

Nonparaxial equation for linear and nonlinear optical propagation

B. Crosignani and P. Di Porto

Dipartimento di Fisica, Università dell'Aquila, 67010 L'Aquila, Italy

A. Yariv

Department of Applied Physics, California Institute of Technology, Pasadena, California 91125

Received January 31, 1997

The formalism of coupled-mode theory, specialized to the continuum of radiation modes, allows us to extend the standard parabolic wave equation to include nonparaxial terms and vectorial effects, and, in particular, to generalize the nonlinear Schrödinger equation that describes propagation in the presence of an intensity-dependent refractive index. © 1997 Optical Society of America

Most of the problems of beam propagation in optics are treated in the so-called paraxial approximation. Within this framework, after the electric field of a generic monochromatic electromagnetic wave is written in the form $E(\mathbf{r}, z, t) = A(\mathbf{r}, z)\exp(i\omega t - ikz)$, where $\mathbf{r} = (x, y)$, $k = (\omega/c)n_1 = k_0 n_1$ (n_1 being the unperturbed refractive index of the propagation medium), the evolution of the beam in the presence of a small refractive-index variation $\delta n(\mathbf{r}, z) \ll 1$ is described by the parabolic (or Fock–Leontovich) wave equation:

$$\left(\frac{\partial}{\partial z} + \frac{i}{2k}\nabla_{\perp}^2\right)A = -i\frac{k}{n_1}\delta nA, \quad (1)$$

where $\nabla_{\perp} = \hat{x}\partial/\partial x + \hat{y}\partial/\partial y$.

Equation (1) can be easily deduced starting from the Helmholtz equation for the electric field \mathbf{E} :

$$\nabla^2\mathbf{E} + 2\nabla(\mathbf{E} \cdot \nabla \ln n) + k_0^2 n^2 \mathbf{E} = 0, \quad (2)$$

which is in turn directly derived from Maxwell's equations written in the presence of a refractive index $n(\mathbf{r}, z) = n_1 + \delta n$. Once the expression of the electric field $\mathbf{E}(\mathbf{r}, z, t) = \mathbf{A}(\mathbf{r}, z)\exp(i\omega t - ikz)$ is inserted into Eq. (2), Eq. (1) is immediately recovered for any component A if one neglects both the term $\partial^2\mathbf{A}/\partial z^2$ (slowly varying approximation) and the polarization-scrambling term $2\nabla(\mathbf{E} \cdot \nabla \ln n)$ (scalar approximation). These approximations limit the transverse scale of variation w of \mathbf{A} (i.e., the beam size) to values much larger than λ and neglect any possible coupling between the x , y , and z components of the field. In particular, the second assumption prevents the divergence equation $\nabla \cdot (n^2\mathbf{E}) = 0$ from following as a consequence of Eq. (2),¹ so once the term $2\nabla(\mathbf{E} \cdot \nabla \ln n)$ is omitted, the divergence equation is no longer automatically satisfied by the solutions of Eq. (2).

The problem of describing beam propagation beyond the paraxial approximation was considered, mainly in the frame of numerical analysis, by Feit and Fleck,² starting again from Helmholtz equation (2) without the polarization-scrambling term, in connection with the analysis of catastrophic beam collapse in self-

focusing media. In this context, the problem was revisited recently, both analytically and numerically.^{3–5} These approaches make it possible, in particular, to write approximate equations that describe beam propagation in bulk media in the presence of the refractive-index variation $\delta n = n_2|E|^2$ associated with the Kerr nonlinearity but completely neglect the vectorial effects and thus are unable to satisfy the divergence equation.

