J-QE/25/12//30964

Pulse Propagation in an Elliptically Birefringent Kerr Medium

Curtis R. Menyuk

Reprinted from
IEEE JOURNAL OF QUANTUM ELECTRONICS
Vol. 25, No. 12, December 1989

Pulse Propagation in an Elliptically Birefringent Kerr Medium

CURTIS R. MENYUK, SENIOR MEMBER, IEEE

Abstract—The coupled nonlinear Schrodinger equation, which describes optical propagation in a birefringent Kerr medium, is derived. It is shown that when the ellipticity angle $\theta \simeq 35°$, Manakov's equation results. Consequences for switching applications are discussed. In particular, if $\theta \neq 35°$, shadows form when two pulses of opposite polarization interact, i.e., the emerging pulses no longer have their original polarizations. This problem disappears at $\theta = 35°$.

I. Introduction

In a Kerr medium, nonlinear wave evolution is due to the $\chi^{(3)}$ response. The $\chi^{(2)}$ response is zero or can be neglected. Examples include certain $LiNbO_3$ and GaAs waveguides and optical fibers. We focus on a pair of orthogonally polarized eigenmodes which are degenerate when birefringence can be neglected. In a single-mode optical fiber, these correspond to the only modes which can propagate. They couple to each other nonlinearly through the Kerr effect.

The nature and strength of the coupling depends critically on the birefringence. In linearly birefringent fibers [1], the cross coupling between modes is only two thirds as strong as the self-coupling. There is another coupling term which at low birefringence leads to ellipse rotation. In circularly birefringent fibers [2], the cross coupling is twice as strong as the self-coupling. These results suggest that there might be an ideal elliptical birefringence at which the self- and cross coupling are identical. That is indeed the case when the ellipticity angle $\theta \simeq 35°$, as we will show.

This cross coupling may play an important role in switching applications. One possible configuration, shown schematically in Fig. 1, has been proposed by Lattes *et al.* [3] and LaGasse *et al.* [4]. A signal pulse of one polarization is divided in two with each portion going down one arm of a Mach–Zender interferometer. In one of the arms, a switching pulse may be introduced at the other polarization. This switching pulse then shifts the phase of the signal pulse in that arm so that when the signal pulses in both arms recombine, they interfere destructively rather than constructively. In its optical fiber implementation [4], the two arms of the interferometer must be temporally, rather than spatially, separated in order to avoid the effect of differing parameter fluctuations in the two arms. One of the signal pulses is temporally delayed, the two signal pulses go through the same fiber, and the delayed pulse is advanced before the two pulses are recombined. To conveniently advance and delay one of the signal pulses without affecting the other, they should have opposite polarizations. The switching pulse must have the opposite polarization from the signal pulse whose phase it is altering.

Manuscript received April 3, 1989; revised July 21, 1989.

The author is with the Department of Electrical Engineering, University of Maryland, Baltimore, MD 21228.

IEEE Log Number 8930964.

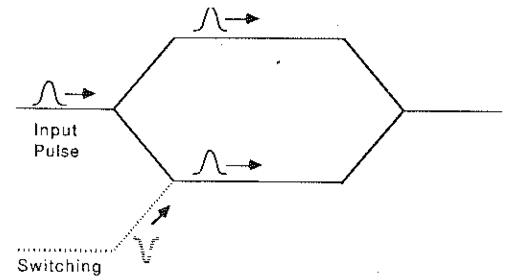


Fig. 1. Schematic illustration of a possible Kerr switch configuration. The signal pulse, shown solid, travels down the two arms of a Mach–Zender interferometer. A switching pulse, shown dashed, changes the phase of the signal pulse in the lower arm.

Other switching configurations have been proposed with varying advantages and drawbacks [5]–[9]. A potential drawback with the scheme just proposed is that when the switching pulse moves through the signal pulse, the switching pulse not only shifts the phase of the signal pulse, but it can leave behind a shadow—a portion of the signal pulse which is in the same polarization as the switching pulse. This shadow is slaved to the signal pulse and does not separate from it. Shadows have been observed numerically [10].

Shadows will no longer form when the medium is elliptically birefringent so that the self-coupling and cross coupling are equal. In this case, Manakov [11] has shown that the coupled nonlinear Schrodinger equation which governs the wave evolution can be solved using nonlinear spectral transform methods [12]. As a consequence, solitons exist. When a soliton of one polarization interacts with an arbitrarily shaped pulse of the opposite polarization, it will undergo a uniform phase shift with some spatial displacement, but will suffer no change in polarization or shape. As long as the signal pulse is a soliton, the shape of the switching pulse can be chosen for convenience.

An elliptically birefringent Kerr medium can be obtained, for example, by twisting an appropriately doped optical fiber preform during the drawing stage [13]–[15].

0018-9197/89/1200-2674$01.00 © 1989 IEEE

The remainder of this paper is organized as follows. In Section II, we derive in detail the coupled nonlinear Schrodinger equation. We discuss, in particular, the physical approximations which are made and the limits in which they hold. In Section III, we study the Manakov equation, considering both the application to switches and possible experiments to study the basic physical phenomena. Section IV contains the conclusions and acknowledgment.

II. Coupled Nonlinear Schrodinger Equation

For simplicity, we will present here a derivation of the basic equations which assumes plane wave propagation. This approach allows us to elucidate the basic physical issues without considering transverse geometric effects. The method for taking into account the detailed geometry is well known [16], and these geometric effects do not change the basic structure of the equations, but merely alter their coefficients somewhat.

The starting point is Maxwell's wave equation which may be written for plane waves

$$\frac{\partial^2 \boldsymbol{E}}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{D}}{\partial t^2} = 0 \tag{1}$$

where $\boldsymbol{E}$ is the electric field, $\boldsymbol{D}$ is the dielectric response, c is the speed of light, and z and t are propagation distance and time. We stress that both $\boldsymbol{E}$ and $\boldsymbol{D}$ are the observed fields which are real, not complex.

Our first goal is to relate $\boldsymbol{D}$ to $\boldsymbol{E}$. We write, as usual,

$$\boldsymbol{D} = \boldsymbol{E} + 4\pi \boldsymbol{P} \tag{2}$$

where $\boldsymbol{P}$ is the polarizability. We shall assume that the linear response of the medium is anisotropic, so that the medium is birefringent along the z direction. Hence, considering only the linear response, $\boldsymbol{P}$ and $\boldsymbol{E}$ are related through a tensor $\boldsymbol{\chi}$, such that

$$\boldsymbol{P}(z, t) = \int_{-\infty}^{t} \boldsymbol{\chi}(t - t') \cdot \boldsymbol{E}(z, t')\, dt'. \tag{3}$$

The nonlinear response will be treated separately. It is a consequence of causality that $\boldsymbol{P}$ at time t can only depend on $\boldsymbol{E}$ at earlier times.

