Progress In Electromagnetics Research B, Vol. 19, 205–231, 2010

NONLINEAR PULSE PROPAGATION IN A WEAKLY BIREFRINGENT OPTICAL FIBER PART 1: DERIVATION OF COUPLED NONLINEAR SCHRODINGER EQUATIONS (CNLSE)

N. Kalyanasundaram

Jaypee Institute of Information Technology University
Noida, India

P. Muthuchidambaranathan

National Institute of Technology
Tiruchirappalli, India

Abstract—A systematic derivation of the Coupled Nonlinear Schrodinger Equations (CNLSE) governing nonlinear pulse propagation in a weakly birefringent monomode optical fiber based on a multiple-scale perturbation solution of the semilinear vector wave equation for the electric field in a (randomly) birefringent fiber medium is presented. The analysis of the nonlinear propagation characteristics of optical pulses based on a numerical solution of the CNLSE is deferred to the second part of this contribution.

1. INTRODUCTION

It is now generally accepted that nonlinear pulse propagation in a birefringent optical fiber is governed by a pair of Coupled Nonlinear Schrodinger-type Equations (CNLSE). The version of CNLSE that has found favor with many researchers in this area [1–3] was proposed by Menyuk [4] on the basis of a plane-wave approximation to the guided propagation in an optical fiber. In his 1999 paper [5] on multiple-length-scale methods for optical fiber transmission, Menyuk observed that "It is remarkable that no derivation of the nonlinear Schrodinger equation that is valid for physically realistic optical fibers exists within the scientific literature. First, all derivations that have been published in text books assume that optical fibers are perfectly

Corresponding author: N. Kalyanasundaram (n.kalyanasundaram@jiit.ac.in).

circular. Not only is this assumption false, but, in fact, just the opposite is true. The magnitude of an effect is inversely proportional to its corresponding scale length, and thus relative to the Kerr effect and chromatic dispersion, the birefringence must be considered large but rapidly varying. A conceptually correct derivation must take this essential fact into account. Second, even if a perfectly round fiber is assumed, most of the published derivations [6–10] — with notable exceptions such as Kodama's elegant derivation [11] — contain contradictory assumptions and errors".

Menyuk's 'derivation' [5] of the CNLSE is no doubt a substantial improvement over his earlier derivation [4] especially with regard to taking the fiber geometry into account. Nevertheless, even this improved version fails to completely meet the projected objective, viz., a conceptually correct derivation of the CNLSE from Maxwell's field equations. Menyuk approaches the problem by considering the effects of chromatic dispersion, fiber birefringence and nonlinearity one at a time (a separate evolution equation for the complex envelopes is deduced by assuming that only one of the effects is present), and then superposing the individual equations with the aid of certain ordering parameters (the precise definitions of which are not given in the paper) to arrive at the final evolution equation (CNLSE). Needless to say, such an approach completely ignores any possible interaction among the effects of chromatic dispersion, fiber birefringence and nonlinearity, and can hardly be termed as a systematic and consistent derivation of the CNLSE from Maxwell's field equations. A more serious drawback of his 'derivation' is the introduction of a birefringence parameter $\Delta\beta$ without either a definition or even an indication of how it is related to the basic parameters characterizing a fiber as birefringent. What is the interpretation of $\Delta\beta$ in the context of guided wave propagation through an optical fiber? Interpreting $\Delta\beta$ as $(\beta_1 - \beta_2)/2$ [12], where β_1 and β_2 are the propagation constants of the two orthogonally polarized modes, is nothing short of a self-contradiction since the orthogonally polarized degenerate modes have identical propagation constants in a fiber fabricated out of an isotropic and azimuthally homogeneous dielectric. Birefrigence exhibited by real fibers arises from the anisotropic and the non-axisymmetric random perturbations to which the isotropic and the azimuthally homogenous dielectric tensor is invariably subjected during the manufacturing and the installation stages. It is thus seen that a systematic derivation of the CNLSE governing nonlinear pulse propagation through a (randomly) birefringent fiber that is free from all inconsistences and self-contradictions is yet to emerge.

In this paper, the CNLSE governing the evolution of the wave packet-representation of the two orthogonal polarization modes of a weakly birefringent monomode fiber incorporating the cumulative effects of chromatic dispersion, birefringence and nonlinearity are derived starting from Maxwell's field equations using a multiple-scale perturbation approach.

2. A MODEL FOR BIREFRINGENCE

The vector wave equation for the electric field vector following from Maxwell's equations is [13]

$$-\nabla \times \nabla \times \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\varepsilon_0 c^2}\frac{\partial^2 \mathbf{P}}{\partial t^2} \tag{1}$$

where $\mathbf{P}$ is the dielectric polarization vector of the material medium, c is the speed of light in and ε_0 is the absolute permittivity of free space.

For a linearly anisotropic and centro-symmetric dielectric material, that is also transversely nonhomogeneous, we have the constitutive relation:

$$\frac{1}{\varepsilon_0}\left[\mathbf{P}_L(\mathbf{x},t) + \mathbf{P}_{NL}(\mathbf{x},t)\right] = \int_{-\infty}^{t} \boldsymbol{\chi}^{(1)}\left(\left|\mathbf{x}_\perp\right|, t-t_1\right)\, \mathbf{E}(\mathbf{x},t_1)\, dt_1$$
$$+\int_{-\infty}^{t}\int_{-\infty}^{t}\int_{-\infty}^{t} \chi^{(3)}(|\mathbf{x}_\perp|, t-t_1, t-t_2, t-t_3)(\mathbf{E}(t_1)\cdot\mathbf{E}(t_2))\mathbf{E}(t_3)dt_1dt_2dt_3, \tag{2}$$

where we have assumed that the first-order linear susceptibility tensor $\boldsymbol{\chi}^{(1)} \equiv \left[\chi_{ij}^{(1)}\right]$, $i,j = 1,2,3$, and the third order (nonlinear) scalar susceptibility $\chi^{(3)}$ depend on the position vector $\mathbf{x}$ only through the magnitude of its transverse part $\mathbf{x}_\perp$. The third-order susceptibility is further assumed to be axisymmetric; that is, its dependence on $\mathbf{x}_\perp$ is only through its magnitude $|\mathbf{x}_\perp|$. The axial asymmetry of the linear susceptibility tensor is one of the causes of birefringence exhibited by real fibers. The presence of the core-cladding boundary is the main cause of radial nonhomogeneity in a monomode fiber.

Computing the Fourier transform of (2) with respect to the time variable, and suppressing the spatial variables, we have

$$\begin{aligned}\frac{1}{\varepsilon_0}\hat{\mathbf{P}}(\omega) &= \frac{1}{\varepsilon_0}\left[\hat{\mathbf{P}}_L(\omega) + \hat{\mathbf{P}}_{NL}(\omega)\right]\\ &= \left[\hat{\boldsymbol{\chi}}^{(1)}(\omega)\right]\hat{\mathbf{E}}(\omega) + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\chi}^{(3)}(\omega_1,\omega_2,\omega_3)\left(\hat{\mathbf{E}}(\omega_1)\cdot\hat{\mathbf{E}}(\omega_2)\right)\\ &\quad \hat{\mathbf{E}}(\omega_3)\delta(\omega_1+\omega_2+\omega_3-\omega)\, d\omega_1 d\omega_2 d\omega_3 \end{aligned} \tag{3}$$

where

$$\hat{\mathbf{P}}(\omega) = \int_{-\infty}^{\infty} \mathbf{P}(t)\, e^{j\omega t} dt, \quad \hat{\mathbf{E}}(\omega) = \int_{-\infty}^{\infty} \mathbf{E}(t)\, e^{j\omega t} dt, \tag{4}$$

$$\hat{\boldsymbol{\chi}}^{(1)}(\omega) = \int_{-\infty}^{\infty} \boldsymbol{\chi}^{(1)}(t)\, e^{j\omega t} dt, \tag{5a}$$

and

$$\hat{\chi}^{(3)}(\omega_1, \omega_2, \omega_3) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi^{(3)}(t_1, t_2, t_3) e^{j(\omega_1 t_1 + \omega_2 t_2 + \omega_3 t_3)} dt_1 dt_2 dt_3 \tag{5b}$$

The effective range of integration in (5a) and (5b) is from 0 to ∞ only, because causality requires, $\chi_{ij}^{(1)}(t)$, $i, j = 1, 2, 3$, and $\chi^{(3)}(t_1, t_2, t_3)$ to be zero for negative values of the arguments. Suppressing the independent variables, the linear part of (3) may be written as

$$\hat{P}_{Li} = \varepsilon_0 \sum_{j=1}^{3} \hat{\chi}_{ij}^{(1)} \hat{E}_j, \quad i, j = 1, 2, 3 \tag{6}$$

For any dielectric material of general anisotropy, the susceptibility tensor (in the frequency domain) may be represented by a 3×3 symmetric matrix $\left[\hat{\chi}_{ij}^{(1)}\right]$ with complex entries [14]. The imaginary part of $\hat{\chi}_{ij}^{(1)}$ is smaller in magnitude than the real part by several orders of magnitude in a practical low-loss fiber material. The causality requirement implies that the real and imaginary parts of $\hat{\chi}_{ij}^{(1)}$ are related by Kramers-Kronig dispersion relation [14]. We may choose coordinates along the principal-axes of the crystal so that the off-diagonal entries of the matrix become zero. That is, in this principal-axes coordinate system, the linear susceptibility matrix becomes diagonal with complex diagonal entries $\hat{\chi}_{11}, \hat{\chi}_{22}, \hat{\chi}_{33}$.

In any fixed coordinate system (x, y, z) which is related to the principal-axes coordinate system (x', y', z) through a counter clockwise rotation through an angle θ in the $xy\,(x'y')$-plane about the z-axis (Fig. 1), the susceptibility tensor $\hat{\boldsymbol{\chi}}^{(1)}$ becomes

$$\begin{bmatrix} \hat{\chi}'_{11} & \hat{\chi}'_{12} & 0 \\ \hat{\chi}'_{21} & \hat{\chi}'_{22} & 0 \\ 0 & 0 & \hat{\chi}'_{33} \end{bmatrix}$$

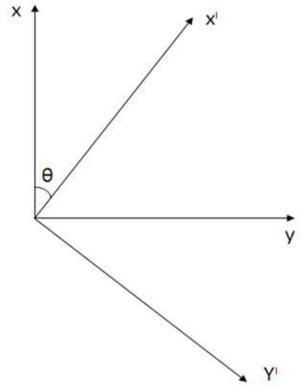

Figure 1. Rotational transformation of coordinates.

so that

$$\begin{aligned}\hat{P}_x &= \varepsilon_0\left[\hat{\chi}'_{11}\hat{E}_x + \hat{\chi}'_{12}\hat{E}_y\right]\\ \hat{P}_y &= \varepsilon_0\left[\hat{\chi}'_{21}\hat{E}_x + \hat{\chi}'_{22}\hat{E}_y\right]\\ \hat{P}_z &= \varepsilon_0\hat{\chi}'_{33}\hat{E}_z = \varepsilon_0\hat{\chi}_{33}\hat{E}_z\end{aligned} \tag{7a}$$

whereas

$$\hat{P}_{x'} = \varepsilon_0\hat{\chi}_{11}\hat{E}_{x'} \quad \hat{P}_{y'} = \varepsilon_0\hat{\chi}_{22}\hat{E}_{y'} \tag{7b}$$

Since

$$\begin{bmatrix}\hat{E}_{x'}\\ \hat{E}_{y'}\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\hat{E}_x\\ \hat{E}_y\end{bmatrix} \text{ and } \begin{bmatrix}\hat{P}_{x'}\\ \hat{P}_{y'}\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\hat{P}_x\\ \hat{P}_y\end{bmatrix},$$

we deduce from (7b) that

$$\begin{aligned}\cos\theta\hat{P}_x + \sin\theta\hat{P}_y &= \varepsilon_0\hat{\chi}_{11}(\cos\theta\hat{E}_x + \sin\theta\hat{E}_y)\\ -\sin\theta\hat{P}_x + \cos\theta\hat{P}_y &= \varepsilon_0\hat{\chi}_{22}(-\sin\theta\hat{E}_x + \cos\theta\hat{E}_y)\end{aligned} \tag{7c}$$