Recently the effect of interaction between propagation and polarization was considered in a guiding structure for the description of intrafiber geometric rotation of polarization⁶ and, by application of an order-of-magnitude analysis method, in bulk nonlinear Kerr media.⁷

Here we deal with the problem of nonparaxial propagation from a new point of view. We express the propagating field as a superposition of transverse radiation modes of the unperturbed refractive index n_1 and look for its evolution in the presence of the perturbation associated with δn in the frame of coupled-mode theory. This approach, besides being able to describe propagation in the presence of any tensorial (linear or nonlinear) refractive-index variation δn at any order in the smallness parameter λ/w , offers, as opposed to Helmholtz's equation, the advantage of being inherently first order in $\partial/\partial z$ without the necessity of introducing the slowly varying approximation and automatically takes into account the vectorial effects associated with the polarization-scrambling term $2\nabla(\mathbf{E} \cdot \nabla \ln n)$. In fact, coupled-mode theory is an exact theory derived without neglecting this term.⁸

Following Ref. 9, we introduce the continuum set of orthogonally polarized normalized modes of the unbound space (radiation modes):

$$\mathbf{E}(\xi, 1; \mathbf{r})\exp(-i\beta_{\xi}z) = N_1 \exp(-i\xi \cdot \mathbf{r})[\hat{x} - (\xi_x/\beta_{\xi})\hat{z}] \times \exp(-i\beta_{\xi}z), \quad (3a)$$

$$\mathbf{E}(\xi, 2; \mathbf{r})\exp(-i\beta_{\xi}z) = N_2 \exp(-i\xi \cdot \mathbf{r}) \times [(\xi_x\xi_y/\beta_{\xi})\hat{x} - (\beta_{\xi} + \xi_x^2/\beta_{\xi})\hat{y} + \xi_y\hat{z}]\exp(-i\beta_{\xi}z), \quad (3b)$$

and $\mathbf{H}(\xi, \sigma; \mathbf{r}) = (1/\omega\mu_0)(\xi + \beta_\xi \hat{z}) \cdot \mathbf{E}(\xi, \sigma; \mathbf{r})$, $\sigma = 1, 2$, where the propagation constant β_ξ is given by $\beta_\xi = (k^2 - \xi_x^2 - \xi_y^2)^{1/2} = (k^2 - \xi^2)^{1/2}$ and ξ is restricted to values that satisfy the relation $0 \leq \xi \leq k$, whereas $N_1(\xi) = (1/2\pi)[\beta_\xi \omega \mu_0 / (\beta_\xi^2 + \xi_x^2)]^{1/2}$ and $N_2(\xi) = N_1(\xi)/k$. Note that the modes of the first and the second type are orthogonal in the usual geometric sense, i.e., $\mathbf{E}(\xi, 1; \mathbf{r}) \cdot \mathbf{E}(\xi, 2; \mathbf{r}) = 0$, and also satisfy the orthonormalization condition

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} d\mathbf{r} \hat{z} \cdot \mathbf{E}(\xi, \sigma; \mathbf{r}) \cdot \mathbf{H}^*(\xi', \sigma'; \mathbf{r}) = \delta^{(2)}(\xi - \xi') \delta_{\sigma, \sigma'}, \quad \sigma' = 1, 2. \quad (4)$$

The modes are also fully transverse, i.e., they are orthogonal to the propagation direction associated with the wave vector (ξ, β_ξ) , that is, $\mathbf{E}(\xi, \sigma; \mathbf{r}) \cdot (\xi + \beta_\xi \hat{z}) = 0$, or, equivalently, they obey the relation

$$\nabla \cdot [\mathbf{E}(\xi, \sigma; \mathbf{r}) \exp(-i\beta_\xi z)] = 0. \quad (5)$$

The forward-propagating electric field can then be written as

$$\begin{aligned} \mathbf{E}(\mathbf{r}, z, t) &= \sum_{\sigma} \iint d\xi \mathbf{E}(\xi, \sigma; \mathbf{r}) \\ &\times \exp(i\omega t - i\beta_\xi z) c(\xi, \sigma; z) \\ &= \sum_{\sigma} \exp(i\omega t - ikz) \iint d\xi \mathbf{E}(\xi, \sigma; \mathbf{r}) \\ &\times \exp[-i(\beta_\xi - k)z] c(\xi, \sigma; z) \\ &\equiv \sum_{\sigma} \exp(i\omega t - ikz) \mathbf{A}^{(\sigma)}(\mathbf{r}, z) \\ &\equiv \exp(i\omega t - ikz) \mathbf{A}(\mathbf{r}, z), \end{aligned} \quad (6)$$

