Nonlinear Theory of Polarization-Mode Dispersion for Fiber Solitons

A. Levent and S. G. Rajeev
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

F. Yaman and Govind P. Agrawal
The Institute of Optics, University of Rochester, Rochester, New York 14627
(Received 19 April 2002; published 8 January 2003)

We consider the evolution of optical solitons inside a nonlinear dispersive fiber with random birefringence, causing polarization-mode dispersion. We convert the pair of coupled nonlinear Schrödinger equations satisfied by the orthogonally polarized components into a Fokker-Planck equation using the collective-variable approach. We solve this equation and derive expressions for the probability density functions associated with the differential group delay and the pulse width in the limit of large propagation distances.

DOI: 10.1103/PhysRevLett.90.013902 PACS numbers: 42.65.Tg, 42.65.–k, 42.65.Wi, 42.81.Dp

It is well known that pulses of light can propagate inside an optical fiber in a way that preserves their shape when the dispersive and nonlinear effects are properly balanced [1]. Such optical solitons are being pursued for transmitting digital information [2]. If the fiber is randomly birefringent (due to variations in its core diameter or mechanical stresses along it), the two polarization components of the optical signal propagate at different speeds because of their different refractive indices. This phenomenon of *polarization mode dispersion* (PMD) causes differential group delay (DGD), whose stochastic nature impacts the performance of fiber-optic communication systems severely, especially at high bit rates [3–5]. The existing theory of PMD considers each frequency component of an optical pulse separately.

The effects of PMD on solitons have been observed experimentally [6] and studied theoretically [7–13]. It is found that solitons are more robust to PMD than linear pulses because the nonlinearities of the fiber that help preserve the shape of the soliton also inhibit large deviations in the DGD induced by PMD [11]. In this Letter, we study this phenomenon analytically using a fully nonlinear theory. More specifically, we derive a linear partial differential equation that describes the evolution of the probability distribution of the DGD at the fiber output. Our approach is based on a Fokker-Plank equation originally developed by Chandrashekhar [14] to describe globular clusters of stars. The adaptation of this theory allows us to obtain a Fokker-Plank equation that includes the effects of both the PMD and the fiber nonlinearity. The probability distribution we predict is markedly different from the linear theory and can be used to calculate the moments associated with the pulse width.

Let us begin our quantitative study with the basic equations governing the propagation of an optical pulse in a fiber with random birefringence. As is well known [1], Maxwell's equations in this case reduce to a set of two coupled nonlinear Schrödinger (NLS) equations. In the Jones-matrix formalism, these equations can be written in the following compact form [7–9]:

$$i\frac{\partial\psi}{\partial\xi}+\left(\boldsymbol{\sigma}\cdot\boldsymbol{b}_0\psi+i\boldsymbol{\sigma}\cdot\boldsymbol{b}_1\frac{\partial\psi}{\partial t}\right)+\frac{1}{2}\frac{\partial^2\psi}{\partial t^2}+s_0\psi-\frac{1}{3}s_3\sigma_3\psi=0, \tag{1}$$

where $\boldsymbol{\sigma}=(\sigma_1,\sigma_2,\sigma_3)$ is related to Pauli spin matrices. In our notation, the two polarization components of the optical pulse form a spinor $\psi(\xi,t)$. In the solitons units used commonly [1], $z=\xi L_D$ is the distance along the fiber axis and $t'=tT_0$ is the reduced time [1]. Here, L_D is the dispersion length and T_0 is a measure of the pulse width. The vectors $\boldsymbol{b}_0$ and $\boldsymbol{b}_1$ govern the PMD effects resulting from random birefringence changes along the fiber length and thus depend on ξ but not on t. Birefringence fluctuations can be modeled by a Markoffian random process with Gaussian white noise.

The state of polarization at a given location and instant is governed by the Stokes vector $\boldsymbol{s}(t,\xi)=\psi^\dagger\boldsymbol{\sigma}\psi(t,\xi)$; if integrated over all time, it gives the Stokes vector at any location: $\boldsymbol{S}=\int\boldsymbol{s}dt$. The quantity $s_0=\psi^\dagger\psi$ is the local intensity of the field. The PMD effects appear in Eq. (1) through the vectors $\boldsymbol{b}_0$ and $\boldsymbol{b}_1$ defined in the Stokes space.