Solving Equation (7c) for $\hat{P}_x$ and $\hat{P}_y$, we have

$$\begin{aligned}\hat{P}_x &= \varepsilon_0\left(\hat{\chi}_{11}\cos^2\theta + \hat{\chi}_{22}\sin^2\theta\right)\hat{E}_x + \left((\hat{\chi}_{11} - \hat{\chi}_{22})\cos\theta\sin\theta\right)\hat{E}_y\\ \hat{P}_y &= \varepsilon_0\left((\hat{\chi}_{11} - \hat{\chi}_{22})\cos\theta\sin\theta\right)\hat{E}_x + \left(\hat{\chi}_{11}\sin^2\theta + \hat{\chi}_{22}\cos^2\theta\right)\hat{E}_y\end{aligned} \tag{7d}$$

Comparing (7d) with (7a) we get the relations

$$\hat{\chi}'_{11} = \hat{\chi}_{11}\cos^2\theta + \hat{\chi}_{22}\sin^2\theta = \left(\frac{\hat{\chi}_{11} + \hat{\chi}_{22}}{2}\right) + \left(\frac{\hat{\chi}_{11} - \hat{\chi}_{22}}{2}\right)\cos 2\theta$$

$$\begin{aligned}
\hat{\chi}'_{22} &= \hat{\chi}_{11}\sin^2\theta + \hat{\chi}_{22}\cos^2\theta = \left(\frac{\hat{\chi}_{11}+\hat{\chi}_{22}}{2}\right) - \left(\frac{\hat{\chi}_{11}-\hat{\chi}_{22}}{2}\right)\cos 2\theta \\
\hat{\chi}'_{21} &= \hat{\chi}'_{12} = \left(\frac{\hat{\chi}_{11}-\hat{\chi}_{22}}{2}\right)\sin 2\theta \\
\hat{\chi}'_{33} &= \hat{\chi}_{33}
\end{aligned} \tag{8}$$

Since the orientation of the principal axes in the transverse plane is not preserved in a (randomly) birefringent optical fiber, the polarization-state angle θ will be a slowly varying (random) function of the axial coordinate z. It is seen from (8) that the linear susceptibility tensor $\hat{\boldsymbol{\chi}}^{(1)}$ for a birefringent optical fiber admits the decomposition

$$\begin{aligned}
&\hat{\boldsymbol{\chi}}^{(1)}(\mathbf{x}_\perp,\omega) = \hat{\boldsymbol{\chi}}^{(1)}(|\mathbf{x}_\perp|,\omega) + \Delta\hat{\boldsymbol{\chi}}^{(1)}(\mathbf{x}_\perp,\omega) = \hat{\chi}(|\mathbf{x}_\perp|,\omega)\,\mathbf{I} \\
&+\begin{bmatrix}
\begin{array}{l}\Delta\hat{\chi}(\mathbf{x}_\perp,\omega)+(\hat{\chi}_d(|\mathbf{x}_\perp|,\omega)\\ +\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega))\cos 2\theta\end{array} & \begin{array}{l}(\hat{\chi}_d(|\mathbf{x}_\perp|,\omega)\\ +\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega))\sin 2\theta\end{array} & 0 \\
\begin{array}{l}(\hat{\chi}_d(|\mathbf{x}_\perp|,\omega)\\ +\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega))\sin 2\theta\end{array} & \begin{array}{l}\Delta\hat{\chi}(\mathbf{x}_\perp,\omega)-(\hat{\chi}_d(|\mathbf{x}_\perp|,\omega)\\ +\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega))\cos 2\theta\end{array} & 0 \\
0 & 0 & \begin{array}{l}\hat{\chi}_{33}(\mathbf{x}_\perp,\omega)\\ -\hat{\chi}(|\mathbf{x}_\perp|,\omega)\end{array}
\end{bmatrix}
\end{aligned} \tag{9}$$

where

$$\begin{aligned}
\hat{\chi}(|\mathbf{x}_\perp|,\omega) &\triangleq \langle\hat{\chi}_{11}+\hat{\chi}_{22}\rangle/2, \\
\hat{\chi}_d(|\mathbf{x}_\perp|,\omega) &\triangleq \langle\hat{\chi}_{11}-\hat{\chi}_{22}\rangle/2,
\end{aligned} \tag{10a}$$

$$\begin{aligned}
\Delta\hat{\chi}(\mathbf{x}_\perp,\omega) &\triangleq (\hat{\chi}_{11}+\hat{\chi}_{22})/2-\hat{\chi}(|\mathbf{x}_\perp|,\omega), \\
\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega) &\triangleq (\hat{\chi}_{11}-\hat{\chi}_{22})/2-\hat{\chi}_d(|\mathbf{x}_\perp|,\omega),
\end{aligned} \tag{10b}$$

and $\mathbf{I}$ is the 3×3 identity matrix. In (10a), $\langle f\rangle$ stands for the azimuthal average of a function f of the transverse coordinates. Denoting the polar coordinates of a point on the transverse plane by $(|\mathbf{x}_\perp|,\phi)$, the azimuthal average $\langle f\rangle$ of f is given by

$$\langle f\rangle\,(|\mathbf{x}_\perp|) \triangleq \frac{1}{2\pi}\int_0^{2\pi} f(\mathbf{x}_\perp)\,d\phi = \frac{1}{2\pi}\int_0^{2\pi} f(|\mathbf{x}_\perp|,\phi)\,d\phi$$

The azimuthal average will be a radial function of $r = |\mathbf{x}_\perp|$. In order for the asymmetric and anisotropic part $\Delta\hat{\boldsymbol{\chi}}^{(1)}$ of the linear susceptibility tensor to satisfy the trace condition [5]

$$\left\langle tr\left(\Delta\hat{\boldsymbol{\chi}}^{(1)}(\mathbf{x}_\perp,\omega)\right)\right\rangle = 0,$$

it is necessary that

$$\langle\hat{\chi}_{33}\rangle(|\mathbf{x}_\perp|,\omega) = \hat{\chi}(|\mathbf{x}_\perp|,\omega) \tag{11}$$

Since $\hat{\chi}_{33}$ may be assumed to be axisymmetric (in absence of any evidence to the contrary), (11) implies that

$$\hat{\chi}_{33}\left(\left|\mathbf{x}_{\perp}\right|,\omega\right) \equiv \hat{\chi}\left(\left|\mathbf{x}_{\perp}\right|,\omega\right)$$

thereby making the third diagonal entry of the symmetric matrix $\Delta\hat{\boldsymbol{\chi}}^{(1)}\left(\mathbf{x}_{\perp},\omega\right)$ identically equal to zero. The asymmetric and the anisotropic contributions to fiber birefringence may be considered as arising mainly from the $\Delta\hat{\chi}$ term and the $\hat{\chi}_d$ factor in the (non-zero) entries of the matrix $\Delta\hat{\boldsymbol{\chi}}^{(1)}\left(\mathbf{x}_{\perp},\omega\right)$ respectively whereas the factor $\Delta\hat{\chi}_d$ may be considered to be the combined source of a higher order contribution to birefringence from both asymmetry and anisotropy. In general, the azimuthal dependence of $\Delta\hat{\chi}\left(\mathbf{x}_{\perp},\omega\right)$ and $\Delta\hat{\chi}_d\left(\mathbf{x}_{\perp},\omega\right)$ exhibits a (random) slow variation with respect to the axial coordinate z. Since this random slow variation will be in a much larger length scale than that of θ with respect to z, we may assume $\Delta\hat{\chi}$ and $\Delta\hat{\chi}_d$ to be independent of z as far as the derivation of the CNLSE from Maxwell's equations is concerned.

3. PROBLEM FORMULATION

We are now adequately prepared to take up the systematic derivation of the CNLSE governing the evolution of the slowly varying envelopes of the two orthogonal polarization modes of the electric field vector using a multiple scale perturbation method. The non-dimensional perturbation parameter ε in terms of which a perturbation expansion of the electric field is sought is defined as

$$\varepsilon \underline{\Delta} \left(\frac{\omega_{\max}-\omega_{\min}}{2\omega_0}\right) \ll 1 \tag{12}$$

where $\omega_{\max}$ and $\omega_{\min}$ are the maximum and the minimum (radian) frequencies of the significant wave trains making up the initial pulse in the form of a wave packet, and ω_0 is the center frequency of the wave packet.

The working hypothesis underlying the perturbation approach adopted for the derivation of the CNLSE is that the nonlinear change of the dielectric constant due to the electric field of the optical pulse is on the order of ε^2 $(\mathrm{O}(\varepsilon^2))$ [15]. Therefore, the magnitude of the electric field associated with the optical pulse launched into the fiber has to be of $\mathrm{O}(\varepsilon)$ in order to be consistent with the working hypothesis.

Birefringence exhibited by optical fibers may be classified into weak, moderate or strong according to the relative magnitudes of $\hat{\chi}_d, \Delta\hat{\chi}$ and $\Delta\hat{\chi}_d$ as indicated in the following table. In the parameter regime in which present-day optical communication systems operate,

Table 1. Classification of fiber birefringence.

Degree of Birefringence	$\max_{\lvert\mathbf{x}_\perp\rvert}\left\lvert\frac{\hat{\chi}_d(\lvert\mathbf{x}_\perp\rvert,\omega_0)}{\hat{\chi}(\lvert\mathbf{x}_\perp\rvert,\omega_0)}\right\rvert$	$\max_{\mathbf{x}_\perp}\left\lvert\frac{\Delta\hat{\chi}(\mathbf{x}_\perp,\omega_0)}{\hat{\chi}(\lvert\mathbf{x}_\perp\rvert,\omega_0)}\right\rvert$	$\max_{\mathbf{x}_\perp}\left\lvert\frac{\Delta\hat{\chi}_d(\mathbf{x}_\perp,\omega_0)}{\hat{\chi}(\lvert\mathbf{x}_\perp\rvert,\omega_0)}\right\rvert$
Weak	$O(\varepsilon^2)$	$O(\varepsilon^2)$	$O(\varepsilon^3)$
Moderate	$O(\varepsilon)$	$O(\varepsilon^2)$	$O(\varepsilon^3)$
Strong	$O(\varepsilon)$	$O(\varepsilon)$	$O(\varepsilon^2)$

the contribution to fiber birefringence from $\hat{\chi}_d$ and $\Delta\hat{\chi}$ dominate over the much smaller contribution due to the axial asymmetry of the difference susceptibility $\langle\hat{\chi}_{11} - \hat{\chi}_{22}\rangle/2$ which has been reflected in the entries in the third column of the Table 1.

It has also been observed that the birefringence effect dominates over the nonlinear and the dispersive effects in practical fibers mainly because of the smaller length scale on which birefringence effects manifest themselves compared to those of dispersion and nonlinearity. Thus, practical fibers must be modeled as either moderately or strongly birefringent to be consistent with the above observation. Nevertheless, we shall model the fiber to be weakly birefringent for the purpose of the derivation of the CNLSE to be presented in this paper. This, in conjunction with the working hypothesis, means that the birefringence effects are assumed to be of the same order as the nonlinear effects. Moreover, the assumption of the weak birefringence rids the analysis of much mathematical clutter without obscuring the main steps in the perturbation approach adopted for the derivation of the CNLSE, and moreover, prepares the ground for the derivation of the CNLSE in the more realistic cases of moderate to strong birefringence to be taken up in a sequel to this contribution. In accordance with the assumption of weak birefringence, we may decompose the matrix $\Delta\hat{\boldsymbol{\chi}}^{(1)}$ (modeling the deviation of the susceptibility tensor $\hat{\boldsymbol{\chi}}^{(1)}$ from that of an (ideal) isotropic and homogeneous dielectric) as

$$\Delta\hat{\boldsymbol{\chi}}^{(1)} = \varepsilon^2\hat{\chi}_{dN}\mathbf{R}_\theta + \varepsilon^2\Delta\hat{\chi}_N\mathbf{I}_0 + \varepsilon^3\Delta\hat{\chi}_{dN}\mathbf{R}_\theta \tag{13}$$

where

$$\hat{\chi}_{dN} \,\underline{\Delta}\, \frac{\hat{\chi}_d}{\varepsilon^2}, \quad \Delta\hat{\chi}_N \,\underline{\Delta}\, \frac{\Delta\hat{\chi}}{\varepsilon^2}, \quad \Delta\hat{\chi}_{dN} \,\underline{\Delta}\, \frac{\Delta\hat{\chi}_d}{\varepsilon^3},$$

$$\mathbf{I}_0 \,\underline{\Delta} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{R}_\theta \,\underline{\Delta} \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ \sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{14}$$

The normalized susceptibilities $\hat{\chi}_{dN}, \Delta\hat{\chi}_N$ and $\Delta\hat{\chi}_{dN}$ will then be of the same order as $\hat{\chi}$. In practical low loss fibers, the imaginary part $\hat{\chi}_I$ of $\hat{\chi}$ will be several orders smaller than its real part $\hat{\chi}_R$. We make this fact explicit by assuming $|\hat{\chi}_I/\hat{\chi}_R|$ to be of at most $O(\varepsilon^2)$. Finally we assume that the polarization state angle θ does not vary too rapidly with the axial coordinate so that θ may be assumed to remain constant at least over every sufficiently short section of the fiber without any significant error.

