^* = \vec{\mathbf{w}}_{10}, \tag{2.19}$$

and also the relation $\overline{\psi_{1,3}} = \hat{\sigma}_1\psi_{1,3}^*$ [see Eq. (2.15)] and the explicit form of the localized states $\hat{\phi}_1$ through $\hat{\phi}_2^D$, one can easily obtain the following symmetry relations for the expansion coefficients:

$$\overline{g}_{1,3}(k) = g_{1,3}^*(k), \tag{2.20a}$$

$$h_n = h_n^*, \quad n=1,\dots,4, \quad h_{1,2}^D = (h_{1,2}^D)^*. \tag{2.20b}$$

Finally, substituting expansion (2.14) into Eq. (2.7) and using the inner products Eqs. (2.17) and (2.18), we obtain

$$ih_{1,z} - 8i\eta^2 h_1^D + 2i\eta(\dot{\alpha}_1 - 2\xi\dot{\tau}_{c1} + \dot{\varphi}_1\cos 2\beta) = \frac{1}{2i}\langle(\phi_1^D)^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21a}$$

$$ih_{2,z} - 8i\eta^2 h_2^D + 2i\eta(-2\eta\dot{\tau}_{c1}) = -\frac{1}{2i}\langle(\phi_2^D)^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21b}$$

$$ih_{3,z} + 2i\eta(-\dot{\varphi}_1\sin 2\beta) = \frac{1}{4i}\langle \hat{\phi}_4^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21c}$$

$$ih_{4,z} + 2i\eta\dot{\beta}_1 = -\frac{1}{4i}\langle \hat{\phi}_3^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21d}$$

$$ih_{1,z}^D + 2i\eta\dot{\eta}_1 = -\frac{1}{2i}\langle \hat{\phi}_1^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21e}$$

$$ih_{2,z}^D + 2i\eta\dot{\xi}_1 = \frac{1}{2i}\langle \hat{\phi}_2^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.21f}$$

$$i\partial_z g_{1,3} + 4\eta^2(k^2+1)g_{1,3} = -\frac{1}{2\pi}\langle \psi_{1,3}^\dagger(k)|\hat{\sigma}_3|\vec{R}\rangle. \tag{2.21g}$$

(Above, $h_{1,z}$, etc. stands for dh_1/dz, etc., respectively, while the notation $g_{1,3}$ means either g_1 or g_3.) Equations for $\overline{g}_{1,3}$ follow then from Eqs. (2.21g) and (2.20a). Note also that Eqs. (2.21a)–(2.21f) are consistent with the symmetry (2.20b).

The standard requirement that the expansion coefficients do not grow secularly with z leads to the following evolution equations for the soliton parameters [cf. Eq. (2.4)]:

$$\dot{\eta}_1 = \frac{1}{4\eta}\langle \hat{\phi}_1^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.22a}$$

$$\dot{\xi}_1 = -\frac{1}{4\eta}\langle \hat{\phi}_2^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.22b}$$

$$\dot{\alpha}_1 - 2\xi\dot{\tau}_{c1} + \dot{\varphi}_1\cos 2\beta = -\frac{1}{4\eta}\langle(\hat{\phi}_1^D)^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.22c}$$

$$\dot{\tau}_{c1} = -\frac{1}{4\eta^2}\langle(\hat{\phi}_2^D)^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.22d}$$

$$\dot{\varphi}_1\sin 2\beta = \frac{1}{8\eta}\langle \hat{\phi}_4^\dagger|\hat{\sigma}_3|\vec{R}\rangle, \tag{2.22e}$$

$$\dot{\beta}_1 = \frac{1}{8\eta}\langle \hat{\phi}_3^\dagger|\hat{\sigma}_3|\vec{R}\rangle. \tag{2.22f}$$

Equations (2.22), (2.21g), and (2.14) constitute the complete first-order perturbation theory for the one-soliton solution of the Manakov equation. Note that Eqs. (2.22) guarantee that if h_1 through h_2^D in Eq. (2.14) are zero at $z=0$, then they will remain zero for all z.

III. ORIGIN OF THE MANAKOV EQUATION IN RANDOMLY BIREFRINGENT FIBERS

In this section, we will follow the main steps of Ref. [14], and give the derivation of the Manakov equation in randomly birefringent fibers. We will also estimate the size of the perturbations to this equation due to linear and nonlinear PMD.

The equation describing the evolution of a pulse in a fiber with linearly polarized, orthogonal eigenmodes is (see, e.g., Ref. [14])

$$i\vec{A}_z + K\vec{A} + i\Delta\vec{A}_\tau + \vec{A}_{\tau\tau} + \frac{9}{4}\vec{N} = \vec{0}, \tag{3.1}$$

where vector $\vec{A} = (A_1, A_2)^T$ denotes the envelopes of the electric field in the two eigenmodes, and

$$\vec{N} = \begin{pmatrix} (|A_1|^2 + \frac{2}{3}|A_2|^2)A_1 + \frac{1}{3}A_2^2 A_1^* \\ (|A_2|^2 + \frac{2}{3}|A_1|^2)A_2 + \frac{1}{3}A_1^2 A_2^* \end{pmatrix}. \tag{3.2}$$

All the variables in Eq. (3.1) are appropriately normalized, with z and τ denoting the distance along the fiber and the retarded time in the pulse's reference frame, respectively. The coefficient $(\frac{9}{4})$ in front of the nonlinear term $\vec{N}$ in Eq. (3.1) is chosen in order to have the resultant Manakov equation exactly in the form of Eq. (1.1). The most general form of the matrix K in Eq. (3.1) is

$$K=\kappa_1\sigma_1+\kappa_2\sigma_2+\kappa_3\sigma_3, \tag{3.3}$$

where κ_1, κ_2, and κ_3 are real, independent random functions of z [25,14,31], and σ_1, σ_2, and σ_3 are Pauli matrices, introduced before. The coefficients κ_1, κ_2, and κ_3, which vanish if the fiber has a perfectly circularly symmetric cross section, acquire nonzero values when that circular symmetry is broken due to unavoidable imperfections in the manufacturing or installation of the fiber, or due to environmental perturbations. The coefficients κ_1 and κ_3 correspond to the random birefringence introduced by the geometric or stress factors, and the coefficient κ_2 corresponds to random microtwists of the fiber around its longitudinal axis [25,27]. For a pair of orthogonal eigenmodes of the fiber which at some $z=z_0$ are aligned along some fixed axes in the cross-section plane, the presence of the term $K\vec{A}$ in the evolution Eq. (3.1) causes linear coupling between these modes, as well as an accumulation of a mismatch between their phases, as z increases along the fiber. The coefficients $\kappa_1(z)$ and $\kappa_3(z)$ also depend on the center frequency ω_0 of the carrier wave, whereas $\kappa_2(z)$ obviously does not. Consequently, the matrix Δ can be shown to have the form [27,31]

$$\Delta=\left(\frac{\partial}{\partial\omega}K\right)\bigg|_{\omega=\omega_0}\equiv\kappa_1'\sigma_1+\kappa_3'\sigma_3, \tag{3.4}$$

where $\kappa_{1,3}'=(\partial\kappa_{1,3}/\partial\omega)|_{\omega=\omega_0}$.

There are several important length scales in Eq. (3.1), which we shall now discuss. The distance z and the amplitude of $\vec{A}$ in Eq. (3.1) are normalized in such a way that the dispersion length l_{sol}, which is also on the order of magnitude of the nonlinear length, is unity. (This distance is also sometimes called the soliton period.) With such a normalization, the magnitudes of the fourth and fifth terms in Eq. (3.1) are of order 1, whereas the magnitude of the term $K\vec{A}$ is of the order $(l_{\text{sol}}/l_{\text{beat}})\gg 1$. Here $l_{\text{beat}}\sim\langle\kappa_1^2+\kappa_2^2+\kappa_3^2\rangle^{-1/2}$ is the average beat length between the two linear eigenmodes of the fiber due to the latter's birefringence, and the notation $\langle\cdots\rangle$ stands for the average over an ensemble of fibers or over the distance along a single fiber [16]. In fact, in present-day communication fibers the relative difference $(\Delta n/n)$ between the two refractive indices of the fiber is typically on the order of 10^{-7} [14]. Then, given that the carrier wavelength is $\lambda_0=1.5\ \mu$m, one has $l_{\text{beat}}\lesssim 10$ m. For $l_{\text{sol}}\sim 100$ km, which is of the right order of magnitude for a 20-ps pulse in a dispersion-shifted fiber, one has $\|K\vec{A}\|\sim(l_{\text{sol}}/l_{\text{beat}})\sim 10^4$. Here $\|\cdots\|$ denotes the magnitude of a vector. As we shall demonstrate below, in Eq. (3.1) the magnitude of the third term is of order 1, and thus it is the term $K\vec{A}$ that determines the pulse evolution on the distances that are much shorter than the nonlinear and dispersion lengths. Therefore, we need to analyze that term before we can proceed with the analysis of the other terms.

One can eliminate the fast variations, caused by the random birefringence, of the direction and phases of the field vector $\vec{A}$ by means of the following transformation:

$$\vec{A}=M(z)\vec{u},\quad M(z)=\begin{pmatrix} m_1 & -m_2^* \\ m_2 & m_1^* \end{pmatrix}, \tag{3.5}$$

where $\vec{u}$ is the "slow" field vector and the matrix $M(z)$ satisfies

$$iM_z+KM=0, \tag{3.6}$$

and is unitary:

$$|m_1|^2+|m_2|^2=1. \tag{3.7}$$

Although it would be desirable to carry out the calculations with a most general form of the matrix K, i.e., assuming $\kappa_{1,2,3}$ to be independent random variables, this could only be done [31] in the limit when the correlation length l_{cor} of $\kappa_{1,2,3}$ is *much less* than the beat length l_{beat}. In this limit, it can be shown that the slow evolution of the pulse envelope is also governed by the Manakov equation. However, in communication fibers one has $l_{\text{cor}}\sim 10\ldots 100$ m and $l_{\text{beat}}\lesssim 10$ m [14,15], so it is the opposite limit, namely, $l_{\text{cor}}\gg l_{\text{beat}}$, that applies. Therefore, in what follows we will have to consider a simplified model of random birefringence, originally introduced in Ref. [16] and further studied in Refs. [32,14]. We will now describe the details of that model.

First, it was shown in Ref. [32] by numerical simulations that even if $\kappa_2\equiv 0$ in Eq. (3.3) and $\kappa_{1,3}$ are independent random variables driven by (independent) white-noise processes $w_{1,3}(z)$, so that

$$\frac{d\kappa_n}{dz}=-\frac{\kappa_n}{l_{\text{cor}}}+w_n,\quad n=1,3,$$
$$\langle w_n\rangle=0,\quad \langle w_n(z_1)w_m(z_2)\rangle=2\frac{D_m}{l_{\text{cor}}}\delta_{nm}\delta(z_1-z_2), \tag{3.8}$$

the direction of the Stokes vector of the electric field in the pulse is, after a sufficiently long distance of propagation, uniformly randomized over the Poincaré sphere. As a consequence, the pulse evolution will be governed by the Manakov equation even in this, somewhat reduced, model of the random birefringence [33]. Furthermore, it was shown [32,14] that the above conclusion about the Stokes vector still holds, and hence the Manakov equation results as the main-order evolution equation, even if one further restricts κ_1 and κ_3 to be

$$\kappa_1=b(\omega)\sin\Phi(z),\quad \kappa_3=b(\omega)\cos\Phi(z),$$
$$\frac{d\Phi}{dz}=w_\Phi(z),\quad \langle w_\Phi(z_1)w_\Phi(z_2)\rangle=2D_\Phi\delta(z_1-z_2). \tag{3.9}$$

In Ref. [14], birefringence models (3.8) and (3.9) were called the second and first models, respectively. Thus the first model, which assumes that the birefringence has a fixed total strength $b(\omega)$, yields essentially the same results [32] as the more general second birefringence model. An intuitive explanation of this is as follows. It is the random character of the birefringence angle, rather than that of its strength, that is essential for the uniform ''smearing'' of the Stokes vector over the Poincaré sphere. Indeed, if one considers a model where the birefringence angle [Φ in Eq. (3.9)] is fixed and the birefringence strength b varies randomly, then it is easy to show that the corresponding Eqs. (3.6) can be solved exactly. In particular, it can be shown that the direction of the Stokes vector will never be uniformly randomized over the Poincaré sphere.