Our theory is based on the observation that, in the absence of the PMD effects, Eq. (1) has an exact soliton solution given by [8]

$$\psi=A\text{sech}(t)\exp(i\xi/2), \tag{2}$$

where $A=(\cos\theta,\sin\theta)$ and θ is an arbitrary angle representing the partition of soliton amplitude between its two polarization components. The PMD effects perturb ψ. To study how the PMD perturbs ψ, it is convenient to view the optical field as a Hamiltonian dynamical system, with the distance ξ along the fiber playing the role of an evolution ("time") variable [15]. This variable should be distinguished from the physical time variable t. The set of all possible initial pulse shapes, $M = \{\psi : R \rightarrow g^2\}$, is the phase space of the dynamical system. It is easy to verify that the Hamiltonian

$$H(\psi) = \int \left(\frac{1}{2} \left| \frac{\partial \psi}{\partial t} \right|^2 - \frac{1}{2} s_0^2 + \frac{1}{6} s_3^2 \right) dt \tag{3}$$

provides the equation of motion for the soliton [Eq. (1) without the PMD effects] with the Poisson bracket

$$\begin{aligned} \{\psi_i^*(t), \psi_j(t')\} &= \delta_{ij}\delta(t - t'), \\ \{\psi_i(t), \psi_j(t')\} &= \{\psi_i^*(t), \psi_j^*(t')\} = 0, \end{aligned} \tag{4}$$

where $i, j = 1$ or 2. With the same Poisson brackets, the Hamiltonian

$$\begin{aligned} H_1(\psi) &= H(\psi) + \boldsymbol{b}_0 \cdot \boldsymbol{S} + \boldsymbol{b}_1 \cdot \boldsymbol{S}_1, \\ \boldsymbol{S}_1 &= \int \psi^\dagger \boldsymbol{\sigma} \frac{\partial \psi}{\partial t} dt \end{aligned} \tag{5}$$

yields Eq. (1). The important question is how the soliton of Eq. (2), whose dynamics is governed by H, is affected by the two PMD-induced perturbations $\boldsymbol{S}$ and $\boldsymbol{S}_1$ appearing in H_1.

To answer this question, we note that every observable of a Hamiltonian system generates some canonical transformation. It is not surprising that the Stokes vector generates a rotation of the polarization on the Poincaré sphere as

$$\{\boldsymbol{u} \cdot \boldsymbol{S}, \psi(t)\} = \boldsymbol{\sigma} \cdot \boldsymbol{u}\psi(t), \tag{6}$$

where $\boldsymbol{u}$ is an arbitrary vector in the Stokes space. Only rotations around the third axis are symmetries of H. Thus the term $\boldsymbol{b}_0$ in the Hamiltonian will cause rapid (random) rotations of the polarization of the soliton as it propagates along the fiber. This could be viewed as a random walk on the Poincaré sphere. To understand the effect of the $\boldsymbol{S}_1$ term, we consider the canonical transformation

$$\{\boldsymbol{u} \cdot \mathbf{S}_1, \psi(t)\} = \boldsymbol{\sigma} \cdot \boldsymbol{u} \frac{\partial \psi}{\partial t}. \tag{7}$$

Clearly, $\boldsymbol{S}_1$ induces a differential delay between the two polarization states: $\boldsymbol{u} \cdot \mathbf{S}_1$ advances the polarization component with eigenvalue of $\boldsymbol{\sigma} \cdot \boldsymbol{u}$ equal to $+1$ while delaying the one that has eigenvalue -1. This differential delay is known as the DGD.

We are interested in the probability distribution of the PMD vector because of fluctuations in the variables $\boldsymbol{b}_0$ and $\boldsymbol{b}_1$. Thus, we can restrict our attention to a six-dimensional submanifold of the phase space corresponding to the distortion of the soliton induced by the rotation of the polarization by $\boldsymbol{b}_0$ and the DGD induced by $\boldsymbol{b}_1$. The PMD-perturbed state of the soliton is governed by

$$\psi_{\boldsymbol{\phi},\boldsymbol{\tau}}(t) = \exp\left(i \frac{1}{2} \boldsymbol{\phi} \cdot \boldsymbol{\sigma} \right) \exp\left(\frac{1}{2} \boldsymbol{\tau} \cdot \boldsymbol{\sigma} \frac{\partial}{\partial t} \right) \psi_0(t). \tag{8}$$