We now consider the Fourier transforms of $\boldsymbol{E}$, $\boldsymbol{P}$, and $\boldsymbol{\chi}$. In general, given a quantity $X(z, t)$, we shall define the transform $\tilde{X}(z, \omega)$ such that

$$\tilde{X}(z, \omega) = \int_{-\infty}^{\infty} X(z, t)\, e^{i\omega t}\, dt, \tag{4}$$

from which it follows that

$$X(z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{X}(z, \omega)\, e^{-i\omega t}\, d\omega. \tag{5}$$

We also define the quantities

$$\tilde{X}^+(z, \omega) = \begin{cases} \tilde{X}(z, \omega), & \omega > 0 \\ 0, & \omega < 0 \end{cases} \tag{6}$$

and $\tilde{X}^-(z, \omega) = \tilde{X}(z, \omega) - \tilde{X}^+(z, \omega)$. The corresponding quantities $X^+(z, t)$ and $X^-(z, t)$ are then defined by (5). Although $X(z, t) = X^+(z, t) + X^-(z, t)$ is real, X^+ and X^- are individually complex and conjugate to each other.

In this paper, we will assume that the nonzero contribution to $\boldsymbol{E}$ and $\boldsymbol{D}$ comes from a small region in ω space surrounding some carrier frequency ω_0 and another small region surrounding its opposite $-\omega_0$. The reality of $E(z, t)$ implies $\tilde{E}(z, -\omega) = \tilde{E}^*(z, \omega)$. Hence, if $E(z, \omega)$ is nonzero near $\omega = \omega_0$, it must be nonzero near $\omega = -\omega_0$. Viewed from this perspective, the standard "trick" of using complex fields and then adding the complex conjugate at the end of the calculation is equivalent to simply considering positive frequency components in the small region around $\omega = \omega_0$ and neglecting the negative frequency components which may then be restored at the end of the calculation.

Since (3) is in the form of an autocorrelation, it may be rewritten

$$\tilde{\boldsymbol{P}}^+(z, \omega) = \tilde{\boldsymbol{\chi}}^+(\omega) \cdot \tilde{\boldsymbol{E}}^+(z, \omega). \tag{7}$$

At any frequency ω, $\tilde{\boldsymbol{\chi}}^+$ will have two orthonormal eigenvectors $\hat{e}_1$ and $\hat{e}_2$ which satisfy the relations

$$\hat{e}_1 \cdot \hat{e}_1^* = \hat{e}_2 \cdot \hat{e}_2^* = 1, \qquad \hat{e}_1 \cdot \hat{e}_2^* = 0. \tag{8}$$

Analogous quantities indicating the eigenvectors of $\tilde{\boldsymbol{\chi}}^-$ may be defined. Writing now

$$\begin{aligned} \tilde{\boldsymbol{E}}^+ &= \tilde{E}_1^+ \hat{e}_1 + \tilde{E}_2^+ \hat{e}_2 \\ \tilde{\boldsymbol{P}}^+ &= \tilde{P}_1^+ \hat{e}_1 + \tilde{P}_2^+ \hat{e}_2, \end{aligned} \tag{9}$$

we obtain

$$\tilde{P}_1^+ = \chi_1 \tilde{E}_1^+, \quad \tilde{P}_2 = \chi_2 \tilde{E}_2^+ \tag{10}$$

where χ_1 and χ_2 are the eigenvalues corresponding to $\hat{e}_1$ and $\hat{e}_2$. Specifying $\chi_1(\omega)$, $\chi_2(\omega)$, $\hat{e}_1(\omega)$, and $\hat{e}_2(\omega)$ is equivalent to specifying $\boldsymbol{\chi}(\omega)$. The linear dispersion relations corresponding to the eigenmodes are given by

$$\begin{aligned} k(\omega) &= \frac{\omega}{c}\left[1 + 4\pi\tilde{\chi}_1(\omega)\right]^{1/2} \\ l(\omega) &= \frac{\omega}{c}\left[1 + 4\pi\tilde{\chi}_2(\omega)\right]^{1/2}. \end{aligned} \tag{11}$$

Equations (7) and (11) are general and do not depend on the assumption that $\boldsymbol{E}^+(\omega)$ is zero outside a small range surrounding $\omega = \omega_0$ frequency.

We now use this assumption, and we also suppose that within this frequency range, we may set $\hat{e}_1(\omega) = \hat{e}_1(\omega_0)$ and $\hat{e}_2(\omega) = \hat{e}_2(\omega_0)$ which is equivalent to ignoring linear mode coupling. It follows from this latter assumption that

$$P_1^+(z, t) = \int_{-\infty}^{t} \chi_1(t - t')\, E_1^+(z, t')\, dt' \tag{12a}$$

$$P_2^+(z, t) = \int_{-\infty}^{t} \chi_2(t - t')\, E_2^+(z, t')\, dt'. \tag{12b}$$

We now write

$$P_1^+(z, t) = \rho(z, t)\, e^{ik_0 z - i\omega_0 t},$$
$$E_1^+(z, t) = U(z, t)\, e^{ik_0 z - i\omega_0 t} \tag{13}$$

where $k_0 = k(\omega_0)$ is determined from the dispersion relation, (11). Because the spectra of P_1^+ and E_1^+ are concentrated near $\omega = \omega_0$, the spectra of ρ and U are concentrated near $\omega = 0$. In effect, ρ and U are the envelopes of P_1^+ and E_1^+; the rapid variation at frequency ω_0 has been removed. One then finds

$$\rho(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\chi}_1(\omega + \omega_0)\, \tilde{U}(z, \omega)\, e^{-i\omega t}\, d\omega. \tag{14}$$

Since $\tilde{U}$ is zero outside a small region surrounding $\omega = 0$, we may approximate $\tilde{\chi}_1$ by its Taylor expansion

$$\tilde{\chi}_1(\omega + \omega_0) \simeq \tilde{\chi}_1(\omega_0) + \chi_1'(\omega_0)\,\omega + \tfrac{1}{2}\chi_1''(\omega_0)\,\omega^2 \tag{15}$$

where $\chi_1'(\omega_0) = d\chi_1/d\omega$ and $\chi''(\omega_0) = d^2\chi_1/d\omega^2$, both evaluated at $\omega = \omega_0$. Substituting (15) into (14) and evaluating the Fourier transform, we obtain

$$\rho(z, t) = \tilde{\chi}_1 U(z, t) + i\tilde{\chi}_1' \frac{\partial U(z, t)}{\partial t} - \frac{1}{2}\tilde{\chi}_1'' \frac{\partial^2 U(z, t)}{\partial t^2} \tag{16}$$

where $\tilde{\chi}_1$, $\tilde{\chi}_1'$, and $\tilde{\chi}_1''$ are all evaluated at $\omega = \omega_0$. If we let

$$D_1^+(z, t) = \Delta(z, t)\, e^{ik_0 z - i\omega_0 t}, \tag{17}$$

we then find

$$\Delta(z, t) = \tilde{\epsilon}_1 U(z, t) + i\tilde{\epsilon}_1' \frac{\partial U(z, t)}{\partial t} - \frac{1}{2}\tilde{\epsilon}_1'' \frac{\partial^2 U(z, t)}{\partial t^2} \tag{18}$$

where

$$\tilde{\epsilon}_1(\omega) = 1 + 4\pi\tilde{\chi}_1(\omega) \tag{19}$$

and its derivatives are evaluated at $\omega = \omega_0$. We have thus determined D_1^+ in terms of E_1^+. We may similarly determine D_2^+ in terms of E_2^+. Noting that $D_1^- = D_1^{+*}$ and $D_2^- = D_2^{+*}$, we see that the linear portion of $\boldsymbol{D}$ is completely determined in terms of $\boldsymbol{E}$.