Denoting the inverse Fourier transforms of $\hat{\chi}, \Delta\hat{\chi}_N, \hat{\chi}_{dN}$ and $\Delta\hat{\chi}_{dN}$ respectively by χ, $\Delta\chi_N, \chi_{dN}$ and $\Delta\chi_{dN}$, we may put the vector wave Equation (1) for the electric field vector in the form

$$\Delta\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\mathbf{E} + \int_0^\infty \chi\left(|\mathbf{x}_\perp|, t_1\right)\mathbf{E}(\mathbf{x}, t-t_1)dt_1\right)$$

$$= \frac{\varepsilon^2}{c^2}\frac{\partial^2}{\partial t^2}\int_0^\infty \left[\Delta\chi_N\left(\mathbf{x}_\perp, t_1\right)\mathbf{I}_0 + \chi_{dN}\left(|\mathbf{x}_\perp|, t_1\right)\mathbf{R}_\theta\right]\mathbf{E}(\mathbf{x}, t-t_1)dt_1$$
$$+\frac{\varepsilon^3}{c^2}\frac{\partial^2}{\partial t^2}\int_0^\infty \Delta\chi_{dN}\left(\mathbf{x}_\perp, t_1\right)\mathbf{R}_\theta\,\mathbf{E}(\mathbf{x}, t-t_1)dt_1$$
$$+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_0^\infty\int_0^\infty\int_0^\infty \chi^{(3)}(|\mathbf{x}_\perp|, t_1, t_2, t_3)(\mathbf{E}(\mathbf{x}, t-t_1)\cdot\mathbf{E}(\mathbf{x}, t-t_2))$$
$$\mathbf{E}(\mathbf{x}, t-t_3)dt_1dt_2dt_3 \tag{15}$$

The linear integrodifferential operator acting on the electric field vector $\mathbf{E}\left(\mathbf{x}_\perp, z, t\right)$ on the left of (15) may be represented in Cartesian coordinates (x, y, z) using matrix notation as

$$\tilde{\mathbf{L}}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial t}\right)$$
$$\underline{\Delta}\begin{pmatrix} \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\mathbf{1} + K\left(r\right)*\right) & -\frac{\partial^2}{\partial x\partial y} & -\frac{\partial^2}{\partial x\partial z} \\ -\frac{\partial^2}{\partial x\partial y} & \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\mathbf{1} + K(r)*\right) & -\frac{\partial^2}{\partial y\partial z} \\ -\frac{\partial^2}{\partial x\partial z} & -\frac{\partial^2}{\partial y\partial z} & \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\mathbf{1} + K(r)*\right) \end{pmatrix} \tag{16}$$

where $r = |\mathbf{x}_\perp| = \left(x^2 + y^2\right)^{1/2}$. In (16), $\mathbf{1}$ is the identity operator, and $K(r)*$ denotes the operation of convolution in the t-variable with the function

$$\begin{array}{rl} \tilde{\chi}\left(r, t\right)\underline{\Delta}\chi\left(r, t\right) & \text{for} \quad t \geq 0, \\ 0 & \text{for} \quad t < 0, \end{array}$$

that is

$$(K(r)*\mathbf{E})(\mathbf{x}_\perp,z,t)\ \underline{\Delta}\int_{-\infty}^{\infty}\tilde{\chi}(r,t_1)\mathbf{E}(\mathbf{x}_\perp,z,t-t_1)\,dt_1 = \int_{0}^{\infty}\chi(r,t_1)\mathbf{E}(\mathbf{x}_\perp,z,t-t_1)\,dt_1$$

The semilinear vector wave Equation (15) for the electric field vector may then be rewritten in operator notation as

$$\begin{aligned}&\tilde{\mathbf{L}}\left(\frac{\partial}{\partial \mathbf{x}},\frac{\partial}{\partial y},\frac{\partial}{\partial z},\frac{\partial}{\partial t}\right)\mathbf{E}(\mathbf{x}_\perp,z,t)\\ =\ &\frac{\varepsilon^2}{c^2}\frac{\partial^2}{\partial t^2}\int_{0}^{\infty}[\Delta\chi_N(\mathbf{x}_\perp,t_1)\mathbf{I}_0+\chi_{dN}(r,t_1)\mathbf{R}_\theta]\mathbf{E}(\mathbf{x}_\perp,z,t-t_1)\ dt_1\\ &+\frac{\varepsilon^3}{c^2}\frac{\partial^2}{\partial t^2}\int_{0}^{\infty}\Delta\chi_{dN}(\mathbf{x}_\perp,t_1)\mathbf{R}_\theta\mathbf{E}(\mathbf{x}_\perp,z,t-t_1)\,dt_1\\ &+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_{0}^{\infty}\int_{0}^{\infty}\int_{0}^{\infty}\chi^{(3)}(r,t_1,t_2,t_3)\left(\mathbf{E}(\mathbf{x}_\perp,z,t-t_1)\cdot\mathbf{E}(\mathbf{x}_\perp,z,t-t_2)\right)\\ &\mathbf{E}(\mathbf{x}_\perp,z,t-t_3)dt_1dt_2dt_3\end{aligned}\tag{17}$$

Equation (17) is the starting point in the derivation of the CNLSE using the perturbation method of multiple scales.

Consider the linear homogeneous problem for the electric field vector $\mathbf{E}$ resulting from setting the right side of (17) to zero:

$$\tilde{\mathbf{L}}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},\frac{\partial}{\partial t}\right)\mathbf{E}(\mathbf{x}_\perp,z,t)=0\tag{18}$$

Taking the Fourier transform of (18) in the time variable, we have

$$\hat{\mathbf{L}}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},-j\omega\right)\hat{\mathbf{E}}(\mathbf{x}_\perp,z,\omega)=-j\frac{\omega^2\varepsilon^2}{c^2}\hat{\chi}_{IN}(r,\omega)\hat{\mathbf{E}}(\mathbf{x}_\perp,z,\omega)\tag{19}$$

where

$$\hat{\mathbf{E}}(\mathbf{x}_\perp,z,\omega)\,\underline{\Delta}\int_{-\infty}^{\infty}\mathbf{E}(\mathbf{x}_\perp,z,t)e^{j\omega t}dt$$

is the Fourier transform of $\mathbf{E}(\mathbf{x}_\perp,z,t)$,

$$\hat{\chi}_{IN}(r,\omega)\,\underline{\Delta}\,\hat{\chi}_I(r,\omega)/\varepsilon^2$$

and $\hat{\mathbf{L}}$ is the matrix differential operator

$$\hat{\mathbf{L}}\underline{\Delta}\begin{pmatrix} \frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}+\frac{n^2(r,\omega)\omega^2}{c^2} & -\frac{\partial^2}{\partial \mathbf{x}\partial y} & -\frac{\partial^2}{\partial \mathbf{x}\partial z} \\ -\frac{\partial^2}{\partial \mathbf{x}\partial y} & \frac{\partial^2}{\partial \mathbf{x}^2}+\frac{\partial^2}{\partial z^2}+\frac{n^2(r,\omega)\omega^2}{c^2} & -\frac{\partial^2}{\partial y\partial z} \\ -\frac{\partial^2}{\partial \mathbf{x}\partial z} & -\frac{\partial^2}{\partial y\partial z} & \frac{\partial^2}{\partial \mathbf{x}^2}+\frac{\partial^2}{\partial y^2}+\frac{n^2(r,\omega)\omega^2}{c^2} \end{pmatrix} \tag{20}$$

In (20), $n(r,\omega)$ is the refractive index defined by

$$n^2(r,\omega)=1+\hat{\chi}_R(r,\omega)$$

The refractive index $n(r,\omega)$ is an even function of ω in view of the reality condition

$$\hat{\chi}(r,-\omega)=\hat{\chi}^*(r,\omega)$$

We first take up the linear homogeneous problem for the (real) symmetric differential operator $\hat{\mathbf{L}}$ by dropping the ε^2-order term on the right side of (19). The dropped ε^2-order term will reappear as one of the nonhomogeneous terms in the linear nonhomogeneous problem arising at the second stage of the perturbation expansion. We seek bounded solutions of the linear homogeneous problem

$$\hat{\mathbf{L}}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},-j\omega\right)\hat{\mathbf{E}}\left(\mathbf{x}_\perp,z,\omega\right)=0 \tag{21}$$

that decay to zero as $r=|\mathbf{x}_\perp|\to\infty$, in the form

$$\hat{\mathbf{E}}\left(\mathbf{x}_\perp,z,\omega\right)=e^{j\beta z}\mathbf{U}\left(\mathbf{x}_\perp,\beta,\omega\right) \tag{22}$$

where the (vector-valued) mode function $\mathbf{U}\left(\mathbf{x}_\perp,\beta,\omega\right)$ gives the transverse structure of the electric field $\hat{\mathbf{E}}\left(\mathbf{x}_\perp,z,\omega\right)$.

Denoting by $L^2\left(\mathbf{R}^2\right)$ the Hilbert space of square integrable complex-valued functions (of x and y) and by $S^1(0,a)$ the circle $\left\{(x,y):x^2+y^2=a^2\right\}$, we define the subspace $\mathbf{H}\left(\mathbf{R}^2\right)$ of $L^2\left(\mathbf{R}^2\right)^3$ to be the set of all those (vector-valued) functions which are infinitely differentiable on $\mathbf{R}^2/S^1(0,a)$ and which, together with all their derivatives, belong to $L^2\left(\mathbf{R}^2\right)^3$. The mode function $\mathbf{U}\in\mathbf{H}\left(\mathbf{R}^2\right)$ which is an inner-product space with the inner product induced from $L^2\left(\mathbf{R}^2\right)^3$. The inner product of $\mathbf{f},\mathbf{g}\in\mathbf{H}\left(\mathbf{R}^2\right)$ is given by

$$\langle\mathbf{f},\mathbf{g}\rangle\,\underline{\Delta}\iint\limits_{\mathbf{R}^2}\left(f_xg_x^*+f_yg_y^*+f_zg_z^*\right)dxdy$$

where the suffixes x,y and z denote the respective components. The cylinder $S^1(0,a)\times\mathbf{R}$ corresponds to the core-cladding boundary of the optical fiber.

The mode function $\mathbf{U}(\mathbf{x}_\perp, \beta, \omega)$ satisfies the linear homogeneous differential equations

$$\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, j\beta, -j\omega\right)\mathbf{U}$$
$$\underline{\Delta}\begin{pmatrix} \frac{\partial^2}{\partial y^2} - \beta^2 + \frac{n^2(r,\omega)\omega^2}{c^2} & -\frac{\partial^2}{\partial x \partial y} & -j\beta\frac{\partial}{\partial x} \\ -\frac{\partial^2}{\partial x \partial y} & \frac{\partial^2}{\partial x^2} - \beta^2 + \frac{n^2(r,\omega)\omega^2}{c^2} & -j\beta\frac{\partial}{\partial y} \\ -j\beta\frac{\partial}{\partial x} & -j\beta\frac{\partial}{\partial y} & \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{n^2(r,\omega)\omega^2}{c^2} \end{pmatrix}\mathbf{U}$$
$$= \mathbf{0} \tag{23}$$

Nontrivial solution for $\mathbf{U}$ exist only for those functions $\beta(\omega)$ that satisfy the dispersion equation. We now assume that the ratio of the refractive index difference between the core and the cladding to the core refractive index is sufficiently small so that the fiber is monomode supporting only a single, but possibly degenerate, guided mode, similar to the two-fold degenerate dominant HE_{11} mode supported by an ideal step-index fiber. The mode functions corresponding to the two degenerate modes may be chosen to be mutually orthogonal with respect to the L^2-inner product. The function $\mathbf{U}(\mathbf{x}_\perp, \beta(\omega), \omega)$ will then signify any one of the two orthogonal modes.