Here, $\boldsymbol{\tau}$ is the PMD vector and $\psi_0 = A\text{sech}(t)$ is the initial pulse launched at $\xi = 0$. The matrix $\exp(\frac{i}{2}\boldsymbol{\phi} \cdot \boldsymbol{\sigma})$ is an element of SU(2) describing the rotation of the soliton polarization on the Poincaré sphere. The set of these rotated-delayed solitons configurations forms a six-dimensional manifold $N = \text{SU}(2) \times R^3$ with coordinates $(\boldsymbol{\phi}, \boldsymbol{\tau})$. It is important to stress that the form of the PMD-perturbed soliton in Eq. (8) is not chosen in an *ad hoc* fashion but follows from a systematic use of classical perturbation theory.

In a deterministic theory of solitons, it is common to study the soliton dynamics in a reduced-dimensional space using the *collective-variable method* [16–18], a method similar in spirit to the variational method used in the literature on optical solitons. We apply the same approach for the stochastic theory developed in this Letter. By substituting Eq. (8) in Eq. (3) and performing the integration, the *effective Hamiltonian* restricted to the six-dimensional submanifold takes the form

$$\begin{aligned} H_{\text{eff}}(\boldsymbol{\tau}) = &\left\{ 2[\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{\alpha}} - (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{\alpha}})(\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}})]^2/3 - \left[1 + \frac{1}{3}(\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}})^2 \right][1 - (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{\alpha}})^2] \right\} \frac{\tau \cosh\tau - \sinh\tau}{\sinh^3\tau} \\ &+ (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}})(\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{\alpha}})[\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{\alpha}} - (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{\alpha}})(\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}})] \frac{\sinh(2\tau) - 2\tau}{3\sinh^3\tau} + [1 + (\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{\alpha}})^2][(\hat{\boldsymbol{\tau}} \cdot \hat{\boldsymbol{n}})^2 - 3]/9. \end{aligned} \tag{9}$$

Here, $\hat{\boldsymbol{\alpha}}$ is the initial polarization state, i.e., $\hat{\boldsymbol{\alpha}} = A^\dagger \boldsymbol{\sigma} A$, $\hat{\boldsymbol{n}}$ is the unit vector obtained by rotating this initial polarization, $\hat{\boldsymbol{n}} = A^\dagger \exp(-\frac{i}{2}\boldsymbol{\phi} \cdot \boldsymbol{\sigma})\boldsymbol{\sigma} \times \exp(\frac{i}{2}\boldsymbol{\phi} \cdot \boldsymbol{\sigma})A$, and $\hat{\boldsymbol{\tau}}$ is the unit vector in the direction of $\boldsymbol{\tau}$.

This effective Hamiltonian describes the cost in "energy" for rotating or differentially delaying the soliton. It has a minimum at $\boldsymbol{\tau} = \boldsymbol{\phi} = 0$ and grows to a finite constant as $\tau \rightarrow \infty$. Physically speaking, the nonlinear effects provide a "restoring force" that resists the PMD-induced rotations and delays. But when the delay becomes large, this restoring force goes to zero. We are interested in the asymptotic ($\xi \gg 1$) probability distribution of the magnitude of the group delay τ, after averaging over the angles $\boldsymbol{\phi}$, $\hat{\boldsymbol{\tau}}$.

Before we calculate this asymptotic probability distribution, it is important to note from the fluctuation-dissipation theorem that there will also be some dissipation due to the effects not included in the effective Hamiltonian—manifested mainly as the transfer of energy from the soliton to the continuum radiation, also known as the dispersive waves. To include dissipation, we note that in the absence of nonlinearities we would expect the soliton to execute a random walk in the six-dimensional subspace. As usual, this random walk is governed by the diffusion equation [14]

$$\frac{\partial p(\boldsymbol{\tau}, \boldsymbol{\phi}, \xi)}{\partial \xi} = (D\nabla_{\boldsymbol{\tau}}^2 + D'\nabla_{\boldsymbol{\phi}}^2)p(\boldsymbol{\tau}, \boldsymbol{\phi}, \xi), \tag{10}$$

where D and D' are the diffusion constants and $p(\boldsymbol{\tau}, \boldsymbol{\phi}, \xi)$ is the joint probability density in the six-dimensional subspace $p(\boldsymbol{\tau}, \boldsymbol{\phi})$ in which random walk occurs. This equation should be solved with the initial condition $p(\boldsymbol{\tau}, \boldsymbol{\phi}, 0) = \delta^3(\boldsymbol{\tau})\delta^3(\boldsymbol{\phi})$.