In (11), we have written the dispersion relation which corresponds to forward-propagating waves. Since Maxwell's wave equation, (1), is second order in z, it will also have backward-propagating solutions which correspond to choosing a negative sign in (11). It is, however, a consequence of our assumption that the frequency spectrum of E_1^+ is concentrated near $\omega = \omega_0$ that the wavenumber spectrum is concentrated near $k = k_0$ or $k = -k_0$. In other words, except for a brief transient, an optical pulse consists entirely of forward-going or backward-going waves. We shall assume that the optical pulse consists of forward-going waves with no loss in generality.

It is now possible to reduce Maxwell's wave equation so that it is only first order in z. Substituting the expressions for E_1^+ and D_1^+ into Maxwell's wave equation, one obtains

$$-k_0^2 U + 2ik_0 \frac{\partial U}{\partial z} + \frac{\partial^2 U}{\partial z^2} + \frac{\omega_0^2}{c^2}\epsilon U + i\left(\frac{\omega_0^2}{c^2}\epsilon' + 2\frac{\omega_0}{c^2}\epsilon\right)\frac{\partial U}{\partial t} - \left(\frac{\omega_0^2}{2c^2}\epsilon'' + 2\frac{\omega_0}{c^2}\epsilon' + \frac{1}{c^2}\epsilon''\right)\frac{\partial^2 U}{\partial t^2} = 0. \tag{20}$$

In (20), only second- and lower order time derivatives have been kept, consistent with the earlier expansion of Δ in terms of U where only derivatives up to second order were kept. The goal is to eliminate the term $\partial^2 U/\partial z^2$ in favor of a term containing only time derivatives. Because U is slowly varying in time and space, we conclude

$$|k_0^2 U| \gg \left|k_0 \frac{\partial U}{\partial z}\right| \gg \left|\frac{\partial^2 U}{\partial z^2}\right|, \tag{21}$$

and similarly,

$$|\omega_0^2 U| \gg \left|\omega_0 \frac{\partial U}{\partial t}\right| \gg \left|\frac{\partial^2 U}{\partial t^2}\right|. \tag{22}$$

We may thus expand (20) in order of the number of derivatives. At lowest order, this procedure yields

$$k_0^2 - \frac{\omega_0^2}{c^2}\epsilon = 0, \tag{23}$$

which in essence fixes the linear dispersion relation, (11). At next order, one finds

$$i\frac{\partial U}{\partial z} + i\left(\frac{\omega_0^2}{2k_0 c^2}\epsilon' + \frac{\omega_0}{k_0 c^2}\epsilon\right)\frac{\partial U}{\partial t} \equiv i\frac{\partial U}{\partial z} + ik'\frac{\partial U}{\partial t} = 0 \tag{24}$$

where $k' = \partial k/\partial\omega$ is evaluated from the dispersion relation, (11), at $\omega = \omega_0$. At this order, the overall motion of an optical pulse is determined to be $v_g = 1/k'$. From (24), we find that to second order

$$\frac{\partial^2 U}{\partial z^2} = (k')^2 \frac{\partial^2 U}{\partial t^2}. \tag{25}$$

Substituting (25) into (20), we conclude

$$i\frac{\partial U}{\partial z} + ik'\frac{\partial U}{\partial t} - \left[\frac{\omega_0^2}{4k_0 c^2}\epsilon'' + \frac{\omega_0}{k_0 c^2}\epsilon' + \frac{1}{2k_0 c^2}\epsilon - \frac{1}{2k_0}(k')^2\right]\frac{\partial^2 U}{\partial t^2} \equiv i\frac{\partial U}{\partial z} + ik'\frac{\partial U}{\partial t} - \frac{1}{2}k''\frac{\partial^2 U}{\partial t^2} = 0 \tag{26}$$

where $k'' = \partial^2 k/\partial\omega^2$ is evaluated at $\omega = \omega_0$. In similar fashion, if we let

$$V(z, t) = E_2^+(z, t)\, e^{il_0 z - i\omega_0 t} \tag{27}$$

where $l_0 = l(\omega)$ is evaluated at $\omega = \omega_0$, it follows from (11) that

$$i\frac{\partial V}{\partial z} + il'\frac{\partial V}{\partial t} - \frac{1}{2}l''\frac{\partial^2 V}{\partial t^2} = 0 \tag{28}$$

where $l' = \partial l/\partial\omega$ and $l'' = \partial^2 l/\partial\omega^2$ are evaluated at $\omega = \omega_0$. The terms which contain second derivatives in time lead to pulse spreading or dispersion.

It is no accident that the coefficients of the time derivatives just involve derivatives of the dispersion relation. This result can be made apparent by using a Green's function or the Fourier–Laplace transform approach. The Fourier transform of Maxwell's wave equation for $\tilde{E}_1^+$ yields

$$\frac{\partial^2 \tilde{E}_1^+}{\partial z^2} + k^2(\omega)\tilde{E}_1^+. \tag{29}$$

If we write the Laplace transform in the form

$$\bar{E}_1^+(k, \omega) = \int_0^\infty \tilde{E}_1(z, \omega)\, e^{-ikz}\, dz \tag{30}$$

where $\mathrm{Im}\,(k) < 0$, we find that (29) becomes

$$\left[k^2(\omega) - k^2\right]\bar{E}_1^+ = \tilde{E}_0' + ik\tilde{E}_0 \tag{31}$$

where $\tilde{E}_0' = \partial\tilde{E}(z, \omega)/\partial z$ and $\tilde{E}_0 = E(z, \omega)$ are evaluated at $z = 0$. Demanding that our light pulse consist of only forward-going waves is equivalent to demanding that $E_0' = ik(\omega)E_0$. In this case, (31) becomes

$$i\left[k - k(\omega)\right]\bar{E}_1^+ = \tilde{E}_0. \tag{32}$$

Expanding $k(\omega)$ in a Taylor series about the frequency $\omega = \omega_0$ yields

$$i\left[(k - k_0) - k'(\omega - \omega_0) - \tfrac{1}{2}k''(\omega - \omega_0)^2\right]\bar{E}_1^+ \simeq \tilde{E}_0. \tag{33}$$

Using the definition of $U(z, t)$, one may verify that (33) is just the Fourier–Laplace transform of (26). This approach yields the linear wave equation more easily than the approach previously described where one directly eliminates $\partial^2 U/\partial z^2$ in the z domain; however, this approach does not generalize in any simple way to nonlinear problems, while the previous approach does.