We seek a solution for the electric field vector $\mathbf{E}(\mathbf{x}_\perp, z, t)$ in the form of a wave packet, that is, as a continuous linear superposition of wavetrains with the frequencies of all the significant wavetrains lying in an ε-neighborhood of a center frequency ω_0:

$$\mathbf{E}(\mathbf{x}_\perp, z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{E}}(\mathbf{x}_\perp, z, \omega)\, e^{-j\omega t} d\omega \tag{24}$$

where almost all the 'energy' of the Fourier transform $\hat{\mathbf{E}}(\mathbf{x}_\perp, z, \omega)$ of the electric field vector is contained in an ε-neighborhood of ω_0. Thus

$$\hat{\mathbf{E}}(\mathbf{x}_\perp, z, \omega) = \tilde{\mathbf{U}}(\mathbf{x}_\perp, \omega)\hat{A}(z, \omega - \omega_0)e^{j\beta_0 z} + \tilde{\mathbf{U}}^*(\mathbf{x}_\perp, -\omega)\hat{A}^*(z, -\omega - \omega_0)e^{-j\beta_0 z} \tag{25}$$

where

$$\beta_0 \,\underline{\Delta}\, \beta(\omega_0), \quad \tilde{\mathbf{U}}(\mathbf{x}_\perp, \omega)\,\underline{\Delta}\mathbf{U}(\mathbf{x}_\perp, \beta(\omega), \omega),$$

the superscript * denotes the complex conjugate, and $\hat{A}(z, \omega)$ is essentially zero for $\omega \notin [-\omega_0\varepsilon, \omega_0\varepsilon]$. The z-dependence of the frequency-domain mode-envelope function $\hat{A}(z, \omega)$ arises mainly due to attenuation, dispersion and nonlinearity. In complete absence of attenuation, dispersion and nonlinearity, each wavetrain making up the wave packet will be traveling at a constant speed with constant amplitude. As a result, the wave packet will be translating itself along the z-axis at the constant speed $(\omega_0/\beta_0 = 1/\beta'(\omega_0))$ without any change of shape, and the mode-envelope function will be independent of z in a frame of reference moving with this constant speed along the z-axis. The effect of a small, but nonzero, amount of (linear) dispersion $\left(0 < \omega_0^2 \left|\beta''(\omega_0)\big/\beta(\omega_0)\right| \ll 1\right)$ would be a gradual spreading as seen in a frame of reference moving at the group speed $1/\beta'(\omega_0)$ in the axial direction. The effect of a small amount of attenuation would be gradual decay of the amplitudes of all the constituent wavetrains with propagation distance. Finally, the effect of the weak nonlinearity would be a gradual compression of the wavepacket in the axial direction partially compensating for the dispersive-spreading. When there is an exact balance between the opposing effects of dispersion and nonlinearity, an envelope soliton, capable of maintaining the pulse shape over large propagation distances in absence of attenuation, can in principle be formed. Thus, the time-domain mode envelope function $\tilde{A}(z,t)$, the Fourier transform of which is essentially supported in an ε-neighborhood of the origin on the ω-axis, will be a slowly varying function of both z and t.

Expanding the modal functions $\tilde{\mathbf{U}}(\mathbf{x}_\perp,\omega)$ and $\tilde{\mathbf{U}}^*(\mathbf{x}_\perp,-\omega)$ appearing in (25) in Taylor series around ω_0 and $-\omega_0$, and computing the inverse Fourier transform of the resulting infinite-series representation of $\hat{\mathbf{E}}\left(\mathbf{x}_\perp,z,\omega\right)$ term-by-term, we obtain a time-domain wave-packet representation for the electric field vector:

$$\mathbf{E}(\mathbf{x}_\perp,z,t) = e^{j(\beta_0 z-\omega_0 t)}\left(\sum_{n=0}^{\infty}\frac{1}{n!}\frac{\partial^n}{\partial\omega^n}\tilde{\mathbf{U}}\left(\mathbf{x}_\perp,\omega_0\right)\left(j\frac{\partial}{\partial t}\right)^n\right)\tilde{A}(z,t)$$
$$+e^{-j(\beta_0 z-\omega_0 t)}\left(\sum_{n=0}^{\infty}\frac{1}{n!}\frac{\partial^n}{\partial\omega^n}\tilde{\mathbf{U}}^*\left(\mathbf{x}_\perp,-\omega_0\right)\left(-j\frac{\partial}{\partial t}\right)^n\right)\tilde{A}^*(z,t) \quad (26)$$

where

$$\tilde{A}(z,t)\underline{\Delta}\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{A}(z,\omega)e^{-j\omega t}d\omega$$

will be a slowly-varying complex-valued function of z and t. Denoting the differential operators of infinite order (with variable coefficients) acting on the slowly-varying envelope functions $\tilde{A}(z,t)$ and $\tilde{A}^*(z,t)$ respectively by $\tilde{\mathbf{U}}\left(\mathbf{x}_\perp,\omega_0+j\frac{\partial}{\partial t}\right)$ and $\tilde{\mathbf{U}}^*\left(\mathbf{x}_\perp,-\omega_0-j\frac{\partial}{\partial t}\right)$, we end up with a compact time-domain representation of the wavepacket:

$$\begin{aligned}\mathbf{E}(\mathbf{x}_\perp, z, t) \;=\; & e^{j(\beta_0 z-\omega_0 t)}\tilde{\mathbf{U}}\left(\mathbf{x}_\perp, \omega_0 + j\frac{\partial}{\partial t}\right)\tilde{A}(z,t) \\ & + e^{-j(\beta_0 z-\omega_0 t)}\tilde{\mathbf{U}}^*\left(\mathbf{x}_\perp, -\omega_0 - j\frac{\partial}{\partial t}\right)\tilde{A}^*(z,t) \qquad (27)\end{aligned}$$

We now make explicit the slowly-varying nature of the complex envelope $\tilde{A}(z,t)$ as a function of z and t with the help of one slow time-scale variable and two slow space-scale variables defined by

$$T\underline{\Delta}\varepsilon t, \quad Z_1\underline{\Delta}\varepsilon z, \quad Z_2\underline{\Delta}\varepsilon^2 z \qquad (28)$$

Thus, we may represent $\tilde{A}(z,t)$ in terms of the slow-scale variables as

$$\tilde{A}(z,t)\,\underline{\Delta}\,A(Z_1, Z_2; T)\,.$$

The corresponding wave-packet representation of the electric field vector will then be

$$\begin{aligned} & \mathbf{E}\left(\mathbf{x}_\perp, z, t\right)\underline{\Delta}\tilde{\mathbf{E}}\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T; \varepsilon\right) \\ = \; & e^{j(\beta_0 z'-\omega_0 t')}\tilde{\mathbf{U}}\left(\mathbf{x}_\perp, \omega_0 + j\varepsilon\frac{\partial}{\partial T}\right)A(Z_1, Z_2, T) + c.c. \qquad (29)\end{aligned}$$

since

$$\frac{\partial}{\partial t}\tilde{A}(z,t) = \varepsilon\frac{\partial}{\partial T}A(Z_1, Z_2, T) \qquad (30a)$$

In (29), and in the sequel, ‘*c.c.*’ denotes the complex conjugate of the preceding expression. The corresponding transformation rules for the partial derivatives of $\mathbf{E}$ and $\tilde{\mathbf{E}}$ with respect to the axial coordinate and the time variable are

$$\frac{\partial\mathbf{E}}{\partial z} = \frac{\partial\tilde{\mathbf{E}}}{\partial z'} + \varepsilon\frac{\partial\tilde{\mathbf{E}}}{\partial Z_1} + \varepsilon^2\frac{\partial\tilde{\mathbf{E}}}{\partial Z_2}, \quad \frac{\partial\mathbf{E}}{\partial t} = \frac{\partial\tilde{\mathbf{E}}}{\partial t'} + \varepsilon\frac{\partial\tilde{\mathbf{E}}}{\partial T} \qquad (30b)$$

Accordingly

$$\begin{aligned} & \mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial t}\right)\mathbf{E}\left(\mathbf{x}_\perp, z, t\right) \\ = \; & \mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z'} + \varepsilon\frac{\partial}{\partial Z_1} + \varepsilon^2\frac{\partial}{\partial Z_2}, \frac{\partial}{\partial t'} + \varepsilon\frac{\partial}{\partial T}\right)\tilde{\mathbf{E}}\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T; \varepsilon\right) \\ = \; & e^{j(\beta_0 z'-\omega_0 t')}\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, j\beta_0 + \varepsilon\frac{\partial}{\partial Z_1} + \varepsilon^2\frac{\partial}{\partial Z_2}, -j\omega_0 + \varepsilon\frac{\partial}{\partial T}\right) \\ & \tilde{\mathbf{U}}\left(\mathbf{x}_\perp, \omega_0 + j\varepsilon\frac{\partial}{\partial T}\right)A(Z_1, Z_2, T)\end{aligned}$$

$$+e^{-j(\beta_0 z'-\omega_0 t')}\mathbf{L}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},-j\beta_0+\varepsilon\frac{\partial}{\partial Z_1}+\varepsilon^2\frac{\partial}{\partial Z_2},j\omega_0+\varepsilon\frac{\partial}{\partial T}\right)$$
$$\tilde{\mathbf{U}}^*\left(\mathbf{x}_\perp,-\omega_0-j\varepsilon\frac{\partial}{\partial T}\right)A^*(Z_1,Z_2,T), \tag{31}$$

In (31), $\mathbf{L}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},\frac{\partial}{\partial t}\right)$ is the integro-differential operator derived from $\tilde{\mathbf{L}}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},\frac{\partial}{\partial t}\right)$ by replacing $\chi\left(r,t\right)$ by the inverse Fourier transform $\chi_R\left(r,t\right)\left(\underline{\Delta}\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\hat{\chi}_R(r,\omega)e^{-j\omega t}d\omega\right)$ of the real part of $\hat{\chi}\left(r,\omega\right)$. We seek a solution of the semilinear wave Equation (17) in the form of a superposition of two wave packets centered around the same frequency ω_0. The peak values of the two constituent wavepackets are assumed to be $\mathrm{O}(\varepsilon)$ in accordance with the working hypothesis. Substituting $\tilde{\mathbf{E}}\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T;\varepsilon\right)\underline{\Delta}\,\varepsilon\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T;\varepsilon\right)$ for $\mathbf{E}(\mathbf{x}_\perp,z,t)$ in (17), we have the following semilinear vector wave equation for $\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T;\varepsilon\right)$ correct to ε^2-order:

$$\mathbf{L}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z'}+\varepsilon\frac{\partial}{\partial Z_1}+\varepsilon^2\frac{\partial}{\partial Z_2},\frac{\partial}{\partial t'}+\varepsilon\frac{\partial}{\partial T}\right)\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T;\varepsilon\right)$$
$$=j\frac{\varepsilon^2}{c^2}\frac{\partial^2}{\partial t'^2}\int\limits_{-\infty}^{\infty}\chi_{IN}(r,t_1)\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t'-t_1;Z_1,Z_2,T;\varepsilon\right)dt_1$$
$$+\frac{\varepsilon^2}{c^2}\frac{\partial^2}{\partial t'^2}\int\limits_0^{\infty}[\Delta\chi_N(\mathbf{x}_\perp,t_1)\mathbf{I}_0+\chi_{dN}(r,t_1)\mathbf{R}_\theta]\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t'-t_1;Z_1,Z_2,T;\varepsilon\right)dt_1$$
$$+\frac{\varepsilon^2}{c^2}\frac{\partial^2}{\partial t'^2}\int\limits_0^{\infty}\int\limits_0^{\infty}\int\limits_0^{\infty}\chi^{(3)}(r,t_1,t_2,t_3)\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t'-t_1;Z_1,Z_2,T;\varepsilon\right)$$
$$\cdot\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t'-t_2;Z_1,Z_2,T;\varepsilon\right)\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t'-t_3;Z_1,Z_2,T;\varepsilon\right)dt_1dt_2dt_3, \tag{32}$$

where $\chi_{IN}\left(r,t\right)$ is the inverse Fourier transform of $\hat{\chi}_{IN}\left(r,\omega\right)$.