It was shown in Ref. [14] that, for the first birefringence model, the average of any polynomial involving the quantities S_n [to be defined in Eqs. (3.11) below] and $\cos N\Phi$ or $\sin N\Phi$ for any N can be explicitly computed, with no restriction being put on the ratio $(l_{\rm cor}/l_{\rm beat})$. On the other hand, the same for the second birefringence model can be done exactly only in the limit $l_{\rm cor} \ll l_{\rm beat}$ (cf. Appendix A). These averages turn out to be important in calculations in Sec. IV, and therefore we will mostly use the results for the first birefringence model in this work (except for the end of Sec. V and Appendix B). According to what was said in the preceding paragraph, restricting our considerations to a specific model of random birefringence does not appear to limit the validity of our results.

Now we will list the most important physical lengths in terms of the parameters of the first birefringence model:

$$l_{\rm cor}=\frac{1}{D_\Phi}, \quad l_{\rm beat}=\frac{2\pi}{b}, \tag{3.10a}$$

$$l_{\rm diffusion}\approx l_{\rm cor} \quad (\text{for } l_{\rm cor}\gg l_{\rm beat}). \tag{3.10b}$$

Here $l_{\rm diffusion}$ is the distance over which the field, measured with respect to the local axes of birefringence, loses memory of its initial orientation relative to those axes. As was shown in Refs. [32,14], $l_{\rm diffusion}$ is also of the order of magnitude of the distance over which the entries of the matrix M, defined in Eq. (3.5), may change significantly. Next, in what follows it will be convenient to use the following sets of notations:

$$S_1=|m_1|^2-|m_2|^2, \quad S_4=2m_1m_2,$$
$$S_2=m_1^*m_2+m_1m_2^*, \quad S_5=m_2^2-m_1^2, \tag{3.11a}$$
$$S_3=i(m_1^*m_2-m_1m_2^*), \quad S_6=i(m_1^2+m_2^2),$$

$$\tilde{S}_{n+1}=S_{n+1}\cos\Phi+S_{n+2}\sin\Phi,$$
$$\tilde{S}_{n+2}=-S_{n+1}\sin\Phi+S_{n+2}\cos\Phi, \tag{3.11b}$$
$$\tilde{S}_{n+3}=S_{n+3}, \quad n=0,3.$$

Note that, when in Eq. (3.5) the vector $\vec{u}=(1,0)^T$, then $\vec{S}=(S_1,S_2,S_3)^T$ is the Stokes vector. Moreover, we will require an explicit form of the matrix Δ, defined in Eq. (3.4). Following Refs. [25,16,14], we take

$$\kappa_1'=b'\sin\Phi, \quad \kappa_3'=b'\cos\Phi, \quad b'\equiv\left.\frac{db}{d\omega}\right|_{\omega=\omega_0}, \tag{3.12}$$

with b' being independent of z. The assumption that the angle Φ does not depend on ω, that has allowed us to obtain Eq. (3.12) from Eq. (3.9), is based on the experimental fact that in fibers, the group and phase velocities coincide within about 10% (see, e.g., Ref. [16]).

Now an evolution equation for the vector $\vec{u}=(u,v)^T$ follows from Eqs. (3.1), (3.5), and (3.6). It is

$$i\vec{u}_z+\vec{u}_{\tau\tau}+2(\vec{u}^\dagger\vec{u})\vec{u}=-\{i\Omega\vec{u}_\tau+[\tfrac{9}{4}M^{-1}\vec{N}-2(\vec{u}^\dagger\vec{u})\vec{u}]\}, \tag{3.13}$$

where $\vec{u}^\dagger=(u^*,v^*)$, and

$$\Omega=M^{-1}\Delta M=\kappa_1'\begin{pmatrix} S_2 & -S_5^* \\ -S_5 & -S_2\end{pmatrix}+\kappa_3'\begin{pmatrix} S_1 & -S_4^* \\ -S_4 & -S_1\end{pmatrix}. \tag{3.14}$$

For the first birefringence model, the form of Ω can be simplified:

$$\Omega=b'\begin{pmatrix} \tilde{S}_1 & -\tilde{S}_4^* \\ -\tilde{S}_4 & -\tilde{S}_1\end{pmatrix}, \tag{3.15}$$

where we used Eqs. (3.12) and (3.11b). The left-hand side of Eq. (3.13) is the Manakov equation and right-hand side is a random perturbation with a zero mean value [14].

Let us now estimate the size of the first term on the right-hand side of Eq. (3.13), which represents the linear PMD. The magnitude of the entries of the matrix Ω is on the order of $O(\kappa_1'+\kappa_3')=O[(\kappa_1+\kappa_3)/\omega_0]$, where ω_0 is the carrier frequency. Next, the magnitude of $\vec{u}_\tau$ is of the order $O(\|\vec{u}\|/T_p)$, where T_p is the pulse width. Therefore, the size of the linear PMD term in Eq. (3.1) is $(1/\omega_0T_p)$ times the size of the term $K\vec{A}$. For a 20-ps pulse, $(\omega_0T_p)\sim10^4$, and thus $\|\Omega\vec{u}_\tau\|\sim(l_{\rm sol}/l_{\rm beat})\times10^{-4}\gtrsim1$ in Eq. (3.13). At first sight, this seems to be not a small perturbation, since the size of all the terms on the left-hand side of Eq. (3.7) is also $O(1)$. However, as will be clear from the results of Sec. IV, it is *not* the rms value of $\Omega\vec{u}_\tau$, i.e., $\sqrt{\langle\|\Omega\vec{u}_\tau\|^2\rangle}$, but rather the integral of its autocorrelation function (i.e., the intensity) that will determine the effect of that term on the soliton. The autocorrelation function of each of the components of the vector $\Omega\vec{u}_\tau$ can be approximately given by the following expression:

$$\langle a(z)a(z_1)\rangle=\left(\frac{D}{L}\right)\exp\left\{-\frac{|z-z_1|}{L}\right\}, \tag{3.16}$$

where $a(z)$ denotes either of the components of $\Omega\vec{u}_\tau$, and D and L are its intensity and correlation length, respectively. Note that

$$\int_{-\infty}^{\infty}\langle a(z)a(z+z_1)\rangle dz_1=2D.$$

From Eq. (3.16), $(D_{\rm lin}/L_{\rm lin}) = \langle \|\Omega \vec{u}_\tau\|^2 \rangle \sim 1\ldots10$ in units of Eq. (3.13), where the subscript ''lin'' stands for the linear PMD. $L_{\rm lin}$ is the length scale of the most rapidly changing term of the matrix Ω, i.e., $L_{\rm lin} = \min(l_{\rm cor}, l_{\rm diffusion})$. As we mentioned earlier, in communication-grade fibers one has $l_{\rm cor} \sim 10...100$ m, $l_{\rm beat} \lesssim 10$ m, and so $l_{\rm cor} > l_{\rm beat}$. Then, from Eq. (3.10b), $L_{\rm lin} \sim l_{\rm cor}$, and therefore

$$D_{\rm lin} \sim \frac{l_{\rm cor}}{l_{\rm sol}} \cdots 10\, \frac{l_{\rm cor}}{l_{\rm sol}} \sim 10^{-4} \ldots 10^{-2}. \tag{3.17}$$

Note that $D_{\rm lin}$ is the analog of the quantity $(\delta^2 l_{\rm cor})$, introduced in Sec. I. Then from Eq. (3.17) one can see that the criterion (1.4) of nonsplitting of the soliton by the random birefringence is almost certainly satisfied in real fibers, and consequently the effect of the linear PMD on the soliton can be treated as a small perturbation.

Similarly, we assume the autocorrelation function of the components of the nonlinear PMD vector, represented by the second group of terms on the right-hand side of Eq. (3.13), to be of the form (3.16). As shown in Ref. [14] for the first birefringence model, the corresponding intensity

$$D_{\rm nl} < 10^{-1} D_{\rm lin} \tag{3.18}$$

when $l_{\rm cor} > l_{\rm beat}$. Thus the effect of the nonlinear PMD on a $\sim$20-ps soliton is to be much weaker than that of the linear PMD. The corresponding scale $L_{\rm nl}$ is now the distance over which various quadratic combinations of the quantities S_1 through S_6, defined in Eq. (3.11), may change considerably [14]. Since the magnitude of the nonlinear PMD terms in Eq. (3.13) is of order 1, then from Eq. (3.18) it follows that $L_{\rm nl} < l_{\rm cor} \sim L_{\rm lin}$, which also agrees with the analytical results of Ref. [14].

Finally, since the above estimates for the coefficients $D_{\rm lin}$ and $D_{\rm nl}$ pertain to the particular choice of the pulse and fiber parameters, we will give here the general formulas from which these coefficients can always be estimated:

$$D_{\rm lin} \sim \left(\frac{l_{\rm sol}}{l_{\rm beat}(\omega_0 T_p)} \right)^2 \frac{l_{\rm cor}}{l_{\rm sol}}, \quad D_{\rm nl} \sim \frac{L_{\rm nl}}{l_{\rm sol}}. \tag{3.19}$$

Since $l_{\rm sol}$ scales as T_p^2, $D_{\rm lin}$ does not depend on T_p, whereas $D_{\rm nl} \sim T_p^{-2}$.

IV. EFFECT OF LINEAR AND NONLINEAR PMD ON A SOLITON

In this section, we will use the machinery developed in Sec. II to calculate the effect of linear and nonlinear PMD's on the soliton, and also to calculate the amount of radiation which is generated by these perturbations. To this end, we will first use Eq. (3.13) to separate explicitly the linear and nonlinear PMD contributions to the perturbation terms R_u and R_v in Eqs. (1.1):

$$\epsilon \begin{pmatrix} R_u \\ R_v \end{pmatrix} \equiv \epsilon \begin{pmatrix} R_u \\ R_v \end{pmatrix}_{\rm lin} + \epsilon \begin{pmatrix} R_u \\ R_v \end{pmatrix}_{\rm nl}, \tag{4.1}$$

where

$$\epsilon \begin{pmatrix} R_u \\ R_v \end{pmatrix}_{\rm lin} = -i\Omega \vec{u}_\tau, \tag{4.2a}$$

$$\epsilon \begin{pmatrix} R_u \\ R_v \end{pmatrix}_{\rm nl} = -[\tfrac{9}{4} M^{-1} \vec{N} - 2(\vec{u}^\dagger \vec{u})\vec{u}], \tag{4.2b}$$

and M, N, and Ω were defined in Eqs. (3.6), (3.2), and (3.14), respectively. For the first-order calculations, it is adequate to take $\vec{u} = \vec{u}_0$, where $\vec{u}_0$ was defined in Eq. (2.1). From the estimates made in Sec. III, it follows that the magnitude of the small parameter ϵ is on the order of $\sqrt{D_{\rm lin}} \sim 10^{-1} \ldots 10^{-2}$, and $\sqrt{D_{\rm nl}} < \sqrt{D_{\rm lin}}$. In this section, we will not assign a specific value to ϵ, but rather will simply use the vector $\epsilon\vec{R}$ in place of $\vec{R}$ when calculating the evolution of the soliton parameters and the radiation from Eqs. (2.21) and (2.22).