We turn now to consideration of the nonlinear contribution to the polarizability $\boldsymbol{P}$. We shall suppose that no second-order nonlinearity appears, so that the lowest order nonlinearity is third order. We shall also suppose that the medium is only weakly anisotropic, so that the nonlinear response can be considered isotropic. Leaving aside second harmonic generation, which only appears in special circumstances, both these assumptions apply to optical fibers. The nonlinear polarizability must have the form

$$\begin{aligned}\boldsymbol{P}(z, t) = &\int_{-\infty}^t dt_1 \int_{-\infty}^t dt_2 \int_{-\infty}^t \\ &\cdot dt_3\, \chi(t - t_1, t - t_2; t - t_3) \\ &\cdot \left[\boldsymbol{E}(z, t_1)\cdot\boldsymbol{E}(z, t_2)\right]\boldsymbol{E}(z, t_3).\end{aligned} \tag{34}$$

This combination of $\boldsymbol{E}$ vectors is the only combination which is invariant under rotations and mirror reflections. From the form of (34), it follows that $\chi(\tau_1, \tau_2; \tau_3)$ is invariant under the interchange $\tau_1 \leftrightarrow \tau_2$, but not necessarily under the interchanges $\tau_1 \leftrightarrow \tau_3$ and $\tau_2 \leftrightarrow \tau_3$. Since the spectrum of $\boldsymbol{E}$ is concentrated primarily in small spectral regions surrounding $\omega = \omega_0$ and $\omega = -\omega_0$, it follows that $\boldsymbol{P}$ will be concentrated primarily in spectral regions surrounding $\omega = -3\omega_0$, $-\omega_0$, ω_0, and $3\omega_0$. Assuming, as is certainly the case in fibers, that if waves propagate at $\omega = \omega_0$ they cannot propagate at $\omega = 3\omega_0$, we can ignore the contributions of $\boldsymbol{P}$ at $\pm 3\omega_0$ to the electric field. Designating $\boldsymbol{P}^+$ as the contribution to $\boldsymbol{P}$ concentrated near $\omega = \omega_0$, we find that it consists of all combinations of the $\boldsymbol{E}$ field containing two + and one − contribution. It follows that

$$\begin{aligned}\boldsymbol{P}^+(z, t) = &\int_{-\infty}^t dt_1 \int_{-\infty}^t dt_2 \int_{-\infty}^t \\ &\cdot dt_3\, \chi(t - t_1, t - t_2; t - t_3) \\ &\cdot \left\{2\left[\boldsymbol{E}^+(z, t_1)\cdot\boldsymbol{E}^-(z, t_2)\right]\boldsymbol{E}^+(z, t_3) \right.\\ &\left.+ \left[\boldsymbol{E}^+(z, t_1)\cdot\boldsymbol{E}^+(z, t_2)\right]\boldsymbol{E}^-(z, t_3)\right\}.\end{aligned} \tag{35}$$

In the Fourier domain, (35) becomes

$$\begin{aligned}\tilde{\boldsymbol{P}}^+(z, \omega) = &\frac{1}{(2\pi)^2}\int_{-\infty}^\infty d\omega_1 \int_{-\infty}^\infty d\omega_2\, \tilde{\chi}(\omega_1, \omega_2; \omega_3) \\ &\cdot \left\{2\left[\tilde{\boldsymbol{E}}^+(z, \omega_1)\cdot\tilde{\boldsymbol{E}}^-(z, \omega_2)\right]\tilde{\boldsymbol{E}}^+(z, \omega_3) \right.\\ &\left.+ \left[\tilde{\boldsymbol{E}}^+(z, \omega_1)\cdot\tilde{\boldsymbol{E}}^+(z, \omega_2)\right]\tilde{\boldsymbol{E}}^-(z, \omega_3)\right\}\end{aligned} \tag{36}$$

where $\omega_3 = \omega - \omega_1 - \omega_2$. The first term in (36) is concentrated in the spectral region $\omega_1 = -\omega_2 = \omega_3 = \omega_0$. The second term in (36) is concentrated in the spectral region $\omega_1 = \omega_2 = -\omega_3 = \omega_0$. If we Taylor expand $\tilde{\chi}(\omega_1, \omega_2; \omega_3)$ just as when deriving the linear response, but retain only the lowest order contribution, we obtain

$$\begin{aligned}\tilde{\boldsymbol{P}}^+(z, \omega) = &\frac{1}{(2\pi)^2}\int_{-\infty}^\infty d\omega_1 \int_{-\infty}^\infty d\omega_2 \left\{2a\left[\tilde{\boldsymbol{E}}^+(z, \omega_1)\right.\right. \\ &\left.\cdot\, \tilde{\boldsymbol{E}}^-(z, \omega_2)\right]\tilde{\boldsymbol{E}}^+(z, \omega_3) \\ &\left.+ b\left[\tilde{\boldsymbol{E}}^+(z, \omega_1)\cdot\tilde{\boldsymbol{E}}^+(z, \omega_2)\right]\tilde{\boldsymbol{E}}^-(z, \omega_3)\right\}\end{aligned} \tag{37}$$

where $a = \tilde{\chi}(\omega_0, -\omega_0; \omega_0)$ and $b = \tilde{\chi}(\omega_0, \omega_0; -\omega_0)$. Neglecting higher order contributions is equivalent to neglecting the contribution to the nonlinear polarizability of terms which include the factors $\partial U/\partial t$, $\partial V/\partial t$, and higher order envelope derivatives. This assumption is valid in optical fibers as long as the optical pulses are longer than several hundred femtoseconds. In the opposite limit, terms containing envelope derivatives lead to the Raman self-frequency shift, a physically important phenomenon [17].

Returning to the time domain yields

$$P^+(z, t) = 2a[E^+(z, t) \cdot E^-(z, t)] E^+(z, t) + b[E^+(z, t) \cdot E^+(z, t)] E^-(z, t). \quad (38)$$

It should be emphasized that in deriving (38), we have assumed that the field *envelopes* vary slowly compared to the dielectric response time, not the E field itself. When the dielectric response times are so fast that they may be regarded as instantaneous, i.e., they are much greater than ω_0^{-1}, then

$$\tilde{\chi}(\omega_0, -\omega_0; \omega_0) = \tilde{\chi}(\omega_0, \omega_0; -\omega_0) = \tilde{\chi}(0, 0; 0) \quad (39)$$

so that $a = b$ and the number of independent Kerr coefficients is reduced from two to one. To make this point explicit, we return to (35) and note that if $\chi(\tau_1, \tau_2; \tau_3) \to 0$ so rapidly that the variation of E can be neglected, then

$$P^+(z, t) = \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t} \cdot dt_3\, \chi(t - t_1, t - t_2, t - t_3) \cdot \left\{2[E^+(z, t) \cdot E^-(z, t)] E^+(z, t) + [E^+(z, t) \cdot E^+(z, t)] E^-(z, t)\right\}. \quad (40)$$

Noting that

$$\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t} dt_3\, \chi(t - t_1, t - t_2; t - t_3) = \tilde{\chi}(0, 0; 0) \quad (41)$$

and comparing (40) to (38), we arrive at (39). In optical fibers, the nonlinear dielectric response can be viewed as instantaneous and one does find that $a = b$. The reduction in the number of independent Kerr coefficients from two to one when the nonlinear dielectric response becomes instantaneous is implicit in the previous results of Maker and Terhune [18].