4. PERTURBATION SOLUTION

We seek an asymptotic solution of (32) in the form of a power series in the perturbation parameter ε:

$$\bar{\mathbf{E}}\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T;\varepsilon\right)=\sum_{n=0}^{\infty}\varepsilon^n\mathbf{E}_n\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T\right) \tag{33}$$

Substituting (33) into (32) and setting $\varepsilon = 0$, we have the following homogeneous problem for $\mathbf{E}_0\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T\right)$ at the ε^0-order:

$$\mathbf{L}\left(\frac{\partial}{\partial \mathbf{x}}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z'}, \frac{\partial}{\partial t'}\right) \mathbf{E}_0\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T\right) = 0 \tag{34}$$

We assume a solution for $\mathbf{E}_0\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T\right)$ as a superposition of wavepackets corresponding to the two degenerate orthogonal modes of a monomode fiber:

$$\begin{aligned}&\mathbf{E}_0\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T\right)\\ &= e^{j(\beta_0 z' - \omega_0 t')} \sum_{l=1}^{2} \mathbf{U}_l\left(\mathbf{x}_\perp, \beta(\omega_0), \omega_0\right) A_l(Z_1, Z_2, T) + c.c. \end{aligned} \tag{35}$$

where $\mathbf{U}_l\left(\mathbf{x}_\perp, \beta(\omega_0), \omega_0\right)$, $l = 1, 2$, are the mode functions associated with the two orthogonally polarized modes evaluated at the common center frequency ω_0 of the wavepackets, and $A_l(Z_1, Z_2, T)$, $l = 1, 2$, are their slowly-varying complex envelopes. We assume without loss in generality that the mode functions $\mathbf{U}_l$, $l = 1, 2$, are normalized with respect to the norm induced by the L^2-innerproduct, that is

$$\|\mathbf{U}_l\|^2 \,\underline{\Delta}\, \langle \mathbf{U}_l, \mathbf{U}_l \rangle = 1 \quad \text{for } l = 1, 2.$$

The mode functions $\mathbf{U}_l\left(\mathbf{x}_\perp, \beta(\omega), \omega\right)$, $l = 1, 2$, satisfy the linear homogeneous differential equations in the transverse (Cartesian) coordinates (x, y):

$$\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, j\beta(\omega), -j\omega\right) \mathbf{U}_l\left(\mathbf{x}_\perp, \beta(\omega), \omega\right) = 0, \;\; l = 1, 2. \tag{36}$$

Since (36) is satisfied for all ω, we have the identity

$$\frac{d}{d\omega}\left(\mathbf{L}_0 \mathbf{U}_l\right)(\omega) \equiv 0, \;\; l = 1, 2 \text{ for all } \omega \tag{37}$$

where we have resorted to the compact notation $\mathbf{L}_0 = \mathbf{L}_0\left(j\beta, -j\omega\right)$ for the operator $\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, j\beta, -j\omega\right)$. Carrying out the ω-differentiation in (37), we get another identity:

$$\left(\mathbf{L}_0 \frac{\partial \tilde{\mathbf{U}}_l}{\partial \omega}\right)(\omega) = -\left(\frac{d\mathbf{L}_0}{d\omega} \tilde{\mathbf{U}}\right)(\omega) = -j\left[\left(\beta'(\omega)\,\mathbf{L}_1 - \mathbf{L}_2\right) \tilde{\mathbf{U}}_l\right](\omega), \; l = 1, 2, \tag{38a}$$

for all ω, where $\tilde{\mathbf{U}}_l = \tilde{\mathbf{U}}_l\left(x_\perp, \omega\right) \,\underline{\Delta}\, \mathbf{U}_l\left(x_\perp, \beta\left(\omega\right), \omega\right)$, $l = 1, 2$, and we have denoted $\frac{\partial \mathbf{L}_0}{\partial (j\beta)}$ by $\mathbf{L}_1$, $\frac{\partial \mathbf{L}_0}{\partial (-j\omega)}$ by $\mathbf{L}_2 = j\frac{\partial}{\partial \omega}\left(\frac{n^2(r,\omega)\omega^2}{c^2}\right)\mathbf{I}$, and $\frac{d\beta(\omega)}{d\omega}$ by $\beta'(\omega)$. Taking the L^2-innerproduct of (38a) with $\tilde{\mathbf{U}}_m\left(x_\perp,\omega\right) \underline{\Delta} \mathbf{U}_m\left(x_\perp,\beta\left(\omega\right),\omega\right),\ m=1,2$, we have

$$0=\left\langle \frac{\partial \tilde{\mathbf{U}}_l}{\partial \omega}, \mathbf{L}_0 \tilde{\mathbf{U}}_m \right\rangle(\omega)=\left\langle \mathbf{L}_0 \frac{\partial \tilde{\mathbf{U}}_l}{\partial \omega}, \tilde{\mathbf{U}}_m \right\rangle(\omega)$$
$$=-j\left\langle \left(\beta'\left(\omega\right)\mathbf{L}_1-\mathbf{L}_2\right)\tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_m \right\rangle(\omega),\ \ l,m=1,2, \tag{38b}$$

where we have appealed to the symmetry property of the operator $\mathbf{L}_0$ with respect to the L^2-innerproduct, viz.,

$$\langle \mathbf{L}_0\mathbf{f},\mathbf{g}\rangle=\langle \mathbf{f},\mathbf{L}_0\mathbf{g}\rangle \quad \forall \mathbf{f},\mathbf{g}\in \mathbf{H}(\mathbf{R}^2)$$

which follows easily by integration by parts since $\tilde{\mathbf{U}}_l$ and $\frac{\partial \tilde{\mathbf{U}}_l}{\partial \omega}\in \mathbf{H}\left(\mathbf{R}^2\right)$ for $l=1,2$, and the functions in $\mathbf{H}\left(\mathbf{R}^2\right)$ vanish at infinity together with all their derivatives with respect to x and y. From (38b), we obtain the useful identity:

$$\langle \mathbf{L}_2\mathbf{U}_l,\mathbf{U}_m\rangle(\omega)=\beta'\left(\omega\right)\langle \mathbf{L}_1\mathbf{U}_l,\mathbf{U}_m\rangle(\omega),\ \ l,m=1,2, \tag{39}$$

for all ω. Differentiating the identity (38a) once again with respect to ω, we obtain yet another identity:

$$\left(\mathbf{L}_0\frac{\partial^2\tilde{\mathbf{U}}_l}{\partial\omega^2}\right)(\omega)=-j\beta''(\omega)\left(\mathbf{L}_1\tilde{\mathbf{U}}_l\right)(\omega)+\left[\left(\beta'(\omega)^2\mathbf{L}_{11}-\beta'(\omega)\mathbf{L}_{12}+\mathbf{L}_{22}\right)\mathbf{U}_l\right](\omega)$$
$$-2j\left[\left(\beta'\left(\omega\right)\mathbf{L}_1-\mathbf{L}_2\right)\frac{\partial\tilde{\mathbf{U}}_l}{\partial\omega}\right](\omega),\ \ l=1,2 \tag{40}$$

In (40)

$$\mathbf{L}_{11}\underline{\Delta}\frac{\partial}{\partial(j\beta)}\mathbf{L}_1,\ \mathbf{L}_{12}\underline{\Delta}\frac{\partial}{\partial(-j\omega)}\mathbf{L}_1=-2\beta'\left(\omega\right)\mathbf{I}_0 \text{ and } \mathbf{L}_{22}\underline{\Delta}\frac{\partial}{\partial(-j\omega)}\mathbf{L}_2$$

The scalar identity (39) and the vector identities (38a) and (40) turn out to be quite useful for the derivation of the CNLSE.

We now turn to the ε^1-order problem. Substituting (33) into (32), differentiating the resulting equation with respect to ε, and setting $\varepsilon=0$, we have to ε^1-order:

$$\mathbf{L}\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z'},\frac{\partial}{\partial t'}\right)\mathbf{E}_1\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T\right)$$
$$=\left\{-j\mathbf{L}_0^{(1)}\left(\sum_{l=1}^{2}\frac{\partial\tilde{\mathbf{U}}_l\left(\mathbf{x}_\perp,\omega_0\right)}{\partial\omega}\frac{\partial A_l}{\partial T}\right)-\mathbf{L}_1^{(1)}\left(\sum_{l=1}^{2}\tilde{\mathbf{U}}_l\left(\mathbf{x}_\perp,\omega_0\right)\frac{\partial A_l}{\partial Z_1}\right)\right.$$
$$\left.-\mathbf{L}_2^{(1)}\left(\sum_{l=1}^{2}\tilde{\mathbf{U}}_l\left(\mathbf{x}_\perp,\omega_0\right)\frac{\partial A_l}{\partial T}\right)\right\}e^{j(\beta_0 z'-\omega_0 t')}+c.c. \tag{41}$$

where

$$\mathbf{L}_0^{(1)} \underline{\Delta} \mathbf{L}_0(j\beta_0, -j\omega_0),\ \mathbf{L}_1^{(1)} \underline{\Delta} \mathbf{L}_1(j\beta_0, -j\omega_0) \text{ and } \mathbf{L}_2^{(1)} \underline{\Delta} \mathbf{L}_2(j\beta_0, -j\omega_0)$$

Using the operator identities (38a), the expression for the coefficient of $e^{j(\beta_0 z' - \omega_0 t')}$ in (41) may be simplified to

$$\mathbf{F}_1(\mathbf{x}_\perp; Z_1, Z_2, T; \omega_0) = -\sum_{l=1}^{2} \mathbf{L}_1^{(1)} \tilde{\mathbf{U}}_l (\mathbf{x}_\perp, \omega_0) \left(\frac{\partial A_l}{\partial Z_1} + \beta'(\omega_0) \frac{\partial A_l}{\partial T} \right) \quad (42)$$

Assuming a solution of (41) in the form

$$\mathbf{E}_1 \big(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T \big) = \mathbf{V}_1(\mathbf{x}_\perp; Z_1, Z_2, T; \omega_0)\, e^{j(\beta_0 z' - \omega_0 t')} + c.c., \quad (43)$$

and substituting into (41), we find that $\mathbf{V}_1 (\mathbf{x}_\perp; Z_1, Z_2, T; \omega_0)$ satisfies the equation

$$\mathbf{L}_0^{(1)} \mathbf{V}_1(\mathbf{x}_\perp; Z_1, Z_2, T; \omega_0) = -\sum_{l=1}^{2} \mathbf{L}_1^{(1)} \tilde{\mathbf{U}}_l (\mathbf{x}_\perp, \omega_0) \left(\frac{\partial A_l}{\partial Z_1} + \beta'(\omega_0) \frac{\partial A_l}{\partial T} \right) \quad (44)$$

Taking the L^2-innerproduct of (44) with $\tilde{\mathbf{U}}_m (\mathbf{x}_\perp, \omega_0)$, $m = 1, 2$, we have

$$\begin{aligned} 0 &= \left\langle \mathbf{V}_1, \mathbf{L}_0^{(1)} \tilde{\mathbf{U}}_m \right\rangle = \left\langle \mathbf{L}_0^{(1)} \mathbf{V}_1, \tilde{\mathbf{U}}_m \right\rangle \\ &= -\sum_{l=1}^{2} \left\langle \mathbf{L}_1^{(1)} \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_m \right\rangle \left(\frac{\partial A_l}{\partial Z_1} + \beta'(\omega_0) \frac{\partial A_l}{\partial T_1} \right), \quad m = 1, 2 \end{aligned} \quad (45)$$

But, the identities (39) and the formula $\mathbf{L}_2^{(1)} = j \frac{\partial}{\partial \omega} \left(\frac{n^2(r, \omega_0) \omega_0^2}{c^2} \right) \mathbf{I}$ imply that

$$\left\langle \mathbf{L}_1^{(1)} \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_m \right\rangle = \beta'(\omega_0)^{-1} \left\langle \mathbf{L}_2^{(1)} \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_m \right\rangle \begin{aligned} &= 0 \quad \text{if} \quad l \neq m, \\ &\neq 0 \quad \text{if} \quad l = m. \end{aligned} \quad (46)$$

Thus, we deduce from (45) that

$$\left(\frac{\partial A_m}{\partial Z_1} + \beta'(\omega_0) \frac{\partial A_m}{\partial T} \right) = 0, \quad m = 1, 2 \quad (47)$$

Once the compatibility conditions (47) are satisfied, Equation (44) reduces to the homogeneous problem

$$\mathbf{L}_0^{(1)} \mathbf{V}_1 = 0 \quad (48)$$

Since a nontrivial solution in the null space of the $\mathbf{L}_0^{(1)}$ has already been included in the leading-order solutions $\mathbf{U}_l (\mathbf{x}_\perp, \beta(\omega_0), \omega_0)$, $l = 1, 2$, we may set $\mathbf{V}_1 (\mathbf{x}_\perp; Z_1, Z_2, T_2; \omega_0)$, and hence $\mathbf{E}_1 (\mathbf{x}_\perp, z', t'; Z_1, Z_2, T)$

identically equal to zero. The compatibility conditions (47) imply that the complex envelopes, A_l, $l = 1, 2$, depend on the slow-scale variables Z_1 and T only through the combination $T - \beta'(\omega_0)Z_1$.