We are interested only in the probability distribution of the DGD, equal to the magnitude of the PMD vector $\boldsymbol{\tau}$. If we write $\boldsymbol{\tau}$ in the spherical polar coordinates and average $p(\boldsymbol{\tau}, \boldsymbol{\phi}, \xi)$ over the five angles, we obtain the following simple diffusion equation for the probability distribution of DGD:

$$\frac{\partial p(\tau, \xi)}{\partial \xi} = D\frac{\partial}{\partial \tau}\left[\tau^2 \frac{\partial}{\partial \tau}[\tau^{-2}p(\tau, \xi)]\right]. \qquad (11)$$

The solution of this equation leads to the Maxwellian distribution, well known in the linear theory of PMD and given by [3]

$$p(\tau, \xi) = N\tau^2 \exp\left(-\frac{\tau^2}{4D\xi}\right). \qquad (12)$$

The diffusion constant can be related to the PMD parameter D_p of the fiber as $D = 3D_p^2/(2|\beta_2|)$, where the dispersion parameter β_2 of the fiber appears because of the use of the soliton units. The normalization constant N ensures that $\int p(\tau)d\tau = 1$.

The nonlinearities modify the Maxwellian distribution in such a way that the less energetic configurations of the soliton become more probable. In statistical mechanics, we understand this as the result of the fluctuations and dissipations being in balance on the average. Physically speaking, the nonlinear effects will cause a drift in the direction of the negative gradient of the "energy." The term in the Fokker-Plank equation describing this combination of drift and dissipation is determined by two factors [14]: first, it involves only a first-order derivative of W; second, the static solution of the Fokker-Plank equation leads to the Boltzmann distribution $e^{-H_{\rm av}/E_b}$, where E_b is a constant analogous to temperature. It measures the ratio of the levels of fluctuations to dissipation. The final Fokker-Plank equation that we obtain is given by

$$\frac{\partial p(\tau, \xi)}{\partial \xi} = D\frac{\partial}{\partial \tau}\left[\tau^2 \frac{\partial}{\partial \tau}(\tau^{-2}p)\right] + \gamma\frac{\partial}{\partial \tau}\left[\frac{\partial H_{\rm av}}{\partial \tau}p\right], \quad (13)$$

where the parameter γ is a measure of the nonlinearity-induced drift. The average value $H_{\rm av}$ of the effective Hamiltonian $H_{\rm eff}$ is obtained by averaging Eq. (9) over all angles and is found to be

$$H_{\rm av}(\tau) = \frac{64(\sinh\tau - \tau\cosh\tau)}{27\sinh^3\tau} - \frac{16}{9}. \qquad (14)$$

One can think of $H_{\rm av}(\tau)$ as the potential well created by the nonlinear effects that opposes the PMD-induced diffusion.

Ueda and Kath also derived a Fokker-Plank equation in the nonlinear regime in a previous study on the effects of random birefringence [19]. Our Fokker-Plank equation (13) is different from theirs. For example, their equation has a *constant* static solution. General principles of statistical physics [14] dictate that the static solution of the Fokker-Plank equation (which is the case of thermodynamic equilibrium) should be the Boltzmann distribution. A reason for this discrepancy turns out to be that the stochastic differential equation used in Ref. [19] includes only fluctuations (no dissipation). It is physically wrong to model a physical system with fluctuations without including dissipation—it violates the fluctuation-dissipation theorem. Without dissipation, the energy of a Hamiltonian system, subject to random fluctuating forces, would increase on the average, so that eventually the system will reach a state of infinite temperature. All states would then be equally probable (constant probability density) instead of the states of lower energy being more probable as the Boltzmann distribution (and physical intuition) requires. In the case of Brownian motion, this dissipation is due to viscosity; in our case it is due to the transfer of energy to continuum radiation.

