Vector theory of self-focusing of an optical beam in Kerr media

Sien Chi and Qi Guo*

Institute of Electro-Optical Engineering, National Chiao Tung University, Hsinchu, Taiwan, China

Received March 6, 1995

The scalar theory of the self-focusing of an optical beam is not valid for a very narrow beam, and a correct description of the beam behavior requires a vector analysis in this case. A vector nonparaxial theory is developed from the vector Maxwell equations by application of an order-of-magnitude analysis method. For the same input beam, the numerical results of self-focusing from both scalar and vector theories are compared. It is found by the vector theory that a linearly polarized circular input beam becomes elliptical in the self-focusing process. © 1995 Optical Society of America

The self-focusing effects of an optical beam in nonlinear media have been studied extensively for more than three decades.[1–8] The early theory, based on a paraxial wave equation, predicts[2,3] the catastrophic collapse of a self-focusing beam in a Kerr medium. Feit and Fleck[6] pointed out that this unphysical collapse is due to the invalidity of the paraxial wave equation in the neighborhood of a self-focus. They applied a nonparaxial algorithm for the scalar wave equation to describe the self-focusing of the beam, and their results showed that the self-focusing is noncatastrophic. Using the same scalar wave equation and assuming the solution with slowly varying amplitude and fast oscillating phase, Akhmediev and co-workers[7,8] developed another nonparaxial method, and their results agree qualitatively with those of Feit and Fleck.

In this Letter we study the beam self-focusing starting from the vector wave equation. A new model is established by an order-of-magnitude analysis method, and the basic idea of this method can be found in Ref. 9.

If the electric-field intensity vector $\mathbf{E}(\mathbf{r}, t)$, the magnetic field intensity vector $\mathbf{H}(\mathbf{r}, t)$, the electric displacement vector $\mathbf{D}(\mathbf{r}, t)$, and the nonlinear polarization vector $\mathbf{P}_{\rm NL}(\mathbf{r}, t)$ are assumed to be $\mathbf{E}(\mathbf{r}, t) = (1/2)\mathbf{E}(\mathbf{r})\exp(-i\omega t) + \text{c.c.}$, $\mathbf{H}(\mathbf{r}, t) = (1/2)\mathbf{H}(\mathbf{r})\exp(-i\omega t) + \text{c.c.}$, $\mathbf{D}(\mathbf{r}, t) = (1/2)\mathbf{D}(\mathbf{r})\exp(-i\omega t) + \text{c.c.}$, and $\mathbf{P}_{\rm NL}(\mathbf{r}, t) = (1/2)\mathbf{P}_{\rm NL}(\mathbf{r})\exp(-i\omega t) + \text{c.c.}$, then, in the absence of free charges in a nonlinear nonmagnetic isotropic medium, time-harmonic Maxwell equations in the mks system are

$$\nabla \times \mathbf{E}(\mathbf{r}) = i\omega\mu_0\mathbf{H}(\mathbf{r}), \quad \nabla \times \mathbf{H}(\mathbf{r}) = -i\omega\mathbf{D}(\mathbf{r}), \tag{1a}$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = 0 \tag{1b}$$

and the constitutive relation reads as $\mathbf{D}(\mathbf{r}) = \varepsilon_0 n_0^2\mathbf{E}(\mathbf{r}) + \mathbf{P}_{\rm NL}(\mathbf{r})$, where n_0 is a linear refractive index, $\mathbf{P}_{\rm NL}(\mathbf{r})$ expressed as $(P_{\rm NL})_i = (3\varepsilon_0/4)\sum_{j,k,l}\chi^{(3)}_{ijkl}(\omega = \omega + \omega - \omega)E_jE_kE_l^*$ is the third-order nonlinear polarization[10–12] (here the subscripts i, j, k, and l refer to the Cartesian components of the fields), and the fourth-rank tensor $\chi^{(3)}(\omega = \omega_1 + \omega_2 + \omega_3)$ is the Fourier transform of the third-order nonlinear susceptibility.[10] From Eqs. (1) and the constitutive relation we have the vector wave equation

$$\nabla^2\mathbf{E} + \frac{\omega^2 {n_0}^2}{c^2}\mathbf{E} + \frac{1}{{n_0}^2\varepsilon_0}\nabla(\nabla \cdot \mathbf{P}_{\rm NL}) + \frac{\omega^2}{c^2\varepsilon_0}\mathbf{P}_{\rm NL} = 0\,. \tag{2}$$