Recalling that the unit vectors $\hat{e}_1 \equiv \hat{e}_1(\omega_0)$ and $\hat{e}_2 \equiv \hat{e}_2(\omega_0)$ define the eigenmodes, and using the orthogonality relations, (8), we find

$$P_1^+ = 2a[E_1^+E_1^- + E_2^+E_2^-]E_1^+ + b[E_1^+E_1^+(\hat{e}_1 \cdot \hat{e}_1) + 2E_1^+E_2^+(\hat{e}_1 \cdot \hat{e}_2) + E_2^+E_2^+(\hat{e}_2 \cdot \hat{e}_2)] \cdot [E_1^-(\hat{e}_1^* \cdot \hat{e}_1^*) + E_2^-(\hat{e}_1^* \cdot \hat{e}_2^*)]. \quad (42)$$

The term in which a appears does not depend on the eigenmode structure, but the term in which b appears does. Hence, the strength of the nonlinear mode coupling will depend on the eigenmode structure. In general, a Kerr medium may be elliptically birefringent. Choosing (with no loss of generality) $\hat{e}_x$ along the major axis of the birefringence ellipse and $\hat{e}_y$ along the minor axis of the birefringence ellipse, we find that we may write

$$\hat{e}_1 = \frac{\hat{e}_x + ir\hat{e}_y}{\sqrt{1 + r^2}}, \quad \hat{e}_2 = \frac{r\hat{e}_x - i\hat{e}_y}{\sqrt{1 + r^2}} \quad (43)$$

where, letting $r = \tan(\theta/2)$, we have that $\hat{e}_1 \cdot \hat{e}_1 = \hat{e}_2 \cdot \hat{e}_2 = \cos\theta$ and $\hat{e}_1 \cdot \hat{e}_2 = \sin\theta$. A linearly birefringent fiber corresponds to $r = 0$ and $\theta = 0$, while a circularly birefringent fiber corresponds to $r = 1$ and $\theta = \pi/2$. In the former case, one finds $\hat{e}_1 = \hat{e}_x$ and $\hat{e}_2 = -i\hat{e}_y$. This choice of eigenvectors differs from the usual choice, $\hat{e}_1 = \hat{e}_x$ and $\hat{e}_2 = \hat{e}_y$, but this difference leads to no change in the evolution equations. In the latter case, one finds $\hat{e}_1 = (\hat{e}_x + i\hat{e}_y)/\sqrt{2}$ and $\hat{e}_2 = (\hat{e}_x - i\hat{e}_y)/\sqrt{2}$ which is standard. Noting that $E_1^- = E_1^{+*}$ and $E_2^- = E_2^{+*}$, (42) now becomes

$$P_1^+ = (2a + b\cos^2\theta)|E_1^+|^2E_1^+ + (2a + 2b\sin^2\theta)|E_2^+|^2E_1^+ + b\cos\theta\sin\theta(E_1^+)^2E_2^- + 2b\cos\theta\sin\theta|E_1^+|^2E_2^+ + b\cos^2\theta(E_2^+)^2E_1^- + b\cos\theta\sin\theta|E_2^+|^2E_2^+. \quad (44)$$

Using the definitions in (13) and (27) for the wave envelopes, we conclude

$$\rho = (2a + b\cos^2\theta)|U|^2U + (2a + 2b\sin^2\theta)|V|^2U + b\cos^2\theta V^2U^*\exp[-2i(k_0 - l_0)z] + b\cos\theta\sin\theta\{U^2V^*\exp[i(k_0 - l_0)z] + (2|U|^2 + |V|^2)V\exp[-i(k_0 - l_0)z]\}. \quad (45)$$

We may now combine (45) which gives the nonlinear polarizability with (16) which gives the linear polarizability and substitute the result into Maxwell's wave equation. We assume that the nonlinear contribution is of the same order as the dispersive contribution because solitons are obtained when these two contributions balance. Substituting the total polarizability into Maxwell's wave equation and reducing the equation so that it is first order in z, just as in the strictly linear case, we obtain

$$i\frac{\partial U}{\partial z} + ik'\frac{\partial U}{\partial t} - \frac{1}{2}k''\frac{\partial^2 U}{\partial t^2} + (2a' + b'\cos^2\theta)|U|^2U + (2a' + 2b'\sin^2\theta)|V|^2U + b'\cos^2\theta V^2U^* \cdot \exp[-2i(k_0 - l_0)z] + b'\cos\theta\sin\theta\{U^2V^*\exp[i(k_0 - l_0)z] + (2|U|^2 + |V|^2)V\exp[-i(k_0 - l_0)z]\} = 0 \quad (46)$$

where $a' = a/2k_0$ and $b' = b/2k_0$. In a similar fashion, it can be shown that

$$
\begin{aligned}
i\frac{\partial V}{\partial z} &+ il'\frac{\partial V}{\partial t} - \frac{1}{2}l''\frac{\partial^2 V}{\partial t^2} + (2a' + 2b'\sin^2\theta)\,|U|^2 V \\
&+ (2a' + b'\cos^2\theta)\,|V|^2 V + b'\cos^2\theta U^2 V^* \\
&\cdot \exp\left[2i(k_0 - l_0)z\right] \\
&+ b'\cos\theta\sin\theta\left\{V^2U^*\exp\left[-i(k_0 - l_0)z\right]\right. \\
&\left.+ \left(|U|^2 + 2|V|^2\right)U\exp\left[i(k_0 - l_0)z\right]\right\} = 0
\end{aligned}
\tag{47}
$$

where we assume that $a/2k_0 \simeq a/2l_0$ and $b/2k_0 \simeq b/2l_0$.

We now reduce (46) and (47) to normalized form. To do so, we assume that light is propagating in the anomalous dispersion regime where $k'' < 0$ and $l'' < 0$. We also assume, as is appropriate for optical fibers, that the small difference between k'' and l'' may be neglected and that $k' - l' \simeq (k_0 - l_0)/\omega_0$. Letting

$$k'' = l'' = -\frac{\lambda_0}{2\pi c^2}D(\lambda_0), \tag{48}$$

we define

$$
\begin{aligned}
&\xi = \frac{\pi z}{2z_0}, \quad z_0 = \frac{\pi^2 c^2 t_0^2}{D(\lambda_0)\lambda_0}, \quad t_0 = 0.568\tau, \\
&s = \frac{1}{t_0}\left(t - \frac{z}{\overline{v}_g}\right), \quad \overline{v}_g = \frac{2}{k' + l'}, \\
&u = (2a' + b'\cos^2\theta)^{1/2}U, \quad v = (2a' + b'\cos^2\theta)^{1/2}V, \\
&\delta = \frac{k' - l'}{2|k''|}t_0 = \frac{\pi c\Delta n}{D(\lambda_0)\lambda_0}t_0, \quad R = \frac{8\pi c}{\lambda_0}t_0, \\
&B = \frac{2a' + 2b'\sin^2\theta}{2a' + b'\cos^2\theta}, \quad C = \frac{b'\cos^2\theta}{2a' + b'\cos^2\theta}, \\
&D = \frac{b'\cos\theta\sin\theta}{2a'\cos^2\theta}
\end{aligned}
$$

where τ is the FWHM pulse intensity and $\Delta n = (k_0 - l_0)c/\omega_0$ is the difference between the indexes of refraction. With these definitions, (46) and (47) become

$$
\begin{aligned}
i\frac{\partial u}{\partial \xi} &+ i\delta\frac{\partial u}{\partial s} + \frac{1}{2}\frac{\partial^2 u}{\partial s^2} + \left(|u|^2 + B|v|^2\right)u \\
&+ Cv^2u^*\exp(-iR\delta\xi) + D\left[u^2v^*\exp(iR\delta\xi/2)\right. \\
&\left.+ \left(2|u|^2 + |v|^2\right)v\exp(-iR\delta\xi/2)\right] = 0
\end{aligned}
\tag{49a}
$$

$$
\begin{aligned}
i\frac{\partial v}{\partial \xi} &- i\delta\frac{\partial v}{\partial s} + \frac{1}{2}\frac{\partial^2 v}{\partial s^2} + \left(B|u|^2 + |v|^2\right)v \\
&+ Cu^2v^*\exp(iR\delta\xi) + D\left[v^2u^*\exp(-iR\delta\xi/2)\right. \\
&\left.+ \left(|u|^2 + 2|v|^2\right)u\exp(iR\delta\xi/2)\right] = 0.
\end{aligned}
\tag{49b}
$$