Let us now determine the explicit form of the perturbation vectors $\epsilon\vec{R}_{\rm lin}$ and $\epsilon\vec{R}_{\rm nl}$. These vectors are defined following the general definition (2.5a) of the vector $\vec{R}$, and Eq. (4.1), which separates the perturbation into the linear PMD and nonlinear PMD parts. Substituting $\vec{u}_0$ from Eqs. (2.1) and (2.2), and Ω from Eq. (3.15) into Eq. (4.2a), and then substituting the result in Eq. (2.5a), one finds

$$\epsilon \vec{R}_{\rm lin} = \begin{pmatrix} R_{\rm lin}^{(1)} \\ -R_{\rm lin}^{(1)*} \\ R_{\rm lin}^{(2)} \\ -R_{\rm lin}^{(2)*} \end{pmatrix}, \tag{4.3}$$

where

$$R_{\rm lin}^{(n)} = -2i\eta Q_n(z) \left(-\frac{i\xi}{\eta} u_{00} + \partial_\theta u_{00} \right), \quad n = 1,2, \tag{4.4}$$

$$Q_1(z) = b' \left(\tilde{S}_1 \cos 2\beta - (\tilde{S}_4 e^{2i\varphi} + \text{ c.c.}) \frac{\sin 2\beta}{2} \right), \tag{4.5a}$$

$$Q_2(z) = -b' \left(\tilde{S}_1 \sin 2\beta + (\tilde{S}_4 e^{2i\varphi} + \text{ c.c.}) \frac{\cos 2\beta}{2} + \tfrac{1}{2} (\tilde{S}_4 e^{2i\varphi} - \text{ c.c.}) \right), \tag{4.5b}$$

and u_{00} is defined in Eq. (2.3b). Recall that in deriving Eq. (4.4), we used the first model of random birefringence, discussed in Sec. III.

Similarly, substituting into Eq. (4.2b) the expressions (2.1) and (2.2) for $\vec{u}_0$, Eq. (3.4) for matrix M, and Eq. (3.2) for vector $\vec{N}$, one obtains that $\epsilon\vec{R}_{\rm nl}$ is of form (4.3), with the subscript ''lin'' being replaced by ''nl.'' It is convenient to separate the θ and z dependence in $R_{\rm nl}^{(1,2)}$, as follows:

$$R_{\rm nl}^{(1,2)} = -u_{00}^3(\theta) n_{1,2}(z) = -8\eta^3 \operatorname{sech}^3\theta\, n_{1,2}(z). \tag{4.6}$$

Here $n_{1,2}(z)$ are some polynomial functions of $m_1(z)$ and $m_2(z)$ and their complex conjugates, as well as of $e^{\pm i\varphi}$ and

$\cos\beta$ or $\sin\beta$. The maximum degree of these polynomials in $m_{1,2}(z)$ and $m^*_{1,2}(z)$ is 4, as it is also for both $e^{i\varphi}$ and $e^{-i\varphi}$, and also for $\cos\beta$ or $\sin\beta$. The explicit forms of $n_{1,2}(z)$ are not given here, because they are extremely cumbersome, and, moreover, later on we will only need the *average* values of $n_{1,2}(z)$. However, for illustrative purpose only, we will write down what a typical term in $n_{1,2}$ looks like:

$$m_1^3 m_2^* \cos\beta \, \sin\beta e^{2i\varphi}.$$

Moreover, one can show that $n_1(z)$ is always real, whereas $n_2(z)$ is, in general, complex. The average values of $n_{1,2}$ could be computed, if needed, by using any symbolic calculations package; however, we will not need to do even that in this work.

Before we go on and apply the results of Sec. II to our particular problem, let us mention the following. The perturbed Manakov equation (1.1) in which the perturbation is given *only* by the term of the form (4.2a), and where, in addition, either κ_1' or κ_3' vanishes and the remaining coefficient does not depend on z, can be reduced to the unperturbed Manakov equation by a simple phase transformation (see, e.g., Ref. [19]). We used this reduction as a check and verified, whenever possible, that our perturbation results, obtained below, were consistent with those obtained in the case of such a reduction. (However, we will not present the corresponding details here.) Having mentioned this point now, we will not be mentioning it again.

A. Evolution of the soliton parameters

We will first evaluate the effect of linear and nonlinear PMD's on the soliton, and then turn to the generation of radiation by the soliton due to these sources. Inserting Eqs. (4.3)–(4.6) into Eqs. (2.22), we arrive at the following system of the first-order evolution equations for the soliton parameters:

$$\dot{\xi}_1 = \dot{\eta}_1 = 0, \tag{4.7a}$$

$$\dot{\tau}_{c1} = Q_1, \tag{4.7b}$$

$$\dot{\alpha}_1 + \dot{\varphi}_1 \cos 2\beta = 4\eta^2 n_1, \tag{4.7c}$$

$$\dot{\beta}_1 = \{-2\xi \, \mathrm{Im} Q_2\} + \{-\tfrac{8}{3}\eta^2 \mathrm{Im} n_2\}, \tag{4.7d}$$

$$\dot{\varphi}_1 \sin 2\beta = \{-2\xi \, \mathrm{Re} Q_2\} + \{-\tfrac{8}{3}\eta^2 \mathrm{Re} n_2\}, \tag{4.7e}$$

where the quantities with the overdot were defined in Eq. (2.4). In the last two equations we used the curly brackets to visually separate the contributions from the linear and nonlinear PMD's.

First, we see that the PMD does not lead, in the first order, to any changes in the soliton's mean frequency, mean velocity, amplitude, and width. Then, the most important effect for the applications in optical telecommunications is the random change in the soliton's center position, because it results in a jitter in the soliton's arrival time at the receiving end of the transmission line. Here by the timing jitter we mean the one that occurs either to similarly polarized pulses propagating in different fibers, or to differently polarized pulses propagating in the same fiber. The mean-square value of this timing jitter is

$$\langle \tau_c^2(z) \rangle = \int_0^z dz_1 \int_0^z dz_2 \langle Q_1(z_1) Q_1(z_2) \rangle. \tag{4.8}$$

Since the value of z (the upper limit of integration) in Eq. (4.8) must be on the order of at least several dispersion lengths l_{sol} in order to be of interest for applications, and since $l_{\mathrm{sol}} \gg L_{\mathrm{lin}}$, then Eq. (4.8) can be approximately rewritten as

$$\langle \tau_c^2(z) \rangle = \int_0^z dz_1 \int_{-\infty}^{\infty} ds \langle Q_1(z_1) Q_1(z_1 + s) \rangle, \tag{4.9}$$

with the relative error of this approximation being of order $L_{\mathrm{lin}}/l_{\mathrm{sol}} < 10^{-3}$. With the same accuracy, one can set $z_1 \gg L_{\mathrm{lin}}$ on the right-hand side of Eq. (4.9). Then the calculations for the first birefringence model yield (see Appendix A) the following result for the timing jitter of the soliton:

$$\langle \tau_c^2(z) \rangle = \frac{2}{3} \frac{b'^2}{D_\Phi} z. \tag{4.10}$$

We will now make three remarks about Eq. (4.10). First, as stated in Sec. I, we need to discuss how Eq. (4.10) is related to the corresponding result in the linear limit of the pulse evolution. In that limit and for a given stretch of fiber, Poole and Wagner showed [34] that there exist two mutually orthogonal directions of polarization of an input pulse, for which the corresponding directions of the output pulse do not, in the first order, depend on the frequency of the signal. These special directions were called in Ref. [34] the principal states of polarization (PSP's). Any sufficiently narrow-band pulse at the input can be decomposed into a vector sum of the two PSP's, and then at the output, the distortion of the pulse's shape will arise as a result of a differential delay time, τ_d, between the PSP's [25]. Thus it is the relation between $\langle \tau_d^2 \rangle$ and $\langle \tau_c^2 \rangle$ in Eq. (4.10) that we will now discuss. By the definition of τ_d, this quantity for a given stretch of fiber is the maximum timing jitter of the output pulse, where the direction of polarization of the input pulse can take on any value between 0 and π. Then $\sqrt{\langle \tau_d^2 \rangle}$ is the rms of that timing jitter averaged over the ensemble of fibers. Using now the obvious fact that the orientations of the PSP's in any, sufficiently large, ensemble of fibers must be uniformly distributed in the interval $[0, 2\pi]$, one can see that $\sqrt{\langle \tau_d^2 \rangle}$ is also the average jitter in the pulse's arrival time, when the polarization of the input pulse is arbitrary. But this is precisely what $\sqrt{\langle (2\tau_c)^2 \rangle}$ is. (Note: The factor 2 in front of τ_c has occurred because τ_c is measured relative to the average arrival time, whereas τ_d is the arrival time difference between the slowest and the fastest pulses.) Moreover, from Eqs. (4.7c) and (4.5a) one can see that the difference in the arrival times for the *orthogonally* polarized solitons (with $\beta = 0$ and $\beta = \pi/2$) is also the maximum; this observation reinforces the statement that $(2\tau_c)$ for solitons is a counterpart of τ_d for linear pulses. Let us emphasize, however, that the effects of the (linear) PMD on a *linear pulse* and on a *soliton* are qualitatively different. The former will be, in general, distorted and spread out at the output (the spreading being proportional to τ_d), while the latter will preserve its shape, although its center may be shifted.

Thus having established the equivalence of the quantities $\sqrt{\langle \tau_d^2 \rangle}$ and $\sqrt{\langle 4\tau_c^2 \rangle}$, we will now compare our result (4.10) for $\sqrt{\langle 4\tau_c^2 \rangle}$ with the result for $\sqrt{\langle \tau_d^2 \rangle}$ found previously. The latter result, as given by Eq. (22) of Ref. [14], reads

$$\langle \tau_d^2 \rangle = 8\frac{b'^2}{D_\Phi} z, \tag{4.11}$$

which is three times as large as that for $\langle 4\tau_c^2 \rangle$. Recall that this result has been derived for the first birefringence model, and that is does not depend on the relative size of $l_{\rm cor}$ and $l_{\rm beat}$ of the fiber. Moreover, one can obtain essentially the same result, i.e., that $\langle \tau_d^2 \rangle = 3\langle 4\tau_c^2 \rangle$ for the second birefringence model. This can be rigorously shown in the case $l_{\rm cor} \ll l_{\rm beat}$, whereas for the general relation between $l_{\rm cor}$ and $l_{\rm beat}$ one needs to assume (since we could not prove it rigorously) that $\langle \tilde{S}_{1,4}^2(z) \rangle = \frac{1}{3}$ as $z \to \infty$ (cf. Appendix A). The latter assumption, however, seems to be confirmed by the results of numerical simulations reported in Ref. [32], and therefore appears to be actually correct. Therefore, we conclude that the timing jitter for solitons is smaller than that for linear pulses, and this appears to be independent of a specific model of the random birefringence. The physical reason for such a reduction will be given in Sec. V.

As to the second remark, we will compare the timing jitter (4.10) for the Manakov soliton with that of the soliton of a single NLS equation with the following random source:

$$iu_z + u_{\tau\tau} + 2u|u|^2 + i\sqrt{2}\kappa_3'(z)u_\tau = 0, \tag{4.12}$$

where $\kappa_3'(z)$ has the same properties as before [see Eqs. (3.12)]. The factor $\sqrt{2}$ in front of κ_3' is used here in order to have $\langle \kappa_3' \rangle = b'^2$. [This example with Eq. (4.12) is purely formal, and the reason for presenting it here will be given in Sec. V when we will explain the above result $\langle \tau_d^2 \rangle = 3\langle 4\tau_c^2 \rangle$.] The last term in Eq. (4.12) can be removed by the coordinate transformation $(z,\tau) \to [\tilde{z} = z, \tilde{\tau} = \tau - \int_0^z \sqrt{2}\kappa_3'(s)ds]$, which immediately yields the value for the timing jitter in this case:

$$\langle \tau_c^2(z) \rangle|_{\rm NLS} = 2\frac{b'^2}{D_\Phi} z. \tag{4.13}$$

In deriving Eq. (4.13), we used the relation $\langle \cos\Phi(z) \rangle = \cos\Phi(0)\exp(-D_\Phi z)$; cf. Eq. (8) in Ref. [14]. Thus the timing jitter for the Manakov soliton due to the linear PMD is three times less than it would be for the single NLS soliton driven by a similar perturbation. As explained in Sec. V, the reason for this is precisely the same as the reason for the jitter for two-component solitons being less than that for linear pulses.

Third and last, this timing jitter is the same for all solitons in the same data string, provided that all these solitons have the same initial state of polarization [cf. Eqs. (4.7b) and (4.5a)], and provided that the random polarization rotations of individual solitons coming from the noise in the lumped amplifiers are ignored [13,35]. (As demonstrated in Ref. [35], the latter rotations lead to a rather small timing jitter, compared to the jitter caused by other sources, for realistic fiber and pulse parameters.) Moreover, the timing jitter given by Eq. (4.7b) does not depend on the velocity of the solitons and thus will be the same for the solitons in any of the WDM (wavelength-division multiplexed) channels. (Note: Recall that a shift in the soliton's velocity translates in the physical units into a shift of its carrier frequency or wavelength.) Thus it seems that problems that such a jitter can cause for a soliton transmission line are benign, whereas they have been shown to be quite detrimental for linear pulses.