The Fokker-Plank equation (13) constitutes our main result. Although it can be solved numerically to study the impact of nonlinear effects on PMD at any distance ξ, one is often interested only in the asymptotic behavior of the solution as $\xi \to \infty$. There is a static solution of the form $p(\tau) \sim \tau^2 \exp(-H_{\rm av}/E_b)$, where $E_b = D/\gamma$. This is the analog of the Boltzmann distribution that appears in many physical problems [14]. In our case, E_b plays the role of the "equilibrium" energy [20].

Usually the probability distribution would tend to this static solution, as $\xi \to \infty$. But this requires that $p(\tau) \to 0$ as $|\tau| \to \infty$ so that $\int p(\tau)d\tau$ can be normalized to 1. This is not possible in our case because the "potential well" has a finite range: $H_{\rm av}(\tau) \to 0$ for large τ. Thus the solution must reduce to that of the linear equation for large τ. Studying the limit of large ξ of the solution of the Fokker-Plank equation results in the following probability distribution for the DGD:

$$p(\tau, \xi) \approx N\tau^2 \exp\left[-\frac{H_{\rm av}(\tau)}{E_b} - \frac{\tau^2}{4D\xi}\right], \qquad (15)$$

where N again ensures that $\int_0^\infty p(\tau)d\tau = 1$. This is the main result of the nonlinear PMD theory developed in this Letter.

From a practical point of view, the effects of PMD manifest through broadening of each optical pulse by a random amount. The root-mean-square (rms) width T_p of any pulse can be quantified using

$$T_p^2 = \frac{\int t^2\psi^\dagger(t)\psi(t)dt}{\int \psi^\dagger(t)\psi(t)dt} = \frac{\pi^2}{12} + \frac{\tau^2}{4}. \qquad (16)$$

This relation can be used to find the probability distribution of T_p as well as the various moments of T_p such as the average value of the pulse broadening factor.

We apply our general results to a specific 40-Gbit/s system designed to transmit 5-ps solitons ($T_0 = 2.84$ ps)

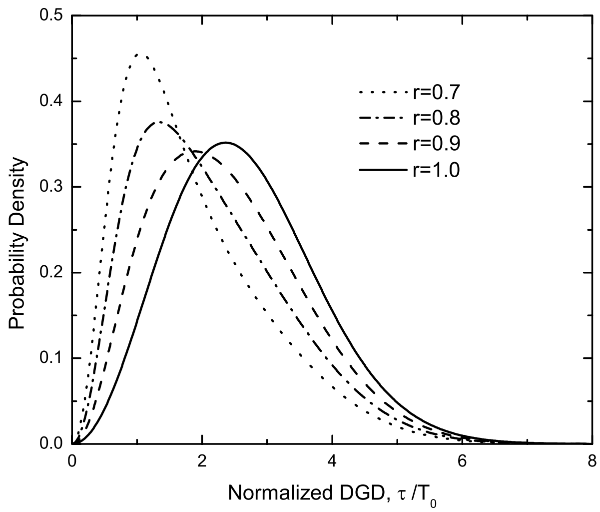

FIG. 1. Probability density $p(\tau)$ as a function of DGD τ for several values of the reduction factor r after the pulse has propagated 3000 km in a fiber with $D_p = 0.15\ \mathrm{ps\,km}^{-1/2}$. The linear case corresponds to $r = 1$.

over 3000 km of optical fiber with the PMD parameter $D_p = 0.15\ \mathrm{ps\,km}^{-1/2}$. The average DGD in the linear case is calculated from Eq. (12) to be $\langle\tau\rangle_L = 7.56$ ps. In the nonlinear case, we expect this value to be reduced. We introduce the reduction factor as $r = \langle\tau\rangle_{NL}/\langle\tau\rangle_L$ and use it to estimate the parameter E_b in Eq. (15). Figure 1 shows how the probability density $p(\tau)$ changes as a function of τ, as r is reduced from 1 to 0.7. Figure 2 shows how the pulse-width distribution $p(T_p)$ changes as a function of T_p under the same conditions. As expected, both distributions are narrower for solitons. This feature indicates that solitons resist the PMD effects and experience much smaller PRD-induced broadening.