We are now ready to take up the ε^2-order problem. Substituting (33), with $\mathbf{E}_1 \equiv 0$, into (32), differentiating the resulting equation twice with respect to ε, and setting $\varepsilon = 0$, we have, to order ε^2:

$$\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z'}, \frac{\partial}{\partial t'}\right)\mathbf{E}_2\left(\mathbf{x}_\perp, z', t'; Z_1, Z_2, T\right)$$
$$= -\frac{1}{2}e^{j(\beta_0 z' - \omega_0 t')}\frac{\partial^2}{\partial\varepsilon^2}\left\{\begin{bmatrix}\mathbf{L}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, j\beta_0 + \varepsilon\frac{\partial}{\partial Z_1} + \varepsilon^2\frac{\partial}{\partial Z_2}, -j\omega_0 + \varepsilon\frac{\partial}{\partial T}\right) \\ + j\frac{\varepsilon^2\omega_0^2}{c^2}\hat{\chi}_{IN}(r, \omega_0)\mathbf{I}\end{bmatrix}\right.$$
$$\left.\left(\sum_{l=1}^{2}\tilde{\mathbf{U}}_l\left(\mathbf{x}_\perp, \omega_0 + \varepsilon\frac{\partial}{\partial T}\right)A_l\left(Z_1, Z_2, T\right)\right)\right\}_{\varepsilon=0} + c.c.$$
$$+\frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\int_0^\infty[\Delta\chi_N(\mathbf{x}_\perp, t_1)\mathbf{I}_0 + \chi_{dN}(r, t_1)\mathbf{R}_\theta]\mathbf{E}_0\left(\mathbf{x}_\perp, z', t' - t_1; Z_1, Z_2, T\right)dt_1$$
$$+\frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\int_0^\infty\int_0^\infty\int_0^\infty\chi^{(3)}(r, t_1, t_2, t_3)\mathbf{E}_0\left(\mathbf{x}_\perp, z', t' - t_1; Z_1, Z_2, T\right)$$
$$\cdot\mathbf{E}_0\left(\mathbf{x}_\perp, z', t' - t_2; Z_1, Z_2, T\right)\mathbf{E}_0\left(\mathbf{x}_\perp, z', t' - t_3; Z_1, Z_2, T\right)dt_1 dt_2 dt_3$$
$$= \mathbf{F}_{21}\left(\mathbf{x}_\perp; Z_1, Z_2, T; \beta_0, \omega_0\right)e^{j(\beta_0 z' - \omega_0 t')} + c.c.$$
$$+\mathbf{F}_{23}\left(\mathbf{x}_\perp; Z_1, Z_2, T; 3\beta_0, 3\omega_0\right)e^{3j(\beta_0 z' - \omega_0 t')} + c.c. \tag{49}$$

The explicit expressions for the coefficients $\mathbf{F}_{21}$ and $\mathbf{F}_{23}$, respectively, of $e^{j(\beta_0 z' - \omega_0 t')}$ and $e^{3j(\beta_0 z' - \omega_0 t')}$ on the right side of (49) are

$$\mathbf{F}_{21}\left(\mathbf{x}_\perp; Z_1, Z_2, T; \beta_0, \omega_0\right)$$
$$= -\left(\mathbf{L}_1^{(1)}\frac{\partial}{\partial Z_2} + \frac{1}{2}\mathbf{L}_{11}^{(1)}\frac{\partial^2}{\partial Z_1^2} + \frac{1}{2}\mathbf{L}_{12}^{(1)}\frac{\partial^2}{\partial Z_1 \partial T} + \frac{1}{2}\mathbf{L}_{22}^{(1)}\frac{\partial^2}{\partial T^2}\right)\left(\sum_{l=1}^{2}\tilde{\mathbf{U}}_l A_l\right)$$
$$-j\left(\mathbf{L}_1^{(1)}\frac{\partial}{\partial Z_1} + \mathbf{L}_2^{(1)}\frac{\partial}{\partial T}\right)\left(\sum_{l=1}^{2}\frac{\partial\tilde{\mathbf{U}}_l}{\partial\omega_0}\frac{\partial A_l}{\partial T}\right)$$
$$-\frac{1}{2}\mathbf{L}_0^{(1)}\left(\sum_{l=1}^{2}\frac{\partial^2\tilde{\mathbf{U}}_l}{\partial\omega_0^2}\frac{\partial^2 A_l}{\partial T^2}\right) - j\frac{\omega_0^2}{c^2}\hat{\chi}_{IN}(r, \omega_0)\left(\sum_{l=1}^{2}\tilde{\mathbf{U}}_l A_l\right)$$
$$-\frac{\omega_0^2}{c^2}\left[\Delta\hat{\chi}_N(\mathbf{x}_\perp, \omega_0)\mathbf{I}_0 + \hat{\chi}_{dN}(r, \omega_0)\mathbf{R}_\theta\right]\sum_{l=1}^{2}\tilde{\mathbf{U}}_l A_l$$

$$-\frac{\omega_0^2}{c^2}\{\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\sum_{l=1}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_l\right)\,\tilde{\mathbf{U}}_l^*\left|A_l\right|^2A_l+2\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)$$

$$\sum_{l=1}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_l^*\right)\tilde{\mathbf{U}}_l\left|A_l\right|^2A_l$$

$$+2\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\sum_{\substack{l=1\\l\neq m}}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_m\right)\,\tilde{\mathbf{U}}_l^*\left|A_l\right|^2A_m+2\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)$$

$$\sum_{\substack{l=1\\l\neq m}}^{2}\left(\left|\tilde{\mathbf{U}}_l\right|^2\tilde{\mathbf{U}}_m+\left(\tilde{\mathbf{U}}_l^*\cdot\tilde{\mathbf{U}}_m\right)\tilde{\mathbf{U}}_l\right)\left|A_l\right|^2A_m$$

$$+\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\sum_{\substack{l=1\\l\neq m}}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_l\right)\,\tilde{\mathbf{U}}_m^*A_l^2A_m^*+2\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)$$

$$\sum_{\substack{l=1\\l\neq m}}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_m^*\right)\tilde{\mathbf{U}}_lA_l^2A_m^*\} \tag{50}$$

where

$$\mathbf{L}_{lm}^{(1)}\triangleq\mathbf{L}_{lm}(j\beta_0,-j\omega_0),\ \ l,m=1,2,\ \ l\leq m,$$
$$\frac{\partial}{\partial\omega_0}\triangleq\left[\frac{\partial}{\partial\omega}\right]_{\omega=\omega_0}\ \text{and}\ \ \frac{\partial^2}{\partial\omega_0^2}\triangleq\left[\frac{\partial^2}{\partial\omega^2}\right]_{\omega=\omega_0}$$

and we have made use of the equality $\hat{\chi}^{(3)}(\omega_0,-\omega_0,\omega_0)=\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)$ to condense the expressions for the coefficients of the nonlinear terms in (50).

$$\mathbf{F}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)=\frac{-9\omega_0^2}{c^2}\hat{\chi}^{(3)}(\omega_0,\omega_0,\omega_0)$$

$$\left\{\sum_{l=1}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_l\right)\tilde{\mathbf{U}}_lA_l^3+2\sum_{\substack{l=1\\l\neq m}}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_m\right)\tilde{\mathbf{U}}_lA_l^2A_m+\sum_{\substack{l=1\\l\neq m}}^{2}\left(\tilde{\mathbf{U}}_l\cdot\tilde{\mathbf{U}}_l\right)\tilde{\mathbf{U}}_mA_l^2A_m\right\} \tag{51}$$

In (50) and (51), we have suppressed the arguments $\mathbf{x}_\perp$ and ω_0 of $\tilde{\mathbf{U}}_l$, $l=1,2$, and the arguments Z_1, Z_2 and T of A_l, $l=1,2$, in the interests of brevity. Making use of the compatibility conditions (47), the sum of the first three terms in (50) may be rewritten as

$$-\left[\mathbf{L}^{(1)}\frac{\partial}{\partial Z_2}+\frac{1}{2}\left(\beta'(\omega_0)^2\mathbf{L}_{11}^{(1)}-\beta'(\omega_0)\mathbf{L}_{12}^{(1)}+\mathbf{L}_{22}^{(1)}\right)\frac{\partial^2}{\partial T^2}\right]\sum_{l=1}^{2}\tilde{\mathbf{U}}_lA_l$$

$$+j\left(\beta'(\omega_0)\mathbf{L}_1^{(1)}-\mathbf{L}_2^{(1)}\right)\sum_{l=1}^{2}\left(\frac{\partial\tilde{\mathbf{U}}_l}{\partial\omega_0}\right)\frac{\partial^2 A_l}{\partial T^2}+\frac{1}{2}\mathbf{L}_0^{(1)}\sum_{l=1}^{2}\left(\frac{\partial^2\tilde{\mathbf{U}}_l}{\partial\omega_0^2}\right)\frac{\partial^2 A_l}{\partial T^2}$$

which collapse, in view of the identity (40), to

$$-\mathbf{L}_1^{(1)}\sum_{l=1}^{2}\tilde{\mathbf{U}}_l\left(\frac{\partial A_l}{\partial Z_2}+j\frac{\beta''(\omega_0)}{2}\frac{\partial^2 A_l}{\partial T^2}\right) \tag{52}$$

We look for a solution of the nonhomogeneous Equation (49) having the same form as the nonhomogeneous terms:

$$\mathbf{E}_2\left(\mathbf{x}_\perp,z',t';Z_1,Z_2,T\right)=\mathbf{V}_{21}(\mathbf{x}_\perp;Z_1,Z_2,T;\beta_0,\omega_0)e^{j(\beta_0 z'-\omega_0 t')}+c.c.$$
$$+\mathbf{V}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)\,e^{3j(\beta_0 z'-\omega_0 t')}+c.c. \tag{53}$$

The (vector) functions $\mathbf{V}_{21}$ and $\mathbf{V}_{23}$ are to be chosen to satisfy, respectively, the nonhomogeneous differential equations

$$\mathbf{L}_0^{(1)}\mathbf{V}_{21}\left(\mathbf{x}_\perp;Z_1,Z_2,T;\beta_0,\omega_0\right)=\mathbf{F}_{21}\left(\mathbf{x}_\perp;Z_1,Z_2,T;\beta_0,\omega_0\right) \tag{54}$$

and

$$\mathbf{L}_0^{(3)}\mathbf{V}_{23}\left(x_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)=\mathbf{F}_{23}\left(x_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right) \tag{55}$$

where

$$\mathbf{L}_0^{(3)}\underline{\Delta}\mathbf{L}_0(3j\beta_0,-3j\omega_0)$$

Let us consider (55) first. Since $3\beta(\omega_0)\neq\beta(3\omega_0)$, in general, the null-space of the operator $\mathbf{L}_0^{(3)}$ is trivial. Consequently, the range of the operator is the whole space $\mathbf{H}\left(\mathbf{R}^2\right)$. Therefore, (55) can be solved uniquely for $\mathbf{V}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)$ in terms of $\mathbf{F}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)$. A homogeneous solution of the form $\sum\limits_{l=1}^{2}\mathbf{U}_l\left(\mathbf{x}_\perp,\beta(3\omega_0),3\omega_0\right)A_l^{(3)}\left(Z_1,Z_2,T\right)$ belonging to the nontrivial null space of the operator $\mathbf{L}_0\left(j\beta(3\omega_0),-3j\omega_0\right)$ may be added to $\mathbf{V}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)$ and the values of $A_l^{(3)}\left(0,0,T\right),\,l=1,2$, chosen to make the resulting total solution corresponding to the wavepacket centered around $3\omega_0$ generated out of the nonlinear interaction vanish at $z=Z_1=Z_2=0$ (launching plane). The semilinear partial differential equations in the slow variables Z_1, Z_2 and T governing the evolution of $A_l^{(3)}\left(Z_1,Z_2,T\right),\,l=1,2$, and the ε^4-order corrections to the complex envelopes $A_l\left(Z_1,Z_2,T\right),\,l=1,2$, arising out of the nonlinear interaction between $\mathbf{V}_{23}\left(\mathbf{x}_\perp;Z_1,Z_2,T;3\beta_0,3\omega_0\right)e^{3j(\beta_0 z'-\omega_0 t')}+c.c.$ and $e^{j(\beta_0 z'-\omega_0 t')}\sum\limits_{l=1}^{2}\mathbf{U}_l\left(\mathbf{x}_\perp,\beta_0,\omega_0\right)A_l\left(Z_1,Z_2,T\right)+c.c.$ can be determined only at the 4th stage of the perturbation. In general, the correction to the complex envelopes $A_l(Z_1, Z_2, T)$, $l = 1, 2$, due to all the odd harmonics of ω_0 (generated out of the nonlinear interaction) up to and including the $(2n+1)$th harmonic will be of the order ε^{2n+2} at least. This happy circumstance makes it possible to study nonlinear pulse propagation in an optical fiber independently of any harmonic or combination frequency generation.