Now, Eqs. (4.7c)–(4.7e) determine the evolution of the phases α and φ and the polarization angle β of the soliton. These quantities are not important for the applications, as long as one is only concerned with the propagation of a single pulse. (Note: They may become important, however, when one considers a collision of solitons in two different WDM channels.) Therefore, and also because of technical difficulties which arise due to the factor of $(\sin 2\beta)$ on the left-hand side of Eq. (4.7e), we have not evaluated the average values of $\langle \alpha^2 \rangle$, $\langle \varphi^2 \rangle$, and $\langle \beta^2 \rangle$ here.

Finally, to conclude the consideration of the effect of the PMD on the soliton parameters, we will discuss one feature of Eqs. (4.7d) and (4.7e) which is worth mentioning. While the changes due to the nonlinear PMD of the phase φ and the polarization angle β are independent of the soliton's velocity, those changes which are due to the linear PMD are proportional to the velocity. The latter observation may seem strange at first glance, because, given a soliton with a nonzero velocity and some central frequency ω_0 of its carrier, one can go to a reference frame where its velocity will be zero by merely choosing a carrier with a certain frequency ω_1 different from ω_0. Thus it would seem to be possible to alter the evolution of the measurable physical quantities φ and β by an unphysical act of choosing a different value for the carrier frequency. The key to resolving this contradiction lies in the fact that in the original evolution Eq. (3.1), from which Eq. (3.13) was derived, the matrix K does depend on ω_0. [Note: Other parameters in Eq. (3.1) also depend on ω_0, but their dependence is not crucial for the present considerations.] Thus, having shifted the value of the carrier frequency from ω_0 to ω_1, one then must modify K, so that $[K(\omega_1) - K(\omega_0)] \sim \Delta$ [cf. Eq. (3.4)]. But then, from Eqs. (3.6) and (3.5), the definition of vector $\vec{u}$ will also be modified as follows:

$$\vec{u}(\omega_0) = [K^{-1}(\omega_0)K(\omega_1)]\vec{u}(\omega_1). \tag{4.14}$$

Then from Eq. (2.1) one has

$$(e^{i\sigma_3\varphi}\hat{B})(\omega_0) = [K^{-1}(\omega_0)K(\omega_1)](e^{i\sigma_3\varphi}\hat{B})(\omega_1), \tag{4.15}$$

and even though $\varphi(\omega_1)$ and $\beta(\omega_1)$ are, according to Eqs. (4.7d) and (4.7e), not affected by the linear PMD, because the velocity of the soliton in the new reference frame is zero, $\varphi(\omega_0)$ and $\beta(\omega_0)$ still will vary with z, simply because $[K^{-1}(\omega_0)K(\omega_1)]$ is not the identity transformation.

B. Generation of radiation by the soliton

Now we will calculate the radiation field generated by the soliton (2.1) due to the perturbations (4.1) and (4.2). Due to the complexity of the resulting formulas, it will be convenient to present the results for the linear and nonlinear PMD's separately. We will consider the case of the linear PMD first.

Let us denote the right-hand side of Eq. (2.21g) as $f_{1,3}(k,z)$, where the subscripts ''1'' and ''3'' correspond to those in Eq. (2.21g). Substituting Eqs. (2.15) and (4.3)–(4.5) into the left-hand side of (2.21g), we obtain

$$f_{1,\mathrm{lin}}(k,z)\equiv 0, \qquad (4.16a)$$

$$f_{3,\mathrm{lin}}(k,z)=i\eta^2(-ik+1)Q_2^*\,\mathrm{sech}\,\frac{\pi k}{2}. \qquad (4.16b)$$

In deriving Eq. (4.16), we used the following integrals:

$$\int_{-\infty}^{\infty} d\theta\, e^{-ik\theta}\,\mathrm{sech}\theta=\pi\,\mathrm{sech}\,\frac{\pi k}{2}\equiv J_1(k), \qquad (4.17a)$$

$$\int_{-\infty}^{\infty} d\theta\, e^{-ik\theta}\,\mathrm{sech}\theta\,\tanh\theta=-ikJ_1(k), \qquad (4.17b)$$

$$\int_{-\infty}^{\infty} d\theta\, e^{-ik\theta}\,\mathrm{sech}^3\theta=\tfrac{1}{2}(k^2+1)J_1(k)\equiv J_3(k). \qquad (4.17c)$$

Let us note that, here and below, the expansion coefficient $g_n(k,z)$ $(n=1,3)$, defined in Eq. (2.14), is related to the corresponding coefficient $f_n(k,z)$ by

$$g_n(k,z)=-i\int_0^z e^{i\lambda(z-z')}f_n(k,z')dz', \qquad (4.18)$$

which follows from Eq. (2.21g); here $\lambda=4\eta^2(k^2+1)$.

Similar to the derivation of Eq. (4.16), we substitute Eq. (4.3), in which R_{lin} is replaced by R_{nl} given by Eq. (4.6), into Eq. (2.21g) and find

$$f_{1,\mathrm{nl}}(k,z)=\eta^3(k+i)^2\,\mathrm{sech}\,\frac{\pi k}{2}n_1(z), \qquad (4.19a)$$

$$f_{3,\mathrm{nl}}(k,z)=\frac{4}{3}\eta^3k(k+i)\,\mathrm{sech}\,\frac{\pi k}{2}n_2^*(z). \qquad (4.19b)$$

In arriving at Eq. (4.19), we also used, in addition to integrals (4.17), the integral

$$\int_{-\infty}^{\infty} d\theta\, e^{-ik\theta}\,\mathrm{sech}^5\theta=\frac{3}{4}\left(1+\frac{k^2}{9}\right)J_3(k)\equiv J_5(k), \qquad (4.20)$$

where $J_3(k)$ has been defined in Eq. (4.17c).

The quantities of physical interest are the averages $\langle|u_{10}|^2\rangle$ and $\langle|v_{10}|^2\rangle$, where u_{10} and v_{10} are defined in Eq. (2.3). Note that the u and v components of the radiation are *not*, in general, equal to u_{10} and v_{10}, respectively, but rather are related to them via the transformation [cf. Eq. (2.3)]

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix}=e^{i\Psi}e^{i\sigma_3\varphi}\hat{B}\begin{pmatrix} u_{10} \\ v_{10} \end{pmatrix}.$$

But it is u_{10} and v_{10} that determine the polarization of the radiation field relative to the polarization of the soliton. Then, from Eqs. (2.14) and (2.5a), one finds that u_{10} and v_{10} are given by the following expansions:

$$u_{10}=\int dk(g_1\psi_1^{(1)}+g_1^*\psi_1^{(2)*}), \qquad (4.21a)$$

$$v_{10}=\int dk\, g_3^*\psi_3^{(4)*}, \qquad (4.21b)$$

where $\psi_n^{(j)}$ denotes the jth component of the eigenfunction ψ_n, $n=1$ and 3. Here and below, integration over k is assumed to be performed over the whole real axis. When writing down Eq. (4.21), we have used the symmetry relation $\overline{\psi}_{1,3}=\sigma_1\psi_{1,3}^*$ and also the fact that only one component of ψ_3 does not vanish. In Eq. (4.21) it is understood that $g_1\equiv g_{1,\mathrm{nl}}$, and $g_3\equiv g_{3,\mathrm{lin}}+g_{3,\mathrm{nl}}$.

Consider first the quantity $\langle|u_{10}|^2\rangle$,

$$\begin{aligned}\langle|u_{10}|^2\rangle=\int\int &[\langle g_1(k')g_1^*(k)\rangle\psi_1^{(2)}(k')\psi_1^{(2)*}(k)\\ &+\langle g_1^*(k')g_1(k)\rangle\psi_1^{(1)*}(k')\psi_1^{(1)}(k)\\ &+\langle g_1(k')g_1(k)\rangle\psi_1^{(2)}(k')\psi_1^{(1)}(k)\\ &+\langle g_1^*(k')g_1^*(k)\rangle\psi_1^{(1)*}(k')\psi_1^{(2)*}(k)]dk\,dk'.\end{aligned} \qquad (4.22)$$

Using Eq. (4.18), we obtain

$$\begin{aligned}\langle g_1^*(k')g_1(k)\rangle=&\int_0^z\int_0^z e^{-i\lambda'(z-z_1)+i\lambda(z-z_2)}\\ &\times\langle f_1^*(k',z_1)f_1(k,z_2)\rangle dz_1dz_2.\end{aligned} \qquad (4.23)$$

Then, from Eqs. (4.23) and (4.19a), one sees that the average on the right-hand side of Eq. (4.23) is proportional to $\langle n_1(z_1)n_1(z_2)\rangle$. Since z in Eq. (4.23) is to be on the order of several soliton dispersion lengths l_{sol}, and the correlation length of $n_1(z)$ is on the order of $L_{\mathrm{nl}}\ll l_{\mathrm{sol}}$ (cf. Sec. III), then one can use the following approximation:

$$\langle n_1(z_1)n_1(z_2)\rangle=2D_{n1}\delta(z_1-z_2), \qquad (4.24)$$

where [cf. Eq. (3.16)]

$$D_{n1}=\langle n_1^2(z)\rangle L_{\mathrm{nl}}. \qquad (4.25)$$

Note that Eq. (4.25) is consistent with Eq. (4.24) provided that one interprets the δ function in Eq. (4.24) as a limiting case of the right-hand side of Eq. (3.16) with $L\to 0$. Now, to be more precise, one needs to take into account the fact that there are several groups of terms in $n_1(z)$ which give a nonzero contribution to the average $\langle n_1(z_1)n_1(z_2)\rangle$ as $z_1\to z_2$. Their individual correlation lengths can, in general, differ from group to group, even though they all are expected to be of the same order of magnitude, namely, L_{nl}. For example, the correlation lengths of the following two terms in $n_1(z)$: $(m_1^{*3}m_2)$ and $(|m_1|^4)$, must be different, simply because the correlation length of the former term is finite when in Eq. (3.6) one (formally) sets $\kappa_1=\kappa_2\equiv 0$, $\kappa_3\neq 0$ [cf. Eq. (3.3)], whereas the correlation length of the latter term is obviously infinite in that case. Therefore, rigorously speaking, Eq. (4.25) needs to be replaced by

$$D_{n1}=\sum_s \langle n_1^2(z)\rangle^{(s)}L_{\rm nl}^{(s)}, \tag{4.25$'$}$$

where the summation is performed over all possible groups, with the terms within one group having the same correlation length $L_{\rm nl}^{(s)}$, and $\langle n_1^2(z)\rangle^{(s)}$ is the corresponding average for the (s)th group. Since the expression for $n_1(z)$, not to mention that for $n_1^2(z)$, is extremely cumbersome, then it would be technically very difficult to perform the separation of terms in $[n_1^2(z)]$ into such groups. Moreover, a calculation, which can be performed using either of the methods outlined in Refs. [14] or [31], of the quantity $L_{\rm nl}^{(s)}$ for each group would also be a formidable task. Therefore, we chose not to evaluate D_{n1} from Eq. (4.25′), but instead will use, as needed, the following simple estimate:

$$D_{n1}=O(1)L_{\rm nl}. \tag{4.26}$$

Thus using Eqs. (4.23), (4.24), (4.19a), and (2.10b), we obtain for the first term in Eq. (4.22)

$$\begin{aligned}I&\equiv\langle g_1^*(k)g_1(k')\rangle\psi_1^{(2)*}(k)\psi_1^{(2)}(k')\\&=D_{n1}\eta^4e^{i(k'-k)\theta}\frac{e^{i(\lambda'-\lambda)z}-1}{2i(k'^2-k^2)}\,\mathrm{sech}\,\frac{\pi k}{2}\,\mathrm{sech}\,\frac{\pi k'}{2}\\&\quad\times[(k^2-1)-2ik\tanh\theta+\mathrm{sech}^2\theta][(k'^2-1)\\&\quad+2ik'\tanh\theta+\mathrm{sech}^2\theta].\end{aligned} \tag{4.27}$$