The time variable s is not proportional to t, absolute time in the laboratory frame. Its variation at any point ξ is proportional to time measured in the laboratory frame, but its origin is ξ dependent. The origin is chosen so that if a signal moved at the group velocity intermediate between that of the two modes, the evolution in ξ of its s profile would appear to be frozen. Hence, terms proportional to $\partial u/\partial s$ and $\partial v/\partial s$ appear with opposite sign in (49) to account for the group velocity difference between the two modes. These terms can, in fact, be removed by making the transformation

$$
\begin{aligned}
\overline{u} &= u\exp\left[-i\frac{\delta^2}{2}\xi + i\delta s\right] \\
\overline{v} &= v\exp\left[-i\frac{\delta^2}{2}\xi - i\delta s\right]
\end{aligned}
\tag{50}
$$

which yields the equations

$$
\begin{aligned}
i\frac{\partial \overline{u}}{\partial \xi} &+ \frac{1}{2}\frac{\partial^2 \overline{u}}{\partial s^2} + \left(|\overline{u}|^2 + B|\overline{v}|^2\right)\overline{u} \\
&+ C\overline{v}^2\overline{u}^*\exp(-iR\delta\xi + 4i\delta s) + D\left[\overline{u}^2\overline{v}^*\right. \\
&\cdot\exp(iR\delta\xi/2 - 2i\delta s) \\
&\left.+ \left(2|\overline{u}|^2 + |\overline{v}|^2\right)\overline{v}\exp(-iR\delta\xi/2 + 2i\delta s)\right] = 0 \\
i\frac{\partial \overline{v}}{\partial \xi} &+ \frac{1}{2}\frac{\partial^2 \overline{v}}{\partial s^2} + \left(B|\overline{u}|^2 + |\overline{v}|^2\right)\overline{v} \\
&+ C\overline{u}^2\overline{v}^*\exp(iR\delta\xi - 4i\delta s) + D\left[\overline{u}^2\overline{v}^*\right. \\
&\cdot\exp(-iR\delta\xi/2 + 2i\delta s) \\
&\left.+ \left(2|\overline{u}|^2 + |\overline{v}|^2\right)\overline{u}\exp(iR\delta\xi/2 - 2i\delta s)\right] = 0.
\end{aligned}
\tag{51}
$$

Physically, this transformation is equivalent to shifting the central frequencies of the two modes just far enough apart that their group velocities become equal. The cost is to add explicit time-dependent factors into the equation. The factor δ is quite sizable in most optical fiber experiments, the exception being when fibers with unusually low birefringence or pulse durations substantially shorter than a picosecond are used [1]. Treatments of nonlinear birefringence that have appeared in the literature in which $\cos\theta = 1$ and $\delta = 0$ apply in this limit [19], [20]. From an experimental standpoint, it is convenient to use (49) rather than (51) since the two eigenmodes are usually injected into the fiber with the same central frequency.

In most cases of experimental interest, $R\delta >> 1$. In these cases, the terms in which exponential factors appear in (49) are rapidly oscillating and can be neglected. This assumption corresponds physically to assuming that the birefringent beat length is small compared to the dispersive scale length. Equation (49) now has the extremely simple form

$$
\begin{aligned}
i\frac{\partial u}{\partial \xi} + i\delta\frac{\partial u}{\partial s} + \frac{1}{2}\frac{\partial^2 u}{\partial s^2} + \left(|u|^2 + B|v|^2\right)u &= 0 \\
i\frac{\partial v}{\partial \xi} - i\delta\frac{\partial v}{\partial s} + \frac{1}{2}\frac{\partial^2 v}{\partial s^2} + \left(B|u|^2 + |v|^2\right)v &= 0.
\end{aligned}
\tag{52}
$$

From the definition of B, we find that $B = 1$ when $\cos^2 \theta = 2 \sin^2 \theta$, or, in other words, when

$$\theta \simeq 35°. \tag{53}$$

This result does not depend on the ratio b/a. Since the terms proportional to $\partial u/\partial s$ and $\partial v/\partial s$ can be transformed away, we find that (52) is just a version of Manakov's equation [11] which is integrable using spectral transform methods [12]. In particular, solitons of one polarization should pass through pulses of the opposite polarization without creating shadows.

Before leaving this section, we review the assumptions which led us to (52). These are as follows.

1) The plane wave approximation. Taking into account the transverse variation in the fiber will not change the form of our final equation. It may lead to multiplying the coefficient B in (52) by some θ dependent factor which would shift the critical angle slightly.

2) The slowly varying envelope approximation. This assumption underpins the entire theoretical development. Most of the subsequent assumptions are only reasonable in the context of this assumption.

3) No linear mode coupling. In the regime of interest to us in this paper, linear mode coupling in fibers is extremely weak compared to the nonlinear coupling.

4) Truncation of the Taylor expansions of $k(\omega)$ and $\chi(\omega_1, \omega_2; \omega_3)$. We stopped at second order in the former case and zeroth order in the latter case. From the standpoint of linear propagation, the first-order term governs the overall velocity of the optical pulse, while the second-order term governs its spreading. Both these effects are readily visible. By contrast, the effect of the third-order term is not readily visible unless the second-order term is zero. Since a soliton forms by balancing dispersion with the zeroth-order nonlinearity, higher order nonlinearity will not be visible unless the pulse duration is quite small.

5) Birefringent beat length small compared to dispersive scale length. This assumption is well obeyed under most circumstances in optical fibers.

6) Other simplifying assumptions. These include $k'' = l''$ and $k' - l' = (k_0 - l_0)/\omega_0$. They are both reasonably well obeyed in optical fibers.