If $\mathbf{E}$ is linearly polarized in the x direction and propagates along the z direction, then if we neglect the $\nabla(\nabla \cdot \mathbf{P}_{\rm NL})$ term Eq. (2) reduces to the scalar wave equation

$$\nabla_T{}^2 E_x + \frac{\partial^2}{\partial z^2}E_x + \frac{\omega^2}{c^2}n^2E_x = 0\,, \tag{3}$$

where the subscript T represents the transverse component of a vector (that is, $\nabla_T{}^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$), $n = n_0 + n_2|E_x|^2$ is the total refractive index, and the Kerr coefficient is $n_2 = 3\chi^{(3)}_{xxxx}/(8n_0)$. Equation (3) is the basic equation of the scalar theory on the beam self-focusing.[2–8] Assuming that the scalar field is $E_x(x, y, z) = A(x, y, z)\exp(ik_0z)$ and using the paraxial approximation condition to drop the $\partial^2A/\partial z^2$ term, from Eq. (3) we obtain the paraxial wave equation[2–5]:

$$i\frac{\partial A}{\partial z} + \frac{1}{2k_0}\nabla_T{}^2A + \gamma|A|^2A = 0\,, \tag{4}$$

where $k_0 = \omega n_0/c$ and $\gamma = k_0n_2/n_0$.

The fundamental relations of the order of magnitude can easily be obtained from the scalar paraxial wave equation (4). Replacing the ∇_T operator by a derivative with respect to x or y, we have

$$\nabla_TA \sim \frac{\partial}{\partial x}A\left(\text{or } \frac{\partial}{\partial y}A\right) \sim \frac{A}{w}\,, \qquad \nabla_T{}^2A \sim \frac{A}{w^2}\,, \tag{5a}$$

where w is the beam width. Because the behavior of the paraxial beam self-focusing is determined equally by the three terms of Eq. (4), the three terms must be of the same order of magnitude. Therefore, from $(\partial A/\partial z)/(\nabla_T{}^2A/k_0) \sim 1$ and $(\gamma|A|^2A)/(\nabla_T{}^2A/k_0) \sim 1$, we can derive

$$\frac{\partial}{\partial z}A \sim \frac{A}{k_0w^2} = \frac{\sigma}{w}A, \quad \frac{n_2}{n_0}|A|^2 \sim \frac{1}{{k_0}^2w^2} = \sigma^2, \tag{5b}$$

where $\sigma = 1/(k_0w) = \lambda/(2\pi w)$. Generally, σ is very small ($\ll 1$). Even if the beam is focused to $w \approx$

λ, σ is only approximately 0.16. From the first of relations (5b), we have

$$\frac{\partial}{\partial z} A/k_0 A \sim \frac{\sigma}{w} A/k_0 A = \sigma^2 . \quad (5c)$$

Now we go back to Eq. (2) to develop the vector theory on self-focusing. If we write the vector electric field $\mathbf{E}(\mathbf{r})$ as

$$\mathbf{E}(x, y, z) = \mathbf{A}(x, y, z)\exp(ik_0 z), \quad (6)$$

then Eq. (2) becomes

$$\nabla_T{}^2\mathbf{A} + \frac{\partial^2}{\partial z^2}\mathbf{A} + 2ik_0\frac{\partial}{\partial z}\mathbf{A} + \left[\frac{1}{n_0{}^2\varepsilon_0}\nabla(\nabla\cdot\mathbf{P}_{\mathrm{NL}}) + \frac{\omega^2}{c^2\varepsilon_0}\mathbf{P}_{\mathrm{NL}}\right]\exp(-ik_0 z) = 0 . \quad (7)$$