To evaluate the double integral of this term, as required by Eq. (4.22), we first consider $\partial I/\partial z$. Taking the z derivative of the right-hand side of Eq. (4.27) eliminates the denominator in that expression, thus making the integration possible in the asymptotic limit $z\to\infty$. In this limit, the integration can be done by the method of stationary phase, which yields

$$\begin{aligned}\int\int dk\,dk'\left.\frac{\partial I}{\partial z}\right|_{z\to\infty}&=\frac{\pi D_{n1}\eta^4}{2z}\,\mathrm{sech}^2\frac{\pi k_0}{2}[(k_0^2+1)^2\\&\quad-2(k_0^2+1)\mathrm{sech}^2\theta+\mathrm{sech}^4\theta],\end{aligned} \tag{4.28a}$$

where

$$k_0=-\frac{\theta}{(8\eta^2z)} \tag{4.28b}$$

is the point of stationary phase. Since the integral (4.28a) converges uniformly in z (for large z), then we interchange the differentiation and double integration on the left-hand side of Eq. (4.28a). Then similar considerations for the other terms in Eq. (4.22) yield the following result:

$$\begin{aligned}\frac{\partial}{\partial z}\langle|u_{10}|^2\rangle\Big|_{z\to\infty}&\simeq\frac{\pi D_{n1}\eta^4}{2z}\,\mathrm{sech}^2\frac{\pi k_0}{2}[(k_0^2+1)^2\\&\quad-2(k_0^2+1)\mathrm{sech}^2\,\theta+2\,\mathrm{sech}^4\theta].\end{aligned} \tag{4.29}$$

Let us note that the z-dependent contribution of the last two terms in Eq. (4.22) is, in the limit considered, much smaller than that of the first two terms, and so it has been neglected. It also turns out that the z-independent contribution of the last two terms to the average quantity $\langle|u_{10}|^2\rangle$ in Eq. (4.22) vanishes exactly.

From Eq. (4.29) we observe, first, that at any given point θ in the soliton's reference frame, the average rate of generation of the radiation decreases inversely proportional to z. However, the quantity $\int_{-\infty}^{\infty}d\theta(\partial/\partial z)\langle|u_{10}|^2\rangle$, which is the rate of generation of the total amount of radiation, over the whole fiber length, in the component parallel to the unperturbed soliton, is independent of z [cf. Eq. (4.28b)]. Second, we can estimate the average width of the radiation field from the first term in Eq. (4.29). Using Eq. (4.28b) again, we find that the radiation spreads out around the soliton with a constant rate equal, on the order of magnitude, to $(4\pi\eta^2)$; this qualitatively agrees with the numerical results of Ref. [17]. Finally, integrating Eq. (4.29) yields the amount of radiation which is generated in the component parallel to the soliton:

$$\underline{\theta=O(1),\quad k_0\ll 1:}\quad\langle|u_{10}|^2\rangle\simeq\frac{\pi D_{n1}\eta^4}{2}(\,\mathrm{sech}^4\theta+\tanh^4\theta)\ln z, \tag{4.30a}$$

$$\underline{\theta\gg k_0\gg 1:}\quad\langle|u_{10}|^2\rangle\simeq 2D_{n1}\eta^4\left(\frac{\theta}{8\eta^2z}\right)^3\exp\left(-\frac{\pi\theta}{8\eta^2z}\right). \tag{4.30b}$$

Now let us consider the average radiation field that is generated in the component perpendicular to the soliton. Obviously,

$$\langle|v_{10}|^2\rangle=\langle|v_{10}|^2\rangle_{\rm lin}+\langle|v_{10}|^2\rangle_{\rm nl}, \tag{4.31}$$

where the first and second terms are generated due to the linear and nonlinear PMD's, respectively. Using then Eqs. (4.21b), (4.16b), (4.18), (2.15), and (A15), we obtain in the limit of $z\to\infty$:

$$\frac{\partial}{\partial z}\langle|v_{10}|^2\rangle_{\rm lin}\simeq\left(\frac{\pi}{4\eta^2z}\right)\frac{4}{3}\frac{b'^2}{D_\Phi}\eta^4\,\mathrm{sech}^2\frac{\pi k_0}{2}(\tanh^2\theta+k_0^2), \tag{4.32}$$

where k_0 was defined in Eq. (4.28b). We recall that this result is derived for the first birefringence model, Eqs. (3.9), and is independent of the relative size of $l_{\rm cor}$ and $l_{\rm beat}$ of the fiber. However, as is discussed in Appendix A, the same result almost certainly holds for the second birefringence-model as well. Similar to Eq. (4.32), in the same limit $z\to\infty$, we obtain that

$$\frac{\partial}{\partial z}\langle |v_{10}|^2\rangle_{\rm nl} \simeq \left(\frac{\pi}{4\eta^2 z}\right) 2D_{n2}\left(\frac{4}{3}\eta^3\right)^2 k_0^2 \,\mathrm{sech}^2\frac{\pi k_0}{2} \times(\tanh^2\theta + k_0^2), \tag{4.33}$$

where D_{n2} was defined similarly to Eq. (4.24).

All the remarks regarding Eq. (4.29), which we made in the text following that equation, also apply to Eqs. (4.32) and (4.33). In addition, we note that the radiation field in both components is much wider than the soliton. Thus at long distances, the soliton will appear to propagate on top of a wide, noisy pedestal formed by the radiation. Such a pedestal could affect the neighboring solitons. However, the numerical simulations of Ref. [36] showed that such an effect would be rather small compared to the effect of certain other perturbations in real fibers.

V. SECOND-ORDER RESULTS FOR THE SOLITON PARAMETERS, AND AN ANALYSIS OF RELATIVE DYNAMICS OF THE SOLITON'S COMPONENTS

All the analytical results that we found in Sec. IV agree well with the numerical results of the earlier studies [12,17]. However, the well-known broadening of the soliton, which was observed in these studies, has not been found in our first-order results. We will address this issue in this section. In so doing, we will first calculate the changes in the soliton's width and amplitude through the second order in ϵ, and show that they obey a linear-diffusion-type equation. In passing, we will also show that the soliton's parameter ξ, which determines its mean velocity and mean frequency, does not change through the second order in ϵ. Next, we will discuss in detail the intuitively appealing picture of relative oscillations of the centers of the soliton's two orthogonal components [23,24], while keeping in mind the following two goals: (i) to show how this picture relates to our perturbation results, and (ii) to compute the average value of the separation distance between the centers of the components. By accomplishing the latter goal, we will also qualitatively confirm criterion (1.4), originally found in Ref. [17] by numerical simulations, for the soliton not to be split up by the random birefringence.

The pulse broadening in a randomly birefringent fiber is a combination of two different effects. First, the pulse loses energy by emitting the radiation; hence its amplitude decreases and the width increases. The average rate of this process can be calculated as follows. Equations (1.1) possess the following conservation law:

$$i\frac{d}{dz}\int_{-\infty}^{\infty}(|u|^2+|v|^2)d\tau = \epsilon\int_{-\infty}^{\infty}(u^*R_u + v^*R_v - \text{c.c})d\tau. \tag{5.1}$$

For the pure soliton, the left-hand side of Eq. (5.1) equals $(4i\,d\eta/dz)$, and since we know [see (4.7a)] that, in first order, $d\eta/dz=0$, then, in this order, the right-hand side of Eq. (5.1) must also vanish. In the second order, Eq. (5.1) becomes

$$i\frac{d}{dz}\int_{-\infty}^{\infty}\{(|u_0|^2+|v_0|^2)+(|u_1|^2+|v_1|^2)\}d\tau = \epsilon\int_{-\infty}^{\infty}(u_1^*R_u[\vec{u}_0]+v_1^*R_v[\vec{u}_0]+u_0^*R_u[\vec{u}_0,\vec{u}_1] + v_0^*R_v[\vec{u}_0,\vec{u}_1] - \text{c.c})d\tau, \tag{5.2}$$

where $R_{u,v}[\vec{u}_0,\vec{u}_1]=R_{u,v}[\vec{u}_0+\vec{u}_1]-R_{u,v}[\vec{u}_0]$, $\vec{u}_0$ is defined in Eq. (2.1), and $\vec{u}_1$ is the second term in Eq. (2.3a) $[\vec{u}_1 = O(\epsilon)]$. In obtaining the left-hand side of Eq. (5.2), we used the orthogonality between the soliton and the radiation. Below we will present calculations, based on Eq. (5.2), only for the linear PMD, since the expressions for $R_{u,v}[\vec{u}_0,\vec{u}_1]$ for the nonlinear PMD are extremely formidable. This, however, does not seem to affect the final result significantly, since, as explained in Sec. III, the nonlinear PMD is much weaker than the linear one for pulses of about 20 ps and longer.

Using Eqs. (2.3), (3.13), and (3.14), and the fact that $(u_{10})_{\rm lin}\equiv 0$ [cf. Eq. (4.16a)], one can rewrite Eq. (5.2) as follows:

$$i\frac{d}{dz}\int_{-\infty}^{\infty}\left(\frac{1}{2\eta}\right)(|u_{00}|^2+|v_{10}|^2)d\theta = (-i\epsilon)\int_{-\infty}^{\infty}\Big[\{\Omega_1(-\sin 2\beta)+\Omega_2\cos^2\beta e^{2i\varphi}-\Omega_2^*\sin^2\beta e^{-2i\varphi}\}v_{10}^* \times\left(-i\frac{\xi}{\eta}u_{00}+\partial_\theta u_{00}\right)+\text{c.c}\Big]d\theta + (-i\epsilon)\int_{-\infty}^{\infty}\Big[\{\Omega_1(-\sin 2\beta) + \Omega_2^*\cos^2\beta e^{-2i\varphi}-\Omega_2\sin^2\beta e^{2i\varphi}\}\left(-i\frac{\xi}{\eta}v_{10}+\partial_\theta v_{10}\right)+\text{c.c}\Big]u_{00}d\theta, \tag{5.3}$$

where $\Omega_{1,2}$ are the components of the matrix Ω, defined in Eq. (3.14):

$$\Omega = \begin{pmatrix}\Omega_1 & \Omega_2^* \\ \Omega_2 & -\Omega_1\end{pmatrix}.$$

Recall that, by our convention, adopted in Sec. IV, we treat v_{10} as a quantity of order ϵ. The terms in Eq. (5.3) that contain the products $v_{10}^*\partial_\theta u_{00}$, $u_{00}\partial_\theta v_{10}^*$, and their complex conjugates, make no contribution to the integral on the right-hand side, because they can be combined into total θ derivatives of $v_{10}^* u_{00}$ and $v_{10}u_{00}$. On the other hand, the coefficients of the products $v_{10}^{*}u_{00}$ and $v_{10}u_{00}$ identically vanish. Thus the right-hand side of Eq. (5.3) is zero.