III. Manakov Equation

Assuming that $B = 1$, Manakov [11] has shown how to solve (52) using the spectral transform method [12]. We do not describe this approach here as it is rather intricate. Instead, we concentrate on its consequences. One can verify by substitution that (52) has the single soliton solutions

$$u = A_1 \exp(iA_1^2\xi/2) \operatorname{sech} \phi_1$$
$$v = 0 \tag{54}$$

where

$$\phi_1 = A_1(s - s_1 - \delta\xi) \tag{55}$$

and

$$u = 0$$
$$v = A_2 \exp(iA_2^2\xi/2) \operatorname{sech} \phi_2 \tag{56}$$

where

$$\phi_2 = A_2(s - s_2 + \delta\xi). \tag{57}$$

The quantities A_1, A_2, s_1, and s_2 are all arbitrary parameters. Equation (52) has these solutions for *any* value of B. When $B = 1$, one has the following additional result. Given an initial condition

$$u = A_1 \operatorname{sech}[A_1(s - s_1)]$$
$$v = D(s) \tag{58}$$

at $\xi = 0$ where we assume that $D(s)$ is real and that the two polarizations pass through each other, then when ξ is large, we will find

$$u = A_1 \exp(iA_1^2\xi/2 + i\psi_1) \operatorname{sech}(\phi_1 - \Delta_1). \tag{59}$$

In other words, the soliton emerges unscathed from the interaction except for a shift in phase ψ_1 and a shift in time Δ_1/A_1. In particular, the polarization of the soliton is unaltered and shadows do not appear.

As an example, we consider two soliton collisions where the solitons are in opposite polarizations. The solution to (52) which we seek has the form

$$u = \frac{2A_1 \exp(iA_1^2\xi/2)[\exp \phi_2 + \Omega_1 \exp(-\phi_2)]}{\exp(\phi_1 + \phi_2) + \exp(\phi_1 - \phi_2) + \exp[-(\phi_1 - \phi_2)] + |\Omega_1|^2 \exp[-(\phi_1 + \phi_2)]} \tag{60a}$$

$$v = \frac{2A_2 \exp(iA_2^2\xi/2)[\exp \phi_1 + \Omega_2 \exp(-\phi_1)]}{\exp(\phi_1 + \phi_2) + \exp(\phi_1 - \phi_2) + \exp[-(\phi_1 - \phi_2)] + |\Omega_2|^2 \exp[-(\phi_1 + \phi_2)]} \tag{60b}$$

where

$$\Omega_1 = \frac{2\delta + i(A_1 - A_2)}{2\delta + i(A_1 + A_2)}$$

$$\Omega_2 = \frac{2\delta + i(A_1 - A_2)}{2\delta - i(A_1 + A_2)}. \tag{61}$$

When $\phi_1 \simeq 0$ and $\xi << 0$, $\phi_2 << 0$, we find

$$u = \frac{2A_1\Omega_1 \exp(iA_1^2\xi/2)}{\exp(\phi_1) + |\Omega_1|^2 \exp(-\phi_1)} \tag{62}$$

and when $\xi >> 0$, $\phi_2 >> 0$,

$$u = \frac{2A_1 \exp(iA_1^2\xi/2)}{\exp(\phi_1) + \exp(-\phi_1)}. \tag{63}$$

Hence, we find for the phase shift

$$\tan \psi_1 = -\frac{\mathrm{Im}(\Omega_1)}{\mathrm{Re}(\Omega_1)} = \frac{4\delta A_2}{4\delta^2 + A_1^2 - A_2^2} \quad (64)$$

and for the time shift

$$\Delta s = \frac{1}{2A_1} \ln \left[\frac{4\delta^2 + (A_1 + A_2)^2}{4\delta^2 + (A_1 - A_2)^2} \right]. \quad (65)$$

The phase change is largest when $A_2 >> A_1$ in which case $\Delta s \simeq 2/A_2$ and

$$\tan \psi_1 \simeq -\frac{4\delta}{A_2}, \quad (66)$$

corresponding to $\psi_1 \simeq \pi - 4\delta/A_2$. When $A_1 = A_2 = A$, we find

$$\Delta s = \frac{1}{2A} \ln \left(1 + \frac{A^2}{\delta^2}\right), \quad \tan \psi_1 = \frac{A}{\delta} \quad (67)$$

in which case $\psi_1 < \pi/2$.

It is interesting to consider (60) in the limit $\delta = 0$. In this case, the solution is stationary, i.e., $|u|^2$ and $|v|^2$ are independent of ξ. For that reason, this solution has been referred to as a soliton solution [20], but it is really a two-soliton (or two-pole) solution.

The result which we have obtained for ψ_1 in (66) can be understood from an elementary viewpoint. We suppose that the pulse in the v polarization has duration s_0 and amplitude A_2 at the point that it interacts with the soliton in the u polarization. We also suppose that the pulse in the v polarization is intense enough to dominate the evolution of the u polarization while the two polarizations interact. Finally, we suppose that the interaction is swift enough that the pulse in the v polarization does not change its shape. It then follows that

$$\frac{\partial \psi_1}{\partial \xi} = |v|^2 (\xi, s) \quad (68)$$

so that

$$\psi_1 = \frac{1}{2\delta} \int_{-\infty}^{\infty} |v|^2 \, ds \simeq \frac{A_2^2}{2\delta} s_0. \quad (69)$$

For a soliton, $s_0 = 2/A_2$ which yields reasonable agreement with (67). Equation (69) is appropriate for v pulses whose integrated intensities are large and whose s derivatives are small.

The existence of shadows when $B \neq 1$ might appear surprising given the well-known robustness of single solitons when the nonlinear Schrodinger equation is perturbed so that is no longer integrable. Under the influence of non-Hamiltonian perturbations such as attenuation or the Raman self-frequency shift, the soliton parameters change secularly. In the first case, the amplitude steadily diminishes [21], and in the second case, the frequency steadily diminishes [16]. In both cases, the soliton maintains its basic shape. When the perturbation is Hamiltonian with no explicit dependence on space or time, the soliton is almost unaffected. Its shape, speed, and wavenumber shift change somewhat, and that is all. From a fundamental viewpoint, solitons can be regarded as poles in spectral transform space [12]. It is difficult for perturbations to destroy the poles in spectral transform space so solitons continue to exist [22]. By contrast, multiple soliton structures are not robust because each of the individual poles which compose it can shift their locations and strengths, in different ways, leading the structure to break up. In a similar sense, when the Manakov equation is perturbed, individual solitons are robust, but their polarizations are not. When they undergo collisions, the polarization will, in general, change along with the speeds and amplitudes.

We now consider a set of example parameters which can be used to experimentally verify the phase shifts which we have predicted. We consider pulses which are 500 fs long, which is the largest size that can be conveniently produced by the soliton laser [23]. We shall also assume $\delta = 5$, corresponding to fairly large birefringence which will ensure a short interaction length and good control over the birefringence. At this pulse size, $\delta = 5$ corresponds to $\Delta n = 1.9 \times 10^{-4}$ where we have set $\lambda_0 = 1.55$ μm and $D(\lambda_0) = 6.5 \times 10^{-3}$. The birefringence is large, but is substantially smaller than the largest birefringences available [13], [14]. The soliton period is $z_0 = 7.1$ m. If we demand that $\Delta s = 20$ for a complete interaction to occur between a soliton in the u polarization and a pulse in the v polarization, we find

$$\Delta z = \frac{1}{\pi} \frac{\Delta s}{\delta} z_0 = 9 \text{ m}. \quad (70)$$

This estimate of the necessary interaction length is conservative; as little as 4.5 m might suffice. Finally, if we suppose that

$$v = A_2 \operatorname{sech} \left[(s - s_2) + \delta\xi \right] \quad (71)$$

before collision, then we find from (69) that

$$\psi_1 \simeq A_2^2/\delta; \quad (72)$$

so, to obtain a phase shift of π, we conclude $A_2 = 4.0$, corresponding to a pulse containing four solitons. To determine the twist length required to obtain $\theta = 35°$, we must know the strength of the electrooptic tensor. In general, we have

$$\alpha = g\tau \quad (73)$$

where α is the birefringent rotation rate and τ is the mechanical rotation rate. For a pure silica fiber with no linear birefringence, $g = 0.16$, but is higher in a fiber which already has substantial linear birefringence. Noting that

$$\alpha = \frac{\omega_0}{c} \Delta n \sin \theta, \quad (74)$$

we find $\tau \lesssim 2600$ m^{-1}, corresponding to a twist length $\gtrsim 2.3$ mm. This value can be obtained by twisting the fiber as it is drawn in the fabrication process [13]–[15].