where we have introduced the so-called slowly varying amplitudes $\mathbf{A}^{(\sigma)}(\mathbf{r}, z)$. Because of the formal properties of the continuum set of modes given in Eqs. (3), the z evolution of the expansion coefficients $c(\xi, \sigma; z)$ can be immediately obtained as a straightforward generalization of the evolution worked out for a set of discrete modes.¹⁰ Neglecting coupling with backward-propagating modes,¹¹ we see that the z evolution is

$$\begin{aligned} \frac{d}{dz} c(\xi, \sigma; z) &= -i\omega \epsilon_0 n_1 \sum_{\sigma'} \iint d\xi' \iint d\mathbf{r}' \\ &\times \exp[i(\beta_\xi - \beta_{\xi'})z] c(\xi', \sigma'; z) \\ &\cdot \mathbf{E}^*(\xi, \sigma; \mathbf{r}') \cdot [\delta n(\mathbf{r}', z) : \mathbf{E}(\xi', \sigma'; \mathbf{r}')], \end{aligned} \quad (7)$$

where, if we adopt the convention of summation over repeated indices, the scalar product $\mathbf{T} : \mathbf{V}$ of a tensor \mathbf{T} times a vector \mathbf{V} produces a vector of components $(\mathbf{T} : \mathbf{V})_i = T_{ij} V_j$.

Multiplying both sides of Eq. (7) by $\mathbf{E}(\xi, \sigma; \mathbf{r}) \exp[-i(\beta_\xi - k)z]$, after integrating over ξ , we get

$$\begin{aligned} &\left(\frac{\partial}{\partial z} + \frac{i}{2k} \nabla_{\perp}^2 - \frac{i}{8k^3} \nabla_{\perp}^4 + \dots \right) \mathbf{A}^{(\sigma)}(\mathbf{r}, z) \\ &= L\mathbf{A}^{(\sigma)} = -i\omega \epsilon_0 n_1 \iint d\mathbf{r}' \iint d\xi [\mathbf{E}(\xi, \sigma; \mathbf{r}) \\ &\times \mathbf{E}^*(\xi, \sigma; \mathbf{r}') \mathbf{E}^*(\xi, \sigma; \mathbf{r}') : [\delta n(\mathbf{r}', z) : \mathbf{A}(\mathbf{r}', z)]. \end{aligned} \quad (8)$$

For us to write the right-hand side of Eq. (8) in differential form, it is expedient to expand the tensors $\mathbf{E}(\xi, \sigma; \mathbf{r}) \mathbf{E}^*(\xi, \sigma; \mathbf{r}')$, ($\sigma = 1, 2$), as the power series in ξ_x and ξ_y . For example, for $\sigma = 1$ we have, recalling Eq. (3a),

$$\begin{aligned} \mathbf{E}(\xi, 1; \mathbf{r}) \mathbf{E}^*(\xi, 1; \mathbf{r}') &= N_1^2 \begin{bmatrix} 1 & 0 & -\xi_x/\beta_\xi \\ 0 & 0 & 0 \\ -\xi_x/\beta_\xi & 0 & \xi_x^2/\beta_\xi^2 \end{bmatrix} \\ &\times \exp[-i(\mathbf{r} - \mathbf{r}') \cdot \xi], \end{aligned} \quad (9)$$

from which we obtain, using the expression of N_1^2 ,

$$\begin{aligned} \mathbf{E}(\xi, 1; \mathbf{r}) \mathbf{E}^*(\xi, 1; \mathbf{r}') &= \frac{1}{(2\pi)^2} \sqrt{\mu_0/\epsilon_1} \\ &\times \begin{bmatrix} 1 + (\xi_y^2 - \xi_x^2)/2k^2 + \dots & 0 & -\xi_x/k + \dots \\ 0 & 0 & 0 \\ -\xi_x/k + \dots & 0 & \xi_x^2/k^2 + \dots \end{bmatrix} \\ &\times \exp[-i(\mathbf{r} - \mathbf{r}') \cdot \xi], \end{aligned} \quad (10)$$