Similarly, one obtains the momentum conservation law of Eqs. (1.1):

$$i\frac{d}{dz}\int_{-\infty}^{\infty}(u\partial_\theta u^*+v\partial_\theta v^*)d\theta=\epsilon\int_{-\infty}^{\infty}(R_u\partial_\theta u^*+R_v\partial_\theta v^* +\text{c.c})d\theta. \tag{5.4}$$

The right-hand side of Eq. (5.4) can also be shown to be zero in second order in ϵ. Then, from Eqs. (5.3) and (5.4), respectively, one obtains the following equations:

$$i\frac{d}{dz}\int_{-\infty}^{\infty}\frac{1}{\eta}(u_{00}^2+|v_{10}|^2)d\theta=O(\epsilon^3),$$

$$i\frac{d}{dz}\int_{-\infty}^{\infty}\frac{i\xi}{\eta}(u_{00}^2+|v_{10}|^2)d\theta=O(\epsilon^3),$$

which immediately yield the following evolutions for η and ξ:

$$8\left\langle\frac{d\eta}{dz}\right\rangle=-\frac{d}{dz}\int_{-\infty}^{\infty}\frac{1}{\eta}\langle|v_{10}|^2\rangle d\theta+O(\epsilon^3), \tag{5.5}$$

$$\frac{d\xi}{dz}=O(\epsilon^3), \tag{5.6}$$

where we have used the explicit form of u_{00}, Eq. (2.3b). Staying with the second-order accuracy (i.e., ϵ^2) and then using Eqs. (4.32) (recall that we are considering only the linear PMD contribution) and (4.28b), one obtains from Eq. (5.6) the following equation:

$$\left\langle\frac{d\eta}{dz}\right\rangle=-\frac{16}{9}\langle\eta^3\rangle\frac{b'^2}{D_\Phi}, \tag{5.7}$$

which is valid in the limit $z\to\infty$. Since $\eta-\langle\eta\rangle\leqslant\mathcal{O}(\epsilon)$, then one can replace $\langle\eta^3\rangle$ by $\langle\eta\rangle^3$ above and hence obtain the following solution of Eq. (5.7):

$$\langle 2\eta\rangle=\frac{2\eta_0}{\sqrt{1+4\eta_0^2(8b'^2/9D_\Phi)_z}}, \tag{5.8}$$

where $\eta_0=\eta(z=0)$. Thus the soliton's amplitude, 2η, decreases by a linear-diffusion-type law, and the width, $1/(2\eta)$, increases accordingly.

Now the second reason for the pulse broadening is the fluctuation in the distance between the centers of the pulse's two components, caused by the random birefringence. Indeed, let us consider a composite soliton, whose components are slightly shifted relative to one another:

$$\vec{A}\equiv\begin{pmatrix}A_1\\A_2\end{pmatrix}=\begin{pmatrix}m_1 2\eta\,\mathrm{sech}[2\eta(\tau+\Delta\tau)]\\m_2 2\eta\,\mathrm{sech}[2\eta(\tau-\Delta\tau)]\end{pmatrix}, \tag{5.9}$$

where $\vec{A}$ is the original field vector, and $m_{1,2}(z)$ were defined in Eq. (3.5). Here we have also assumed that the soliton parameters β and φ [see Eq. (2.1)] are zero, which does not affect the generality of the considerations below. Now suppose that the signal detector at the receiving end of the telecommunication line measures the total intensity $(|u|^2+|v|^2)$, irrespective of the signal's polarization:

$$|u|^2+|v|^2 =4\eta^2[\mathrm{sech}^2\theta+4\eta\Delta\tau(|m_2|^2-|m_1|^2)\mathrm{sech}^2\theta\tanh^2\theta +4\eta^2\Delta\tau^2\mathrm{sech}^2\theta(2-3\,\mathrm{sech}^2\theta)]+O(\Delta\tau^3), \tag{5.10}$$

where $\theta=2\eta\tau$ and we have used Eq. (3.7). The second term on the right-hand side of Eq. (5.10) corresponds to the soliton's timing jitter, which was extensively discussed in Sec. IV A. Meanwhile, the third term appears to increase the width and decrease the amplitude of the total intensity profile, as can be easily verified by plotting the right-hand side of Eq. (5.10) and comparing the two cases of $\Delta\tau=0$ and $\Delta\tau\neq 0$. Note that the pulse broadening due to this mechanism is, on average, independent of z. Therefore, the broadening due to emission of radiation by the soliton, as given by Eq. (5.8), will dominate for large z.

Let us now relate the intuitive picture based on the ansatz (5.9) with the results obtained in Sec. IV. For $\Delta\tau\ll 1$, Eq. (5.9) can be rewritten as follows:

$$\begin{pmatrix}u\\v\end{pmatrix}\approx\begin{pmatrix}2\eta\,\mathrm{sech}\,\theta\\0\end{pmatrix}+2\eta\Delta\tau \times\begin{pmatrix}(|m_2|^2-|m_1|^2)2\eta\,\mathrm{sech}\,\theta\,\tanh\,\theta\\2m_1m_2 2\eta\,\mathrm{sech}\,\theta\,\tanh\,\theta\end{pmatrix}. \tag{5.11}$$

On the right-hand side of Eq. (5.11), the first component of the second term is proportional to the discrete spectrum mode $\hat{\phi}_2$ [cf. Eq. (2.6)], which corresponds to the shift of the soliton's center, whereas the second component contains only a combination of the continuous spectrum modes, ψ_3 and $\overline{\psi_3}$ [cf. Eq. (2.15)]. Thus the intuitive picture of the bound oscillations of the soliton's components about their center of mass is translated, in the language of the perturbation theory, into a timing jitter of the soliton [cf. Eqs. (2.7) and (4.7b)] and radiation in the component orthogonal to the unperturbed soliton, i.e., v_{10}.

Before we proceed, let us make three more observations about Eq. (5.11). First, by comparing it with Eqs. (4.16) and (4.7), one concludes that it is the linear PMD that causes a relative change of positions of the soliton's components, whereas the nonlinear PMD makes no contribution to this process.

Second, we can now give the reason for the result, found in Sec. IV, that the timing jitter for the Manakov soliton is $\sqrt{3}$ times less than both the timing jitter for a linear pulse in a randomly birefringent fiber and that for a scalar NLS soliton. Indeed, as we already noted, the potential energy of the relative displacement of the components of the Manakov soliton goes to causing both the timing jitter and radiation, with the energy going to the radiation being twice that going to the discrete spectrum mode $\hat{\phi}_2$. The latter statement follows upon noticing that the first and second components of the second term in Eq. (5.11) are proportional to $\tilde{S}_1$ and $\tilde{S}_4$, respectively, and that $\langle\tilde{S}_1^2(z)\rangle=\frac{1}{3}$, $\langle|\tilde{S}_4|^2(z)\rangle=\frac{2}{3}$ for $z\to\infty$

[cf. Eqs. (A14), (A15), and (A17) in Appendix A]. It is also reasonable to state that the emitted radiation is not, on average, absorbed back by the soliton at later times, because the perturbations are noncorrelated in z. Thus only one third of the ''work'' done by the random birefringence goes to causing the soliton's timing jitter. On the other hand, for the linear pulse, all the energy of the relative displacement must go to the differential delay time τ_d which was shown to be an analog of the timing jitter between the components, because there is no physical mechanism that would distinguish between a linear pulse and the radiation. Similarly, the perturbation term in the scalar NLS [Eq. (4.12)] does not generate any radiation, and thus all the energy of the perturbation goes to the timing jitter.

Third observation: ansatz (5.9) would actually mimic the exact, first-order solution of Eqs. (1.1) if the perturbation had the form (4.2a) with $\kappa_1' \equiv 0$ [cf. Eq. (3.14)] [compare the second term on the right-hand side of Eq. (5.11) with Eqs. (4.7b) and (4.16b)]. If, however, one intends to study the dynamics of the soliton's components in the case when both $\kappa_1' \neq 0$ and $\kappa_3' \neq 0$, then one needs to use a more complicated form of the ansatz, namely,

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = U \begin{pmatrix} m_1\, 2\eta\, \mathrm{sech}[2\eta(\tau+\Delta\tau)] \\ m_2\, 2\eta\, \mathrm{sech}[2\eta(\tau-\Delta\tau)] \end{pmatrix}, \qquad (5.9')$$

where U is some unitary matrix.

In the remainder of this section, we will obtain some qualitative information about the quantity $\langle \Delta\tau^2 \rangle$. This part of our work is distinctly different from the preceding part that has dealt with the perturbation theory. In contrast, now we will attempt to obtain a criterion for the soliton splitting into two individual components. As was explained in Sec. I, this criterion could not be found with the perturbation theory developed above. Instead, we will employ a less rigorous (from the viewpoint of the IST) model that treats the soliton as a bound state of two interacting components. We will use the simplified form of the ansatz (5.9), as opposed to its more exact form (5.9′). Moreover, we will limit ourselves to considering only the second birefringence model in the limit when $l_{\rm cor} \ll l_{\rm beat}$, because the calculations then are significantly simpler than in other cases. Using other reasonable birefringence models, as well as considering $l_{\rm cor}$ and $l_{\rm beat}$ to be of arbitrary relative size, are expected to produce results that are qualitatively (or, perhaps, even quantitatively) similar to those described below.

We require that our model account for the following two effects: (i) the attraction between the components due to the cross-nonlinearity in Eqs. (1.1), and (ii) the damping of the oscillations due to the emission of radiation. While the former effect has been analyzed in the literature [18,23,24], the latter one, to our knowledge, has not been, at least in this context. In particular, it will be important to determine the form of the damping of the oscillations (see below). Here we will not attempt to analyze this complex and very interesting problem, but will instead present the results for a simple phenomenological model of the oscillations of the components. As was shown in Ref. [23], under certain simplifying assumptions one can describe the dynamics of $\Delta\tau$ by the following equation of the harmonic oscillator with a *nonlinear* damping:

$$\frac{d^2\Delta\tau}{dz^2} + \gamma\left(\frac{d\Delta\tau}{dz}\right)\Delta\tau^2 + \omega^2\Delta\tau = \kappa'(z), \qquad (5.12)$$

where γ is a phenomenological damping constant, ω is the frequency of the free oscillations, computed in Ref. [23], and $\kappa'(z)$ is the δ-correlated (since we have assumed that $l_{\rm cor} \ll l_{\rm beat}$) random force arising due to the group-velocity birefringence. [Using the notation $\kappa'(z)$ here implies that this quantity is related, in some manner which will not be important for the following, to the previously introduced birefringence parameters $\kappa_1'(z)$ and $\kappa_3'(z)$.] Note that, in Ref. [23], there was no damping in the equations. Here we have introduced the nonlinear damping term in Eq. (5.12) because it is known (see, e.g., Sec. IV in Ref. [37]) that the damping of such a form leads to the $1/\sqrt{z}$-decay, as $z\to\infty$, of the free oscillations. The latter type of decay is characteristic of the radiation modes in dispersive evolution equations (see, e.g., references in Ref. [23]). In fact, an equation that reduces to Eq. (5.12) in the small-amplitude limit of the oscillations was derived in Ref. [37], although for a different nonlinear evolution equation (Kadomtsev–Petviashvili I equation). Note also that recently, another equation, that describes the evolution of internal mode oscillations, was derived in Ref. [38], which also predicts the $1/\sqrt{z}$ decay of the free oscillations.

Note that *the magnitude of the damping constant γ must be of order 1*, since no small parameter exists on the left-hand side of Eqs. (1.1). From Eq. (5.12), the following estimate for $\langle \Delta\tau^2 \rangle$ can be obtained (see Appendix B):

$$\langle \Delta\tau^2 \rangle \approx \left(\frac{D}{C\gamma\omega^2}\right)^{1/2}, \qquad (5.13)$$

where $\langle \kappa'(z)\kappa'(z_1) \rangle = 2D\,\delta(z-z_1)$, $D\sim\max(D_1,D_3)$ [cf. Eq. (3.8)], and C is some positive constant, with $C\to 3$ as (formally) $\gamma\to 0$. Since $C,\gamma,\omega \sim O(1)$, then one can rewrite Eq. (5.13) as

$$\sqrt{\langle \Delta\tau^2 \rangle} \sim D^{1/4}. \qquad (5.14)$$

Thus it is the intensity D of the random birefringence that determines the average separation between the soliton's components, and thus this intensity needs to be sufficiently small so that the soliton would not be split up into the separate components. This agrees qualitatively with the numerical criterion (1.4), which was obtained in Ref. [17].

VI. CONCLUSIONS

The two main goals achieved in this work are the following. First, in Sec. II, we found explicit equations for the first-order evolution of the soliton parameters and the radiation in the Manakov equation under the influence of an arbitrary perturbation. Second, in Secs. IV and V, we applied these results to study the effects of linear and nonlinear PMD's, as specified by Eq. (3.13), on the pulse propagation in randomly birefringent fibers.

In developing the perturbation theory for the soliton of the Manakov equation, we did not use the IST formalism for that equation. Instead, we made use of the fact that the exact one-soliton solution (which, in general, has two nonzero vector components) of the Manakov equation can be reduced to the soliton of the NLS (which has a single component) by means of a rotation of the reference frame, cf. Eq. (1.3). Thus we only needed to use the previously known mathematical formalism of the perturbation theory for the NLS soliton. Obviously, this would not be the case if one were to consider perturbations of a general, say, two-soliton solution of the Manakov equation; in the latter case, using the IST for the Manakov equation would be essential.