We now return to (54). Taking the L^2-inner product of (54) with $\tilde{\mathbf{U}}_p(x_\perp, \omega_0)$, $p = 1, 2$, we have

$$0 = \left\langle \mathbf{V}_{21}, \mathbf{L}_0^{(1)} \tilde{\mathbf{U}}_p \right\rangle = \left\langle \mathbf{L}_0^{(1)} \mathbf{V}_{21}, \tilde{\mathbf{U}}_p \right\rangle = \left\langle \mathbf{F}_{21}, \tilde{\mathbf{U}}_p \right\rangle, \quad p = 1, 2. \tag{56}$$

Now

$$\begin{aligned}
&\left\langle \mathbf{F}_{21}, \tilde{\mathbf{U}}_p \right\rangle = -\sum_{l=1}^{2} \left\{ \left\langle \mathbf{L}_1^{(1)} \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_p \right\rangle \left(\frac{\partial A_l}{\partial Z_2} + j \frac{\beta''(\omega_0)}{2} \frac{\partial^2 A_l}{\partial T^2} \right) \right. \\
&\left. + \frac{\omega_0^2}{c^2} \left\langle \left(\Delta \hat{\chi}_N \mathbf{I}_0 + \hat{\chi}_{dN} \mathbf{R}_\theta \right) \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_p \right\rangle A_l \right\} \\
&- j \frac{\omega_0^2}{c^2} \left\langle \hat{\chi}_{IN}(\omega_0) \tilde{\mathbf{U}}_p, \tilde{\mathbf{U}}_p \right\rangle A_p - \frac{\omega_0^2}{c^2} \left\{ \left\langle \hat{\chi}^{(3)}(\omega_0, \omega_0, -\omega_0) \left(\tilde{\mathbf{U}}_p \cdot \tilde{\mathbf{U}}_p \right) \tilde{\mathbf{U}}_p^*, \tilde{\mathbf{U}}_p \right\rangle \right. \\
&+ 2 \left\langle \hat{\chi}^{(3)}(-\omega_0, \omega_0, \omega_0) \left(\tilde{\mathbf{U}}_p \cdot \tilde{\mathbf{U}}_p^* \right) \tilde{\mathbf{U}}_p, \tilde{\mathbf{U}}_p \right\rangle |A_p|^2 A_p \\
&+ 2 \sum_{l=1}^{2} (1 - \delta_{lp}) \left[\left\langle \hat{\chi}^{(3)}(\omega_0, \omega_0, -\omega_0) \left(\tilde{\mathbf{U}}_l \cdot \tilde{\mathbf{U}}_p \right) \tilde{\mathbf{U}}_l^* \right. \right. \\
&\left. \left. + \hat{\chi}^{(3)}(-\omega_0, \omega_0, \omega_0) \left(\left| \tilde{\mathbf{U}}_l \right|^2 \tilde{\mathbf{U}}_p + \left(\tilde{\mathbf{U}}_l^* . \tilde{\mathbf{U}}_p \right) \tilde{\mathbf{U}}_l \right), \tilde{\mathbf{U}}_p \right\rangle \right] |A_l|^2 A_p \\
&+ \sum_{l=1}^{2} (1 - \delta_{lp}) \left[\left\langle \hat{\chi}^{(3)}(\omega_0, \omega_0, -\omega_0) \left(\tilde{\mathbf{U}}_l \cdot \tilde{\mathbf{U}}_l \right) \tilde{\mathbf{U}}_p^* \right. \right. \\
&\left. \left. \left. + 2 \hat{\chi}^{(3)}(-\omega_0, \omega_0, \omega_0) \left(\tilde{\mathbf{U}}_l \cdot \tilde{\mathbf{U}}_p^* \right) \tilde{\mathbf{U}}_l, \tilde{\mathbf{U}}_p \right\rangle \right] A_l^2 A_p^* \right\}, \quad p = 1, 2
\end{aligned} \tag{57}$$

With the help of the identity (39), the compatibility conditions $\left\langle \mathbf{F}_{21}, \tilde{\mathbf{U}}_p \right\rangle = 0, p = 1, 2$, may be translated to

$$\begin{aligned}
&\frac{\partial A_m}{\partial Z_2} + j \frac{\beta''(\omega_0)}{2} \frac{\partial^2 A_m}{\partial T^2} + \alpha(\omega_0) A_m - j \sum_{l=1}^{2} \Lambda_{ml}(\theta, \omega_0) A_l \\
&- j \left(\sum_{l=1}^{2} \left[n_1(\omega_0) \delta_{ml} + n_3(\omega_0)(1 - \delta_{ml}) \right] \; |A_l|^2 \right) A_m
\end{aligned}$$

$$-jn_2\left(\omega_0\right)\left(\sum_{l=1}^{2}\left(1-\delta_{ml}\right)A_l^2\right)A_m^*,\ \ m=1,2, \tag{58}$$

where

$$\alpha\left(\omega_0\right)=\beta'\left(\omega_0\right)\left(\omega_0^2/c^2\right)\left\langle\hat{\chi}_{IN}\left(\omega_0\right)\tilde{\mathbf{U}}_m,\tilde{\mathbf{U}}_m\right\rangle\Big/D\left(\omega_0\right), \tag{59a}$$

$$\begin{aligned}-\Lambda_{11}\left(\theta,\omega_0\right)=\Lambda_{22}\left(\theta,\omega_0\right)=\sigma_{s1}\left(\omega_0\right)-\sigma_a\left(\omega_0\right)\cos 2\theta,\\ \Lambda_{12}\left(\theta,\omega_0\right)=\Lambda_{21}\left(\theta,\omega_0\right)=\sigma_{s2}\left(\omega_0\right)-\sigma_a\left(\omega_0\right)\sin 2\theta,\end{aligned} \tag{59b}$$

$$\begin{aligned}&n_1\left(\omega_0\right)=\beta'\left(\omega_0\right)\left(\omega_0^2/c^2\right)\left\langle\left(\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\left(\tilde{\mathbf{U}}_m\cdot\tilde{\mathbf{U}}_m\right)\tilde{\mathbf{U}}_m^*\right.\right.\\ &\left.\left.+2\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)\right)\left|\tilde{\mathbf{U}}_m\right|^2\tilde{\mathbf{U}}_m,\tilde{\mathbf{U}}_m\right\rangle\Big/D(\omega_0),\\ &n_2\left(\omega_0\right)=\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\left(1-\delta_{mp}\right)\left\langle\left(\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\left(\tilde{\mathbf{U}}_m\cdot\tilde{\mathbf{U}}_m\right)\tilde{\mathbf{U}}_p^*\right.\right.\\ &\left.\left.+2\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)\right)\ \left(\tilde{\mathbf{U}}_p^*\cdot\tilde{\mathbf{U}}_m\right)\tilde{\mathbf{U}}_m,\tilde{\mathbf{U}}_p\right\rangle\Big/D(\omega_0),\\ &n_3\left(\omega_0\right)=n_1\left(\omega_0\right)-n_2\left(\omega_0\right)=2\beta'\left(\omega_0\right)\left(\omega_0^2/c^2\right)\left(1-\delta_{mp}\right)\\ &\left\langle\begin{array}{l}\hat{\chi}^{(3)}(\omega_0,\omega_0,-\omega_0)\left(\tilde{\mathbf{U}}_m.\tilde{\mathbf{U}}_p\right)\tilde{\mathbf{U}}_m^*\\ +\hat{\chi}^{(3)}(-\omega_0,\omega_0,\omega_0)\left(\left|\tilde{\mathbf{U}}_m\right|^2\tilde{\mathbf{U}}_p+\left(\tilde{\mathbf{U}}_m^*.\tilde{\mathbf{U}}_p\right)\tilde{\mathbf{U}}_m\right),\tilde{\mathbf{U}}_p\end{array}\right\rangle\Big/D(\omega_0)\end{aligned} \tag{59c}$$

and where

$$\begin{aligned}\sigma_{s1}\left(\omega_0\right)&=-\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\Delta\hat{\chi}_N\left(\omega_0\right)\tilde{\mathbf{U}}_1,\mathbf{I}_0\tilde{\mathbf{U}}_1\right\rangle\Big/D\left(\omega_0\right)\\ &=\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\Delta\hat{\chi}_N\left(\omega_0\right)\tilde{\mathbf{U}}_2,\mathbf{I}_0\tilde{\mathbf{U}}_2\right\rangle\Big/D\left(\omega_0\right),\\ \sigma_{s2}\left(\omega_0\right)&=\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\Delta\hat{\chi}_N\left(\omega_0\right)\tilde{\mathbf{U}}_1,\mathbf{I}_0\tilde{\mathbf{U}}_2\right\rangle\Big/D\left(\omega_0\right)\\ &=\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\Delta\hat{\chi}_N\left(\omega_0\right)\tilde{\mathbf{U}}_2,\mathbf{I}_0\tilde{\mathbf{U}}_1\right\rangle\Big/D\left(\omega_0\right),\end{aligned} \tag{60a}$$

$$\begin{aligned}\sigma_a\left(\omega_0\right)&=\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\hat{\chi}_{dN}\left(\omega_0\right)\tilde{\mathbf{U}}_1,\mathbf{I}_1\tilde{\mathbf{U}}_1\right\rangle\Big/D\left(\omega_0\right)\\ &=-\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\hat{\chi}_{dN}\left(\omega_0\right)\tilde{\mathbf{U}}_2,\mathbf{I}_1\tilde{\mathbf{U}}_2\right\rangle\Big/D\left(\omega_0\right)\\ &=-\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\hat{\chi}_{dN}\left(\omega_0\right)\tilde{\mathbf{U}}_1,\mathbf{I}_2\tilde{\mathbf{U}}_2\right\rangle\Big/D\left(\omega_0\right)\\ &=-\beta'\left(\omega_0\right)\ \left(\omega_0^2/c^2\right)\ \left\langle\hat{\chi}_{dN}\left(\omega_0\right)\tilde{\mathbf{U}}_2,\mathbf{I}_2\tilde{\mathbf{U}}_1\right\rangle\Big/D\left(\omega_0\right),\end{aligned} \tag{60b}$$

$$D(\omega_0) = -j\left\langle \mathbf{L}_2^{(1)}\mathbf{U}_m, \tilde{\mathbf{U}}_m \right\rangle = \left\langle \frac{\partial}{\partial\omega}\left(\frac{n^2(r,\omega_0)\omega_0^2}{c^2}\right)\tilde{\mathbf{U}}_m, \tilde{\mathbf{U}}_m \right\rangle, \quad (61)$$

$$\mathbf{I}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{I}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (62)$$

Since the susceptibilities $\Delta\hat{\chi}, \hat{\chi}_{dN}$ and $\hat{\chi}^{(3)}$ are all complex with their imaginary parts several orders smaller than their respective real parts, the coefficients of asymmetry $\sigma_{s1}(\omega_0)$ and $\sigma_{s2}(\omega_0)$, the coefficient of anisotropy $\sigma_a(\omega_0)$, and the coefficients of nonlinearity $n_1(\omega_0)$ and $n_2(\omega_0)$ turn out also to be complex with very small imaginary parts.