where the ellipses stand for terms of order higher than the second in ξ_x/k , ξ_y/k . If we insert Eq. (10) into Eq. (8), we can calculate the integral over ξ , which consists of contributions of the form

$$\iint d\xi \exp[-i(\mathbf{r} - \mathbf{r}') \cdot \xi] \xi_{x,y}^m, \quad m = 1, 2, \quad (11)$$

which can be expressed as partial derivatives of order m with respect to x' and y' ($\xi_x \rightarrow -i\partial/\partial x'$ and $\xi_y \rightarrow -i\partial/\partial y'$) of the two-dimensional δ function

$$\delta^{(2)}(\mathbf{r} - \mathbf{r}') = \frac{1}{(2\pi)^2} \iint d\xi \exp[-i(\mathbf{r} - \mathbf{r}') \cdot \xi]. \quad (12)$$

A similar procedure applies when one is evaluating $\mathbf{E}(\xi, 2; \mathbf{r}) \mathbf{E}^*(\xi, 2; \mathbf{r}')$. Finally, exploiting the well-known properties of the two-dimensional δ function and of its derivatives, we can calculate the integral over \mathbf{r}' on the right-hand side of Eq. (8) and express it as a power series in the transverse partial derivatives of $\mathbf{A}(\mathbf{r}, z)$. Keeping only terms up to the second derivatives for simplicity, as we have done in Eq. (10), we obtain the resulting system of equations:

$$\begin{aligned} L\mathbf{A}_x &= -i(k/n_1)(\delta n : \mathbf{A})_x + (1/n_1)(\partial/\partial x)(\delta n : \mathbf{A})_z \\ &- i(1/kn_1)(\partial^2/\partial x \partial y)(\delta n : \mathbf{A})_y \\ &- i(1/2kn_1)(\partial^2/\partial x^2 - \partial^2/\partial y^2)(\delta n : \mathbf{A})_x, \end{aligned} \quad (13)$$

$$\begin{aligned}
LA_y &= -i(k/n_1)(\delta n : \mathbf{A})_y + (1/n_1)(\partial/\partial y)(\delta n : \mathbf{A})_z \\
&\quad - i(1/kn_1)(\partial^2/\partial x \partial y)(\delta n : \mathbf{A})_x \\
&\quad - i(1/2kn_1)(\partial^2/\partial y^2 - \partial^2/\partial x^2)(\delta n : \mathbf{A})_y, \quad (14)
\end{aligned}$$

$$\begin{aligned}
LA_z &= (1/n_1)\frac{\partial}{\partial x}(\delta n : \mathbf{A})_x + (1/n_1)\frac{\partial}{\partial y}(\delta n : \mathbf{A})_y \\
&\quad + i(1/kn_1)\nabla_{\perp}^2(\delta n : \mathbf{A})_z, \quad (15)
\end{aligned}$$

where the differential operator L is defined in Eq. (8).

The set of Eqs. (13)–(15) is the nonparaxial extension, which includes first- and second-order terms in the parameter λ/w of the parabolic Eq. (2). Note that the set is inherently first order in $\partial/\partial z$, and no approximation is required for neglecting the term in $\partial^2/\partial z^2$. This is contrasted with derivations that start with the Helmholtz equation.

Sometimes, instead of Eq. (15), it is more convenient to write

$$A_z = -\frac{ik}{\mu_0\omega^2\epsilon_{zz}}\nabla_{\perp} \cdot \mathbf{A}_{\perp} - \frac{\epsilon_{zx}}{\epsilon_{zz}}A_x - \frac{\epsilon_{zy}}{\epsilon_{zz}}A_y, \quad (16)$$

which is valid to the lowest significant order in λ/w and easily derived from Maxwell's equations, where we introduce the dielectric constant tensor $\epsilon = \epsilon_0 n_1(1 + 2\delta n)$.