In the application of the results of Sec. II to the randomly birefringent fibers, we found that neither the linear nor nonlinear PMD's affect, in the first order, the amplitude and the mean velocity (mean frequency) of the soliton. However, the linear PMD does cause a timing jitter, whose rms value is given by Eq. (4.10). The magnitude of this jitter was shown to be $\sqrt{3}$ times less than that occurring to the linear pulse. (However, we were careful to point out that the effects of this jitter for the linear pulse and for the soliton are qualitatively different.) We also note that since all solitons in a given fiber, which initially have the same state of polarization, will have the same value of the PMD-induced timing jitter, then this jitter would probably not be detrimental for a soliton communication line. There exist another PMD-mediated timing jitter which is caused by the random rotation of polarization of individual solitons due to the noise from lumped amplifiers, and which can be potentially detrimental. However, as shown in Ref. [35], such a PMD-mediated jitter in real fibers is rather small compared to the jitter from other sources.

Next, we showed that both linear and nonlinear PMD's cause a slow, random polarization rotation in the reference frame, which itself rapidly rotates due to the random birefringence of the fiber, cf. Eqs. (3.5) and (3.6). This additional slow rotation seems not to be practically important if one considers propagation of a single pulse. However, it may become important when one considers a collision of two solitons with distinctly different velocities, as occurs when one has several WDM channels in a communication fiber. We postpone a study of this effect to a future work. Note that, according to Eqs. (4.7d) and (4.7e), only the linear PMD will be important in that case.

We also calculated the amount of radiation generated by the soliton. The rate of the emission of radiation, measured at a fixed position in a reference frame moving with the soliton, was found to decrease inversely proportional to the distance z. However, the width of the radiation field increases in a linear proportionality to z, thus making the rate of emission of the *total* amount of radiation constant along the fiber. The amount of radiation generated at some particular point in the soliton's reference frame then grows proportionally to $\ln z$, i.e., much slower than the width. Thus the radiation would form a wide pedestal, on top of which a soliton would propagate down the fiber.

In Sec. V, we found that the soliton's mean velocity (and hence the mean frequency) is not affected by the linear PMD through second order in ϵ. Thus the linear PMD will not interfere with using several different WDM channels, at least if one does not consider collisions of the solitons in different channels (see above). Next, we calculated, also for the linear PMD, the increase in the soliton's width through second order and found that it occurs solely due to the emission of radiation by the soliton. This increase obeys a linear-diffusion-type law, Eq. (5.8). Since the strength of the linear PMD is independent of the pulse width, while the strength of the nonlinear PMD is inversely proportional to its square [cf. Eq. (3.19)], then we expect that the above second-order results will be valid for pulses that are not too short (quantitative estimates, however, would require an exact knowledge of the birefringence parameters, which we do not possess at this time).

In Sec. V, we also discussed how the intuitive picture of oscillations of the centers of the soliton's components can be related to our perturbative results. Finally, in the same section, we used the analogy of the soliton's components with interacting particles to confirm the numerical criterion (1.4) qualitatively for the soliton to not be split up by the random birefringence. We found that, in order to derive such a criterion with the necessary numerical factor, as it appears in Eq. (1.4), one would need to determine the exact form of the damping term in Eq. (5.12). Note that the same problem, although in another context, was posed as early as in Ref. [23], but, to our knowledge, has not been solved.

ACKNOWLEDGMENTS

We gratefully acknowledge many useful discussions with Dr. C. R. Menyuk. This research was supported in part by the Office of Naval Research, under Grant No. N00014-95-1-0323. Acknowledgment and Disclaimer: Effort sponsored in part by the Air Force Office of Scientific Research, Air Force Materials Command, USAF, under Contract No. F49620-96-C-0031. The U. S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors, and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U. S. Government.

APPENDIX A: CALCULATIONS OF SOME AVERAGED QUANTITIES FOR THE FIRST BIREFRINGENCE MODEL

Here we will derive Eq. (4.10) for the soliton timing jitter, as well as discuss some other relevant results regarding the two models of random birefringence, given by Eqs. (3.8) and (3.9).

First, we establish the correspondence between the notations, introduced in Eqs. (3.8), (3.9), (3.11), and (3.12) of this paper with the corresponding notations used in Ref. [14]:

$$(l_{\mathrm{cor}},2D_\Phi,\kappa_1,\kappa_3,2D_1,2D_3,S_n)_{\text{this paper}} = (h_{\mathrm{fiber}},\sigma_\theta^2,y,x,\sigma^2 h_{\mathrm{fiber}},\sigma^2 h_{\mathrm{fiber}},a_n)_{\text{Ref. [14]}}, \quad n=1,\ldots 6 , \tag{A1}$$

and b and b' denote the same quantities as in Ref. [14]. Note that our quantities $S_{1,2,3}$ and $\tilde{S}_{1,2,3}$ become identical to, respectively, $S_{1,2,3}$ and $\tilde{S}_{1,2,3}$ of Ref. [14], if in Eq. (3.5) one sets $\vec{u}=(1,0)^T$. Using Eqs. (3.11a) and (3.6), one finds that the vectors (S_1,S_2,S_3) and (S_4,S_5,S_6) satisfy the same evolution equation:

$$\frac{d}{dz}\begin{pmatrix} S_{n+1} \\ S_{n+2} \\ S_{n+3} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2\kappa_1 \\ 0 & 0 & -2\kappa_3 \\ -2\kappa_1 & 2\kappa_3 & 0 \end{pmatrix} \begin{pmatrix} S_{n+1} \\ S_{n+2} \\ S_{n+3} \end{pmatrix}, \quad n=0,3. \tag{A2}$$

Also, one can easily establish the following identities:

$$S_1^2+S_2^2+S_3^2=1, \quad \tilde{S}_1^2+\tilde{S}_2^2+\tilde{S}_3^2=1, \tag{A3a}$$

$$S_4^2+S_5^2+S_6^2=0, \quad \tilde{S}_4^2+\tilde{S}_5^2+\tilde{S}_6^2=0, \tag{A3b}$$

$$S_1S_4+S_2S_5+S_3S_6=0, \quad \tilde{S}_1\tilde{S}_4+\tilde{S}_2\tilde{S}_5+\tilde{S}_3\tilde{S}_6=0, \tag{A3c}$$

$$|S_4|^2+|S_5|^2+|S_6|^2=2, \quad |\tilde{S}_4|^2+|\tilde{S}_5|^2+|\tilde{S}_6|^2=2. \tag{A3d}$$

Next, to compute $\langle \tau_c^2 \rangle$ in Eq. (4.9), we need to know the correlation function $\langle Q_1(z_1)Q_1(z_1+s)\rangle$ when $z_1 \to \infty$. From the explicit form of $Q_1(z)$, Eq. (4.5), one sees that the following quantities are then required:

$$\langle \tilde{S}_n(z_1)\tilde{S}_m(z_1+s)\rangle, \quad \langle \tilde{S}_n^*(z_1)\tilde{S}_m(z_1+s)\rangle, \quad n,m=1,4. \tag{A4}$$

We used the fact that the soliton parameters β and φ change over much longer distances ($> l_{\text{sol}}$) than S_n do (i.e., $l_{\text{diffusion}}$), and thus can be considered as constant in the calculation of $\langle Q_1(z_1)Q_1(z_1+s)\rangle$. Here we will detail only the calculations for the term $\langle \tilde{S}_1(z_1)\tilde{S}_1(z_1+s)\rangle$; all the other terms can be computed similarly. We first show, using a technique similar to the one used in Refs. [25,14], that the evolution in s of that correlator is the same as the evolution of $\langle \tilde{S}_1(s)\rangle$. Suppose $X(z)$ is a variable satisfying the *stochastic differential equation*

$$\frac{d}{dz}X(z)=A(X)+B(X)w(z), \tag{A5}$$

where A and B are some functions, and $w(z)$ is a white-noise process:

$$\langle w(z)\rangle=0, \quad \langle w(z_1)w(z_2)\rangle=2D_w\delta(z_1-z_2). \tag{A6}$$

Then the probability density $p(X_0,z)$ for X to have the value X_0 at the ''moment'' z, is governed by the Fokker-Plank equation (see, e.g., Ref. [39])

$$\frac{\partial p}{\partial z}=\frac{\partial}{\partial X_0}[A(X_0)p]+D_w\frac{\partial}{\partial X_0}\left(B(X_0)\frac{\partial}{\partial X_0}[B(X_0)p]\right) \equiv G^A p. \tag{A7}$$

Now, the average of any function $f[X]$ of X is, by definition, $\langle f[X]\rangle = \int dX\, p(X,z) f[X]$, and then

$$\frac{d}{dz}\langle f[X](z)\rangle = \int dX\, p(X,z)Gf[X] = \langle Gf[X]\rangle, \tag{A8}$$

where $G \equiv -A(\partial/\partial X)+D_w B(\partial/\partial X)B(\partial/\partial X)$ is the operator adjoint to G^A. Note that Eq. (A8) is the same as Eq. (A16), derived in Ref. [14] by a different means. In deriving Eq. (A8) we used the condition that p and $(\partial/\partial X)p$ vanish at the boundaries of the domain where X is defined. In the calculations that we need to do in this paper, this is always the case. A generalization to the case when X is a vector rather than a single variable is straightforward and can be found, in e.g., Ref. [39], Chap. 3; see also Refs. [25,14]. Now a two-point correlator $\langle f[X](z)f[X](z+s)\rangle$ is defined (Ref. [39], Chap. 2) with the two-point probability density $p(X_1,z,X_2,z+s)$ as follows:

$$\langle f[X](z)f[X](z+s)\rangle \equiv \int dX_1 dX_2 p(X_1,z,X_2,z+s) f[X_1] f[X_2].$$

The density $p(X_1,z,X_2,z+s)$ is known (Ref. [39], Chap. 5) to satisfy the same evolution equation, (A6), in the variable s as $p(X,z)$ does in z, with the operator G now acting only on the variable X_2 but not on X_1. Then

$$\frac{d}{ds}\langle f[X](z)f[X](z+s)\rangle = \langle f[X](z)Gf[X](z+s)\rangle. \tag{A9}$$

Taking $X=(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)$ and using the operator G given in Ref. [14], one obtains [cf. Eq. (30) in Ref. [14]]:

$$\frac{d}{ds}\langle \tilde{S}_1(z_1)\tilde{S}_1(z_1+s)\rangle = -D_\Phi \langle \tilde{S}_1(z_1)\tilde{S}_1(z_1+s)\rangle \quad (s>0). \tag{A10}$$

Note that (i) any of the correlators in Eq. (A4) satisfies the same equation, because $\tilde{S}_1$ and $\tilde{S}_4$ satisfy the same one; (ii) Eq. (A10) is obtained for the first birefringence model, and it holds irrespective of the relation between the birefringence correlation length and the beat length [cf. Eqs. (3.10)]. Now we need $\langle \tilde{S}_1^2(z_1)\rangle$ as the initial condition (at $s=0$) for the correlator in Eq. (A10). As was shown in Ref. [14] for the first birefringence model,

$$\langle \tilde{S}_1^2(z_1)\rangle|_{z_1\to\infty} = \tfrac{1}{3}, \tag{A11}$$

and thus

$$\langle \tilde{S}_1(z_1)\tilde{S}_1(z_1+s)\rangle = \tfrac{1}{3}e^{-D_\Phi s} \quad \left(z_1 \gg l_{\text{cor}} = \frac{1}{D_\Phi}, \quad s>0\right). \tag{A12}$$

Let us note that $\langle \tilde{S}_1^2(z)\rangle$ being nonzero is a direct consequence of Eq. (A3a). Similarly, using the rest of Eqs. (A3) and analogs of Eq. (A10), we find (for $z_1 \to \infty$)

$$\langle \tilde{S}_1(z_1)\tilde{S}_4(z_1+s)\rangle = 0, \quad \langle \tilde{S}_4(z_1)\tilde{S}_4(z_1+s)\rangle = 0, \tag{A13}$$

$$\langle \tilde{S}_4^*(z_1)\tilde{S}_4(z_1+s)\rangle = \tfrac{2}{3}e^{-D_\Phi s} \quad (s>0).$$

Using Eqs. (A12), (A13), and (4.5), one obtains

$$\int_{-\infty}^{\infty} ds \langle Q_1(z_1)Q_1(z_1+s)\rangle|_{z_1\to\infty} = \frac{2}{3}\frac{b'^2}{D_\Phi}, \tag{A14}$$

and then a substitution of Eq. (A14) into Eq. (4.9) yields Eq. (4.10). Here we used the fact that $l_{\rm cor} \sim L_{\rm lin}$, so that one has $z_1 \gg L_{\rm lin}$. Similarly, we find

$$\int_{-\infty}^{\infty} ds \langle Q_2^*(z_1)Q_2(z_1+s)\rangle|_{z_1\to\infty} = \frac{4}{3}\frac{b'^2}{D_\Phi}; \tag{A15}$$

this equation is used in deriving Eq. (4.32).