Since the dependence of the slowly-varying envelope functions $A_m(Z_1, Z_2, T)$, $m = 1, 2$, on Z_1 and T is only through the combination $T - \beta'(\omega_0) Z_1$, we may introduce a new set $\{\zeta, \tau\}$ of independent variables related to the set $\{Z_1, Z_2, T\}$ by

$$\zeta \underline{\Delta} Z_2 = \varepsilon Z_1 = \varepsilon^2 z \quad \tau \underline{\Delta} T - \beta'(\omega_0) Z_1 = \varepsilon\left(t - \beta'(\omega_0) z\right) \quad (63a)$$

In terms of the new dependent variables

$$A(\zeta,\tau)\underline{\Delta} A_1(Z_1, Z_2, T), \quad B(\zeta,\tau)\underline{\Delta} A_2(Z_1, Z_2, T), \quad (63b)$$

Equation (58) becomes

$$\frac{\partial A}{\partial\zeta} + j\frac{\beta''(\omega_0)}{2}\frac{\partial^2 A}{\partial\tau^2} + \alpha(\omega_0)A + j\left(b_1(\theta,\omega_0)A - b_2(\theta,\omega_0)B\right)$$
$$-jn_1(\omega_0)A\left(|A|^2 + |B|^2\right) + jn_2(\omega_0)\left(A|B|^2 - A^*B^2\right) = 0 \quad (64a)$$
$$\frac{\partial B}{\partial\zeta} + j\frac{\beta''(\omega_0)}{2}\frac{\partial^2 B}{\partial\tau^2} + \alpha(\omega_0)B - j\left(b_1(\theta,\omega_0)B + b_2(\theta,\omega_0)A\right)$$
$$-jn_1(\omega_0)B\left(|A|^2 + |B|^2\right) + jn_2(\omega_0)\left(|A|^2 B - A^2B^*\right) = 0 \quad (64b)$$

where

$$\begin{aligned} b_1(\theta,\omega_0) &= \sigma_{s1}(\omega_0) - \sigma_a(\omega_0)\cos 2\theta, \\ b_2(\theta,\omega_0) &= \sigma_{s2}(\omega_0) - \sigma_a(\omega_0)\sin 2\theta \end{aligned} \quad (65)$$

The pair of Equations (64a)–(64b) is the CNLSE governing the evolution of the complex envelopes $A(\zeta,\tau)$ and $B(\zeta,\tau)$ of the wavepackets representing the two orthogonal polarization modes of a weakly birefringent monomode fiber. In (64), $\alpha(\omega_0)$ is the usual coefficient of attenuation, $b_1(\theta,\omega_0)$ and $b_2(\theta,\omega_0)$ are the two birefringence coefficients (associated with terms linear in A and B), and $n_1(\omega_0)$ and $n_2(\omega_0)$ are the coefficients of the nonlinear terms. Since the imaginary parts of $\sigma_{s1}(\omega_0), \sigma_{s2}(\omega_0), \sigma_a(\omega_0), n_1(\omega_0)$ and $n_2(\omega_0)$ are much smaller in magnitude compared to their respective real parts, the complex-valued birefringent coefficients $b_1(\theta, \omega_0)$ and $b_2(\theta, \omega_0)$ and the coefficients of nonlinearity $n_1(\omega_0)$ and $n_2(\omega_0)$ may be replaced by their respective real parts in (64) without any significant error. All the coefficients appearing in the CNLSE (64), henceforth, are understood to be real. The birefringence coefficient $b_2(\theta, \omega_0)$ gives rise to linear coupling between the mode envelope functions $A(\zeta, \tau)$ and $B(\zeta, \tau)$. Since the last term in (64a) and that in (64b) are equal to $2n_2(\omega_0)\,\mathrm{Im}\,(A^*B)\,B$ and $-2n_2(\omega_0)\,\mathrm{Im}\,(A^*B)\,A$ respectively, these terms represent an exchange of energy, that varies with ζ and τ, between the mode envelope functions arising out of the nonlinear coupling between the two modes. The two coupled partial differential Equations (64a)–(64b) may be combined into a single vector partial differential equation for the two-dimensional vector envelope function $\mathbf{A}(\zeta, \tau) \underline{\Delta} [A(\zeta, \tau), B(\zeta, \tau)]^T$:

$$\frac{\partial \mathbf{A}}{\partial \zeta} + j\frac{\beta''(\omega_0)}{2}\frac{\partial^2 \mathbf{A}}{\partial \tau^2} + \alpha(\omega_0)\mathbf{A} - j\mathbf{B}(\theta, \omega_0)\,\mathbf{A} - jn_1(\omega_0)\,|\mathbf{A}|^2\,\mathbf{A}$$
$$+jn_2(\omega_0)\left(\mathbf{A}^H\boldsymbol{\sigma}_3\mathbf{A}\right)\boldsymbol{\sigma}_3\mathbf{A} = \mathbf{0} \qquad (66)$$

where the 2×2 matrices $\mathbf{B}(\theta, \omega_0)$ and $\boldsymbol{\sigma}_3$ are given by

$$\mathbf{B}(\theta, \omega_0) \underline{\Delta} \begin{bmatrix} -b_1(\theta, \omega_0) & b_2(\theta, \omega_0) \\ b_2(\theta, \omega_0) & b_1(\theta, \omega_0) \end{bmatrix}, \quad \boldsymbol{\sigma}_3 \underline{\Delta} \begin{bmatrix} 0 & j \\ -j & 0 \end{bmatrix}$$

and the superscript H denotes the Hermitian (transpose conjugate).

5. DISCUSSION AND CONCLUSIONS

It is, in general, not possible for a weakly birefringent fiber to be polarization preserving due to the presence of 'linear coupling' between the two equations in (64) through the coefficient $b_2(\theta, \omega_0)$. This is so because even if one of the modes, say the mode with envelope B, were initially absent, i.e., $B(0, \tau) \equiv 0, B(\zeta, \tau)$ cannot continue to be zero for $\zeta > 0$ as the nonzero term $-jb_2(\theta, \omega_0)A$ in (64b) forces B to be nonzero for $\zeta > 0$.

It is instructive at this point to compare (66) with the corresponding equation for the complex vector envelope appearing as (46) in [5]. Apart from the fact that two independent coefficients $n_1(\omega_0)$ and $n_2(\omega_0)$ (as against Menyuk's just one) are needed to account for the effects of nonlinearity, and that there are neither any third order linear dispersion term involving $\beta'''(\omega_0)$ nor any birefringence term involving $\partial\mathbf{A}/\partial\tau$ in (66), the nature of the birefringence coefficients appearing in the two equations are quite different. A third order dispersion term involving $\beta'''(\omega_0)$ can appear only in a combined evolution equation for $\mathbf{A}(\zeta,\tau)+\varepsilon\mathbf{A}'(\zeta,\tau)$ if a nonzero ε-order correction to $A_m(Z_1,Z_2,T)$, $m=1,2$, is retained at the first stage of the perturbation by assuming a nonzero solution for $\mathbf{E}_1(x_\perp,z',t';Z_1,Z_2,T)$ in the form $\sum_{l=1}^{2} A'_l(Z_1,Z_2,T)\,\tilde{\mathbf{U}}_l(x_\perp,\omega_0)\,e^{j(\beta(\omega_0)z'-\omega_0 t')}+c.c.$, and determining the evolution equations for $A'_m(Z_1,Z_2,T)$, $m=1,2$, from the compatibility (solvability) conditions arising at the (third) ε^3-stage of the perturbation [15]. Birefringence terms involving $\partial\mathbf{A}/\partial\tau$ are absent from (66) because of the assumption of weak birefringence. It may be helpful at this stage to present the coefficients of the terms proportional to $\mathbf{A}$ in the two equations in tabular form in terms of the two matrices [16, 17] $\boldsymbol{\sigma}_1\underline{\Delta}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$ and $\boldsymbol{\sigma}_2\underline{\Delta}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$:

	Equation (66)	Menyuk's Equation (46) from [5]
Coefficients of $\mathbf{A}$	$\alpha(\omega_0)$	$-g(\omega_0)$
	$-[\sigma_{s1}(\omega_0)-\sigma_a(\omega_0)\cos 2\theta]\,\boldsymbol{\sigma}_1$ $+[\sigma_{s2}(\omega_0)-\sigma_a(\omega_0)\sin 2\theta]\,\boldsymbol{\sigma}_2$	$\Delta\beta\cos\theta\boldsymbol{\sigma}_1+\Delta\beta\sin\theta\boldsymbol{\sigma}_2$

A couple of observations may be in order: (i) Menyuk's equation does not appear to take fiber asymmetry into account and (ii) the sign of the coefficient multiplying the sine function is different in the two equations. The above fundamental differences in the form of the birefringence coefficients may be attributed to the way the fiber birefringence is modeled in the two approaches. Since our θ corresponds to Menyuk's $\theta/2$, the dependence of the trigonometric functions, appearing in the two expressions for the birefringence coefficients, on the polarization state will be the same.

The detailed analysis of the nonlinear propagation characteristics of optical pulses based on a numerical solution of the CNLSE will be presented in the second part of this paper.

REFERENCES

1. Matsumoto, M., Y. Akagi, and A. Hasegawa, "Propagation of solitons in fibers with randomly varying birefringence: Effects of soliton transmission control," *J. Lightwave Technol.*, Vol. 15, 584–589, Apr. 1997.
2. Xie, C., M. Karlsson, P. A. Andrekson, H. Sunnerud, and J. Lie, "Influences of polarization-mode dispersion on soliton

transmission systems," *IEEE J. Selected Topics in Quantum Electron.*, Vol. 8, No. 3, 575–590, 2002.

3. Muthuchidambaranathan, P., N. Kalyanasundaram, and D. Ravikumar, "Soliton propagation in birefringent optical fibers," *Fiber and Integrated Optics*, Vol. 27, 99–111, 2008.
4. Menyuk, C. R., "Nonlinear pulse propagation in birefringent optical fibers," *IEEE J. Quantum Electron.*, Vol. 23, 174–176, 1987.
5. Menyuk, C. R., "Application of multiple length-scale methods to the study of optical fiber transmission," *Journal of Engineering Mathematics*, Vol. 36, 113–136, 1999.
6. Hasegawa, A. and F. D. Tappert, "Transmission of stationary nonlinear pulses optical pulses in dispersive dielectric fibers. I. Anomalous dispersion," *Appl. Phys. Lett.*, Vol. 23, 142–144, 1973.
7. Menyuk, C. R., "Pulse propagation in an elliptically birefringent Kerr medium," *IEEE J. Quantum Electron.*, Vol. 25, 2674–2682, 1989.
8. Shen, Y. R., *The Principles of Nonlinear Optics*, Wiley, New York, 1984.
9. Agrawal, G. P., *Nonlinear Fiber Optics*, Academic, San Diego, 1995.
10. Wai, P. K. A. and C. R. Menyuk, "Polarization decorrelation in optical fibers with randomly varying birefringence," *Opt. Lett.*, Vol. 19, 1517–151, 1994.
11. Kodama, Y., "Optical solitons in monomode fiber," *J. Stat. Phys.*, Vol. 39, 597–614, 1985.
12. Marcuse, D., C. R. Menyuk, and P. K. A. Wai, "Applications of Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, Vol. 15, No. 9, 1735–1746, 1997.
13. Newell, A. C. and J. V. Moloney, *Nonlinear Optics*, Addison-Wesley Publishing Company, 1992.
14. Butcher, P. N. and D. Cotter, *The Elements of Nonlinear Optics*, Cambridege University Press, 1990.
15. Kodama, Y. and A. Hasegawa, "Nonlinear pulse propagation in a monomode dielectric guide," *IEEE Journal of Quantum Electronics*, Vol. 23, No. 5, 510–524, 1987.
16. Okamoto, K., *Fundamentals of Optical Waveguides*, Academic Press, 2006.
17. Hasegawa, A. and M. Matsumoto, *Optical Solitons in Fibers*, 3rd edition, Springer, 2003.