IV. Conclusions

We have considered nonlinear pulse propagation in elliptically birefringent Kerr media, with particular emphasis on optical fibers. We have derived a version of the coupled nonlinear Schrodinger equation, and we have shown that when the angle of ellipticity $\theta = 35°$, cross coupling and self-coupling in the Kerr effect become equal, and pulse evolution is described by Manakov's equation for sufficiently large birefringence.

An important potential application of the Kerr effect is in switches where a switching pulse in the v polarization will rotate the phase of a signal pulse in the u polarization. In general, this use of the Kerr effect generates shadows—the signal pulse develops a component in the v polarization. However, when Manakov's equation applies, shadows no longer develop. If the signal pulse is a soliton, it undergoes a phase shift and some displacement, but no change in polarization and no distortion.

It is possible to experimentally study the phenomena described in this paper by using specially fabricated optical fibers.

Acknowledgment

The author gratefully acknowledges useful conversations with Dr. H. Haus, Dr. R. Stolen, and Dr. P.-K. Wai. He is also grateful to Dr. D. Kaup for discussions of Manakov's equation.

References

[1] C. R. Menyuk, "Nonlinear pulse propagation in birefringent optical fibers," *IEEE J. Quantum Electron.*, vol. QE-23, pp. 174–176, Feb. 1987.

[2] K. J. Blow, N. J. Doran, and D. Wood, "Generation and stabilization of short soliton pulses in the amplified nonlinear Schrödinger equation," *J. Opt. Soc. Amer. B*, vol. 5, pp. 381–390, Feb. 1988.

[3] A. Lattes, H. A. Haus, F. J. Leonberger, and E. P. Ippen, "An ultrafast all-optical gate," *IEEE J. Quantum Electron.*, vol. QE-19, pp. 1718–1723, Nov. 1983.

[4] M. J. LaGasse, D. L. Wong, J. C. Fujimoto, and H. A. Haus, "Ultrafast switching with a single-mode interferometer," *Opt. Lett.*, vol. 14, pp. 311–313, Mar. 1989.

[5] K. Kitayama, Y. Kimura, and S. Seikai, "Fiber-optic logic gate," *Appl. Phys. Lett.*, vol. 46, pp. 317–319, Feb. 1985.

[6] T. Morioka, M. Saruwatari, and A. Takada, "Ultrafast optical multi/demultiplexer utilizing optical Kerr effect in polarization-maintaining single-mode fibers," *Electron. Lett.*, vol. 23, pp. 453–454, Feb. 1987.

[7] N. J. Halas, D. Krökel, and D. Grischkowsky, "Ultrafast light-controlled optical-fiber modulator," *Appl. Phys. Lett.*, vol. 50, pp. 886–888, Apr. 1987.

[8] N. J. Doran and D. Wood, "Nonlinear-optical loop mirror," *Opt. Lett.*, vol. 13, pp. 56–58, Jan. 1988.

[9] S. Trillo, S. Wabnitz, E. M. Wright, and G. I. Stegeman, "Soliton switching in fiber nonlinear directional couplers," *Opt. Lett.*, vol. 13, pp. 672–674, Aug. 1988.

[10] C. R. Menyuk, "Stability of solitons in birefringent optical fibers. II. Arbitrary amplitudes," *J. Opt. Soc. Amer.*, vol. 5, pp. 392–402, Feb. 1988.

[11] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Sov. Phys.—JETP*, vol. 38, pp. 248–253, Feb. 1974 (*Zh. Eksp. Teor. Fiz.*, vol. 65, pp. 505–516, Aug. 1973).

[12] M. J. Ablowitz and H. Seger, *Solitons and the Inverse Scattering Transform.* Philadelphia, PA: SIAM, 1981.

[13] R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Appl. Opt.*, vol. 18, pp. 2241–2251, July 1979.

[14] D. N. Payne, A. J. Barlow, and J. J. Ramskov-Hansen, "Development of low and high birefringence fibers," *IEEE J. Quantum Electron.*, vol. QE-18, pp. 477–488, Mar. 1982.

[15] S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightwave Technol.*, vol. LT-1, pp. 312–331, June 1983.

[16] Y. Kodama, "Optical solitons in a monomode fiber," *J. Stat. Phys.*, vol. 39, pp. 597–614, June 1985.

[17] F. M. Mitschke and L. F. Mollenauer, "Discovery of the soliton self frequency shift," *Opt. Lett.*, vol. 11, pp. 659–661, Oct. 1986.

[18] P. D. Maker and R. W. Terhune, "Study of optical effects due to an induced polarization third order in the electric field strength," *Phys. Rev.*, vol. 137A, pp. 801–818, Feb. 1965.

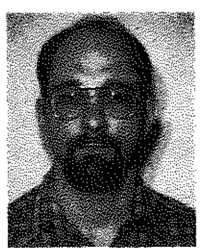
[19] K. J. Blow, N. J. Doran, and D. Wood, "Polarization instabilities for solitons in birefringent fibers," *Opt. Lett.*, vol. 12, pp. 202–204, Mar. 1987.

[20] D. N. Christodoulides and R. I. Joseph, "Vector solitons in birefringent nonlinear dispersive media," *Opt. Lett.*, vol. 13, pp. 53–55, Jan. 1988.

[21] A. Hasegawa and Y. Kodama, "Signal transmission by optical solitons in monomode fiber," *Proc. IEEE*, vol. 69, pp. 1145–1150, Sept. 1981.

[22] C. R. Menyuk, "Origin of solitons in the 'real' world," *Phys. Rev. A*, vol. 33, pp. 4367–4374, June 1986.

[23] L. F. Mollenauer and R. Stolen, "The soliton laser," *Opt. Lett.*, vol. 9, pp. 13–15, Jan. 1984.



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

He has worked as a Research Associate at the University of Maryland, College Park, and at Science Applications International Corporation, McLean, VA. Since 1986 he has been an Associate Professor with the Department of Electrical Engineering, University of Maryland, Baltimore County and College Park. He was the first regular faculty member at the Baltimore County Campus and was Program Coordinator from 1986 to 1988. For the last few years, his primary research area has been theoretical and computational studies of nonlinear optics.

Dr. Menyuk is a member of the American Physical Society, the Optical Society of America, and the Society for Industrial and Applied Mathematics.