It is useful to note that an analog of Eq. (A10), with D_Φ being replaced by $1/l_{\rm cor}$, can also be obtained for the second birefringence model, Eqs. (3.8), where now $l_{\rm beat} = 2\pi l_{\rm cor}/(D_1+D_3)$. However, in that case we can rigorously derive Eq. (A11) only when $l_{\rm cor} \ll l_{\rm beat}$ (the details of the derivation are omitted). For a general relation between $l_{\rm cor}$ and $l_{\rm beat}$, we need to use a mean-field approximation for correlators like $\langle(\kappa_1^2+\kappa_3^2)\widetilde{S}_1\rangle$, i.e.,

$$\langle(\kappa_1^2+\kappa_3^2)\widetilde{S}_1\rangle = C\langle\kappa_1^2+\kappa_3^2\rangle\langle\widetilde{S}_1\rangle, \tag{A16}$$

with some constant $C>0$ (when $l_{\rm cor} \ll l_{\rm beat}$, then $C=1$). However, as the results of numerical simulations in Ref. [32] indicate, the direction of the Stokes vector becomes uniformly distributed over the Poincaré sphere for either of the birefringence models, which means that Eq. (A11), probably, holds for the second model as well. If that is indeed the case, then for that model, Eqs. (A14) and (A15) take on the following forms:

$$\int_{-\infty}^{\infty} ds \langle Q_1(z_1)Q_1(z_1+s)\rangle|_{z_1\to\infty} = \tfrac{2}{3}(D_1+D_3), \tag{A17a}$$

$$\int_{-\infty}^{\infty} ds \langle Q_2^*(z_1)Q_2(z_1+s)\rangle|_{z_1\to\infty} = \tfrac{4}{3}(D_1+D_3). \tag{A17b}$$

APPENDIX B: DERIVATION OF EQ. (5.13)

Here we will derive Eq. (5.13) from Eq. (5.12). Multiply Eq. (5.12) by $\Delta\tau$, take the average, and then look for the stationary case, thereby assuming that $d\langle\Delta\tau^2\rangle/dz=0$, etc., but $\langle(d\Delta\tau/dz)^2\rangle \neq 0$. One obtains:

$$-\left\langle\left(\frac{d\Delta\tau}{dz}\right)^2\right\rangle + \omega^2\langle\Delta\tau^2\rangle = \langle\Delta\tau(z)\kappa'(z)\rangle. \tag{B1}$$

Similarly, multiplying Eq. (5.12) by $(d\Delta\tau/dz)$, one obtains

$$\gamma\left\langle\left(\frac{d\Delta\tau}{dz}\right)^2\Delta\tau^2\right\rangle = \left\langle\frac{d\Delta\tau}{dz}(z)\kappa'(z)\right\rangle. \tag{B2}$$

To calculate the right-hand sides of Eqs. (B1) and (B2), we use the Novikov theorem [40], which states that if $\kappa'(z)$ is a stationary Gaussian process (which we assume here to be the case) and $\Phi[\kappa']$ is any functional of it; then

$$\langle\kappa'(z)\Phi[\kappa']\rangle = \int_0^z dz_1\langle\kappa'(z)\kappa'(z_1)\rangle\left\langle\frac{\delta\Phi[\kappa']}{\delta\kappa'(z_1)}\right\rangle, \tag{B3}$$

where $\delta\Phi[\kappa']/\delta\kappa'(z_1)$ denotes the functional derivative and $z=0$ is the initial point of the evolution. Since $\kappa'(z)$ can be assumed to be δ-correlated [see Eq. (4.9)]

$$\langle\kappa'(z)\kappa'(z_1)\rangle = 2D\,\delta(z-z_1), \tag{B4}$$

then we only need to find

$$\frac{\delta}{\delta\kappa'(z)}[\Delta\tau(z)] \quad \text{and} \quad \frac{\delta}{\delta\kappa'(z)}\left(\frac{d\Delta\tau(z)}{dz}\right).$$

To this end, rewrite Eq. (5.12) as a system of equations:

$$\frac{d\Delta\tau}{dz} = X, \qquad \frac{dX}{dz} = -\gamma X\Delta\tau^2 - \omega^2\Delta\tau + \kappa'(z), \tag{B5}$$

which can be further rewritten in the integral forms

$$\Delta\tau(z) = \int_0^z X(z_1)dz_1,$$

$$X(z) \equiv \frac{d\Delta\tau}{dz} = \int_0^z \{-\gamma X(z_1)[\Delta\tau(z_1)]^2 - \omega^2\Delta\tau(z_1) + \kappa'(z_1)\}dz_1. \tag{B6}$$

From Eq. (B6) one obtains (cf. Ref. [40])

$$\frac{\delta}{\delta\kappa'(z)}[\Delta\tau(z)] = 0 \quad \text{and} \quad \frac{\delta}{\delta\kappa'(z)}\left(\frac{d\Delta\tau(z)}{dz}\right) = 1\,. \tag{B7}$$

Then from Eqs. (B3), (B4), and (B7), one finds that

$$\langle\Delta\tau(z)\kappa'(z)\rangle = 0, \quad \left\langle\frac{d\Delta\tau}{dz}(z)\kappa'(z)\right\rangle = D, \tag{B8}$$

where we used the identity $\int_0^z \delta(z-z_1)dz_1 = \frac{1}{2}$.

One can also approximately decompose the fourth-order moment on the left-hand side of Eq. (B2) into a product of the two second-order moments:

$$\left\langle\left(\frac{d\Delta\tau}{dz}\right)^2\Delta\tau^2\right\rangle \approx C\left\langle\left(\frac{d\Delta\tau}{dz}\right)^2\right\rangle\langle\Delta\tau^2\rangle, \tag{B9}$$

where C is some constant, with the approximate equality becoming exact and with $C=3$ if either $\gamma=0$ or if the damping were linear [in those two cases, $\Delta\tau(z)$ would also be, as $\kappa'(z)$ is, a Gaussian process]. Finally, from Eqs. (B1), (B2), (B8), and (B9) one obtains Eq. (5.13).

[1] S. V. Manakov, Zh. Eksp. Teor. Fiz. **65**, 505 (1973), [Sov. Phys. JETP **38**, 248 (1974)].
[2] V. S. Gerdjikov, Teor Mat. Fiz. **99**, 292 (1994) [Theor. Math. Phys. Phys. **99**, 593 (1994)].
[3] D. J. Kaup, J. Math. Anal. Appl. **54**, 849 (1976).
[4] D. J. Kaup, SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **31**, 121 (1976).
[5] M. Midrio, S. Wabnitz, and P. Franco, Phys. Rev. E **54**, 5743 (1996).
[6] V. S. Shchesnovich and E. V. Doktorov, Phys. Rev. E **55**, 7626 (1997).
[7] V. I. Karpman, Phys. Scr. **20**, 462 (1979).
[8] D. J. Kaup, Phys. Rev. A **42**, 5689 (1990).
[9] J. P. Gordon, J. Opt. Soc. Am. B **9**, 91 (1992).
[10] J. N. Elgin, Phys. Rev. A **47**, 4331 (1993).
[11] B. A. Malomed, Phys. Rev. A **43**, 410 (1991).
[12] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, Opt. Lett. **16**, 1231 (1991).
[13] S. G. Evangelides, Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, J. Lightwave Technol. **10**, 28 (1992).
[14] P. K. A. Wai and C. R. Menyuk, J. Lightwave Technol. **14**, 148 (1996).
[15] C. R. Menyuk (private communication).
[16] C. R. Menyuk and P. K. A. Wai, J. Opt. Soc. Am. B **11**, 1288 (1994).
[17] L. F. Mollenauer, K. Smith, J. P. Gordon, and C. R. Menyuk, Opt. Lett. **14**, 1219 (1989).
[18] Yu. S. Kivshar, J. Opt. Soc. Am. B **7**, 2204 (1990).
[19] D. J. Kaup and B. A. Malomed, Phys. Rev. A **48**, 599 (1993).
[20] V. V. Afanasjev and V. A. Vysloukh, J. Opt. Soc. Am. B **11**, 2385 (1994).
[21] X. D. Cao and C. J. McKinstrie, J. Opt. Soc. Am. B **10**, 1202 (1993).
[22] C. R. Menyuk, Opt. Lett. **12**, 614 (1987); J. Opt. Soc. Am. B **5**, 392 (1988).
[23] T. Ueda and W. L. Kath, Phys. Rev. A **42**, 563 (1990).
[24] D. J. Kaup, B. A. Malomed, and R. S. Tasgal, Phys. Rev. E **48**, 3049 (1993).
[25] G. I. Foschini and C. D. Poole, J. Lightwave Technol. **9**, 1439 (1991).
[26] C. D. Poole, Opt. Lett. **14**, 523 (1989).
[27] C. R. Menyuk and P. K. A. Wai, J. Opt. Soc. Am. B **11**, 1305 (1994).
[28] J. U. Kang, G. I. Stegeman, J. S. Aitchison, and N. N. Akhmediev, Phys. Rev. Lett. **76**, 3699 (1996).
[29] J. W. Haus and T. I. Lakoba (unpublished).
[30] D. J. Kaup and T. I. Lakoba, J. Math. Phys. **37**, 3442 (1996).
[31] T. I. Lakoba, J. Opt. Soc. Am. B **13**, 2006 (1996).
[32] P. K. A. Wai and C. R. Menyuk, Opt. Lett. **19**, 1517 (1994).
[33] In Ref. [31], it was shown that in the case when $l_{\rm cor} \gg l_{\rm beat}$, the averaging yields a system that is, in general, different from the Manakov equation. This, however, does not contradict the conclusion of Ref. [14], where the Manakov equation was shown to result when model (3.8) with an *arbitrary* $l_{\rm cor}$ was used. The reason is that Ref. [31] used the assumption that (in the notations of this work) $|\kappa|/|\partial\kappa/\partial z| \gg l_{\rm beat}$. Clearly, this is not the case for Eq. (3.8), since the white noise has infinite amplitude (see, e.g., Ref. [39]). We also note that the assumptions under which model (3.8) was derived seem to be more realistic than the above assumption $|\kappa|/|\partial\kappa/\partial z| \gg l_{\rm beat}$. Indeed, a major source of randomness of the fiber's refractive indices is the fast and noncorrelated fluctuations in the equipment and the ambient conditions during the fiber drawing process; this is exactly what the white-noise term in Eq. (3.8) models.
[34] C. D. Poole and R. E. Wagner, Electron. Lett. **22**, 1029 (1986).
[35] L. F. Mollenauer and J. P. Gordon, Opt. Lett. **19**, 375 (1994).
[36] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, Opt. Lett. **16**, 1735 (1991).
[37] K. A. Gorshkov and D. E. Pelinovsky, Physica D **85**, 468 (1995).
[38] D. E. Pelinovsky, Yu. S. Kivshar, and V. V. Afanasjev, Physica D (to be published).
[39] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).
[40] E. A. Novikov, Zh. Eksp. Teor. Fiz. **47**, 1919 (1964) [Sov. Phys. JETP **20**, 1290 (1965)]; W. Horsthemke and R. Lefever, *Noise Induced Transitions* (Springer-Verlag, Berlin, 1983).