One can specialize the set of Eqs. (13), (14), and (16) to describe nonlinear propagation in a Kerr medium by using the expression of the nonlinear refractive-index tensor δn , which is valid in the presence of the optical Kerr effect¹²:

$$\delta \mathbf{n} = \frac{2}{3}n_2 \begin{bmatrix} |\mathbf{E}|^2 + (1/2)|E_x|^2 & (1/2)E_y E_x^* & (1/2)E_z E_x^* \\ (1/2)E_x E_y^* & |\mathbf{E}|^2 + (1/2)|E_y|^2 & (1/2)E_z E_y^* \\ (1/2)E_x E_z^* & (1/2)E_y E_z^* & |\mathbf{E}|^2 + (1/2)|E_z|^2 \end{bmatrix}, \quad (17)$$

where n_2 is the nonlinear refractive-index coefficient appearing in the standard expression for the intensity-dependent refractive index $n = n_1 + n_2|E|^2$ (optical Kerr effect). In this case, Eq. (16) can be rewritten, neglecting higher-order terms, in the simple form $A_z \approx -(i/k)\nabla_{\perp} \cdot \mathbf{A}_{\perp}$.

If we consider, for the sake of simplicity, the case $A_y = 0$ [note that A_y will eventually become nonnegligible, because Eq. (14) is not homogeneous in A_y , also if $A_y(z=0) = 0$], it is possible to derive, after straightforward if somewhat tedious algebra, the following equation of evolution:

$$\begin{aligned}
&\left[\frac{\partial}{\partial Z} + \frac{i}{2} \left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} \right) - \frac{i}{8} \left(\frac{\partial^4}{\partial X^4} + \frac{\partial^4}{\partial Y^4} \right) \right] u \\
&\quad + i|u|^2 u = -\frac{i}{2} \left(\frac{\partial^2}{\partial X^2} - \frac{\partial^2}{\partial Y^2} \right) (|u|^2 u) - \frac{i}{3} \left| \frac{\partial u}{\partial X} \right|^2 u \\
&\quad - \frac{i}{3} \left(|u|^2 \frac{\partial^2 u}{\partial X^2} - u^2 \frac{\partial^2 u^*}{\partial X^2} \right), \quad (18)
\end{aligned}$$

where $u = (n_2/n_1)^{1/2}A_x$ and $(X, Y, Z) = k(x, y, z)$, which generalizes the standard nonlinear Schrödinger equation through the terms that appear on its right-hand side to include nonparaxial contributions up to the second order.

We have derived a new set of equations that generalize the standard parabolic equation that describes paraxial propagation by the addition of nonparaxial terms up to the second order in the ratio λ/w between the wavelength and the typical dimension of the beam (the formalism can easily provide the nonparaxial terms to any order in λ/w). This has been accomplished with the coupled-mode theory formalism, applied to the continuum of transverse radiation modes, which automatically includes vectorial effects and is inherently first order in $\partial/\partial z$ without any slowly varying approximation hypothesis.

B. Crosignani gratefully acknowledges the support of collaborative NATO research grant 930240.

References

1. M. Lax, W. H. Louisell, and W. B. McKnight, *Phys. Rev. A* **11**, 1365 (1975).
2. M. D. Feit and J. A. Fleck, Jr., *J. Opt. Soc. Am. B* **5**, 633 (1988).
3. N. Akhmediev, A. Ankiewicz, and J. M. Soto-Crespo, *Opt. Lett.* **18**, 411 (1993).
4. J. M. Soto-Crespo and N. Akhmediev, *Opt. Commun.* **101**, 223 (1993).
5. G. Fibich, *Phys. Rev. Lett.* **76**, 4356 (1996).
6. A. Yu. Savchenko and B. Ya. Zel'dovich, *J. Opt. Soc. Am. B* **13**, 273 (1996).
7. S. Chi and Q. Guo, *Opt. Lett.* **20**, 1598 (1995).
8. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1974), Chap. 3.
9. B. Crosignani and A. Yariv, *J. Opt. Soc. Am.* **1**, 1034 (1984).
10. D. Marcuse, *Bell Syst. Tech. J.* **54**, 985 (1975).
11. The set of equations given by Eq. (8) can be generalized in a straightforward way to include backward-propagating modes, the resulting set of equations being completely equivalent to Maxwell's equation and still first order in z .
12. B. Crosignani, A. Cutolo, and P. Di Porto, *J. Opt. Soc. Am.* **72**, 1136 (1982).