Substitution of Eq. (6) into Eq. (1b) yields

$$\left(\nabla_T\cdot\mathbf{A}_T + ik_0A_z + \frac{\partial}{\partial z}A_z\right)\exp(ik_0 z) + \frac{1}{n_0{}^2\varepsilon_0}\nabla\cdot\mathbf{P}_{\mathrm{NL}} = 0 . \quad (8a)$$

We further assume that the components of vector beam $\mathbf{A}$ also satisfy the fundamental relations (5); then we can prove that $(\nabla\cdot\mathbf{P}_{\mathrm{NL}}/n_0{}^2\varepsilon_0)/(k_0A_z) \sim \sigma^2$ and $\nabla_T\cdot\mathbf{A}_T/(k_0A_z) \sim 1$. Therefore Eq. (8a) reads as

$$A_z = \frac{i}{k_0}\nabla_T\cdot\mathbf{A}_T[1 + O(\sigma^2)], \quad (8b)$$

where the symbol $O(\sigma^2)$ means that the terms neglected are 2 orders of σ less than the term kept. Then it can easily be found that $A_z/A_T = (i\nabla_T\cdot\mathbf{A}_T/k_0)/A_T \sim (A_T/k_0w)/A_T = \sigma$. This means that A_z is 1 order of σ smaller than A_T. In fact, it has been found even for the linear case that both the electric and the magnetic fields have a small longitudinal components in addition to the transverse components. We can also prove from Eqs. (1a) that, if the input beam is linearly polarized along x direction at $z = 0$, although the nonlinear coupling will excite the y component, A_y will be so small that $A_y/A_x \sim \sigma^2$. Therefore A_y can be neglected. Then, for the special case that the nonlinear refractive index results from the nonresonant electronic nonlinearity,[11] by means of the symmetry property of the third-order nonlinear susceptibility in an isotropic medium[10,11] the x component of $\mathbf{P}_{\mathrm{NL}}$ can be reduced to

$$(P_{\mathrm{NL}})_x = 2\varepsilon_0 n_0 n_2 \exp(ik_0 z)\left(|A_x|^2A_x + \frac{2}{3}|A_z|^2A_x + \frac{1}{3}A_z{}^2A_x{}^*\right)[1 + O(\sigma^4)] . \quad (9)$$

It can be found easily by Eq. (8b) that the last two terms in Eq. (9) are an order of σ^2 smaller than the first term.

Using $A_z = (i/k_0)\partial A_x/\partial x$ in Eq. (9), substituting Eq. (9) into Eq. (7), and retaining all terms to the order of σ^2 smaller than the $\nabla_T{}^2A_x$ term, from the x component of Eq. (7) we obtain

$$i\frac{\partial}{\partial z}A_x + \frac{1}{2k_0}\nabla_t{}^2A_x + \gamma|A_x|^2A_x = -\frac{1}{2k_0}\frac{\partial^2}{\partial z^2}A_x - \frac{\gamma}{k_0{}^2}\left[\frac{\partial^2}{\partial x^2}(|A_x|^2A_x) + \frac{2}{3}\left|\frac{\partial}{\partial x}A_x\right|^2A_x - \frac{1}{3}\left(\frac{\partial}{\partial x}A_x\right)^2A_x^*\right] . \quad (10)$$

From relations (5) we know that the terms on the left-hand side of Eq. (10) are of equal order and all terms on the right-hand side are of the same order and are σ^2 smaller than the left-hand-side terms. In Eq. (10) all terms dropped are at least an order of σ^4 smaller than the left-hand-side terms. The second term on the right-hand side is the contribution of $\nabla\cdot\mathbf{E}$, and the last two terms come from the coupling between E_x and E_z. Feit and Fleck[6] and Akhmediev and co-workers[7,8] considered only the first term on the right-hand side.