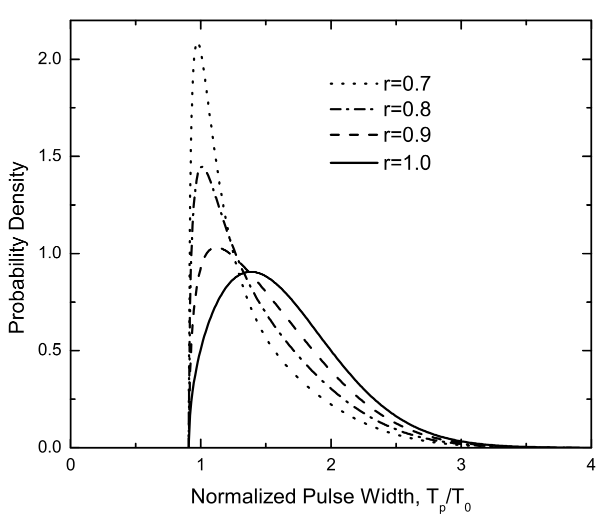

FIG. 2. Probablity density $p(T_p)$ as a function of the rms pulse width T_p for several values of the reduction factor r. The linear case corresponds to $r = 1$. All parameter values are the same as in Fig. 1.

In conclusion, we have developed a nonlinear theory of PMD for solitons propagating inside optical fibers. The theory converts the pair of coupled NLS satisfied by the orthogonally polarized components into a Fokker-Planck equation using the collective-variable approach. We solve this equation in the asymptotic limit (long fiber lengths) and derive expressions for the probability density functions associated with the DGD and the pulse width in the limit of large propagation distances. The predictions of our theory agree with the experimental data.

This research is supported by the National Science Foundation under Grants No. NSF ECS-0123419 and No. DMS-0073923.

[1] G. P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, San Diego, CA, 2001), 3rd ed.
[2] G. P. Agrawal, *Fiber-Optic Communication Systems* (Wiley, New York, 2002), 3rd ed.
[3] C. D. Poole and J. Nagel, in *Optical Fiber Telecommunications III*, edited by I. P. Kaminow and T. L. Koch (Academic Press, San Diego, CA, 1997), Vol. A, Chap. 6.
[4] B. Huttner, C. Geiser, and N. Gisin, IEEE J. Sel. Areas Commun. **6**, 317 (2000).
[5] H. Kogelnik, R. M. Jopson, and L. E. Nelson, *Optical Fiber Telecommunications IV*, edited by I. P. Kaminow and T. Li (Academic Press, San Diego, CA, 2002), Vol. B, Chap. 6.
[6] H. Sunnerud, J. Li, C. J. Xie, and P. A. Andrekson, IEEE Photonics Technol. Lett. **13**, 118 (2001).
[7] S. G. Evangelides, L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, J. Lightwave Technol. **10**, 28 (1992).
[8] P. K. A. Wai and C. R. Menyuk, J. Lightwave Technol. **14**, 148 (1996).
[9] P. K. A. Wai, W. Kath, C. R. Menyuk, and J. W. Zhang, J. Opt. Soc. Am. B **14**, 2967 (1997).
[10] Y. Chen and H. A. Haus, Opt. Lett. **25**, 290 (2000); Chaos **10**, 529 (2000).
[11] C. J. Xie, M. Karlsson, and P. A. Andrekson, IEEE Photonics Technol. Lett. **12**, 801 (2000).
[12] I. Nishioka, T. Hirooka, and A. Hasegawa, IEEE Photonics Technol. Lett. **12**, 1480 (2000).
[13] C. J. Xie, H. Sunnerud, M. Karlsson, and P. A. Andrekson, IEEE Photonics Technol. Lett. **13**, 1079 (2001).
[14] S. Chandrashekhar, Rev. Mod. Phys. **15**, 1 (1943).
[15] L. Faddeev and L. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Springer, New York, 1987).
[16] T. H. R. Skyrme, Proc. R. Soc. London A **260**, 127 (1961); Nucl. Phys. 31, 556 (1962); J. Math. Phys. (N.Y.) **12**, 1735 (1971).
[17] R. Rajaraman, *Solitons and Instantons* (North-Holland, Amsterdam, 1987).
[18] N. S. Manton, Phys. Rev. Lett. **60**, 1916 (1988).
[19] T. Ueda and W. L. Kath, J. Opt. Soc. Am. **11**, 818 (1994); Physica (Amsterdam) **D55**, 166 (1992).
[20] The constant γ can be calculated in principle, but such a calculation is beyond the scope of this Letter.