Because all the terms on the right-hand side of Eq. (10) are of the order of σ^2 smaller than the left-hand-side terms, the evolution of a beam that has a large enough initial w (i.e., a small initial σ) is governed by the paraxial wave equation [the left-hand side of Eq. (10)] in the initial stage of the self-focusing. As w becomes smaller, σ gets bigger. Therefore the effect of all terms on the right-hand side can no longer be neglected. By the normalized transformation

$$A_x(x, y, z) = \sigma\sqrt{n_0/n_2}\,u(\xi, \eta, \zeta), (x, y, z) = (w\xi, w\eta, 1\zeta), \quad (11)$$

where $l = k_0w^2$ is the diffraction length, Eq. (10) becomes

$$i\frac{\partial}{\partial\zeta}u + \frac{1}{2}\nabla^2_{\xi,\eta}u + |u|^2u = -\sigma^2\left\{\frac{1}{2}\frac{\partial^2}{\partial\zeta^2}u + \frac{\partial^2}{\partial\xi^2}\times\left[|u|^2u + \frac{2}{3}\left|\frac{\partial}{\partial\xi}u\right|^2u - \frac{1}{3}\left(\frac{\partial}{\partial\xi}u\right)^2u^*\right]\right\}, \quad (12)$$

where $\nabla^2_{\xi,\eta} = \partial^2/\partial\xi^2 + \partial^2/\partial\eta^2$. In Eq. (12) all terms except the coefficients are of the same order, approximately 1, and the order-of-magnitude relation of the terms is more clearly manifested.

To use the split-step Fourier method (SSFM) to solve Eq. (12) numerically, we must replace the term $\partial^2u/\partial\zeta^2$ by the transverse derivative. Partial derivation of Eq. (12) with respect to ζ gives the expression $\partial^2u/\partial\zeta^2$. Again substituting this expression into Eq. (12) and retaining all terms to the order of σ^2, we have

$$\frac{\partial}{\partial\zeta}u = i\frac{1}{2}\nabla^2_{\xi,\eta}u - i\frac{\sigma^2}{8}\nabla^4_{\xi,\eta}u + i|u|^2u - i\sigma^2N_hu + O(\sigma^4), \quad (13)$$

where $N_h = |u|^4/2 + u^*(\partial^2u/\partial\eta^2 - \partial^2u/\partial\xi^2) + |\partial u/\partial\eta|^2 - 11|\partial u/\partial\xi|^2/3 + u^*[(\partial u/\partial\eta)^2 - 7(\partial u/$

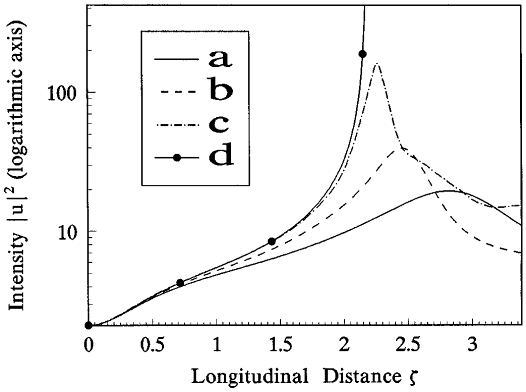

Fig. 1. On-axis normalized intensity versus normalized longitudinal distance ζ for the initial condition: curve a, by our vector model; curve b, by Feit and Fleck's method; curve c, by Akhmediev and co-workers' method; curve d, by the paraxial wave equation.

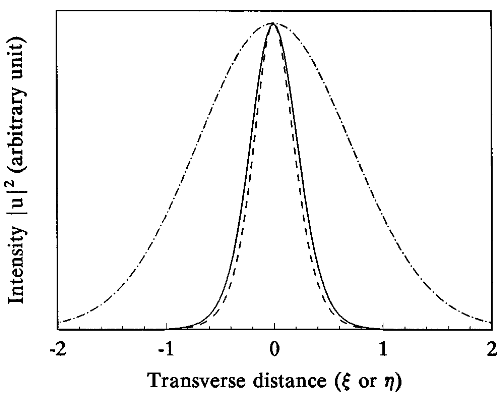

Fig. 2. Transverse normalized intensity distribution at $\zeta = 0$ (dashed–dotted curve), along the x axis at $\zeta = \zeta_{\max}$ (solid curve), and along the y axis at $\zeta = \zeta_{\max}$ (dashed curve). The input intensity is magnified for a comparison of the beam widths.

$\partial\xi)^2/3]/(2u) - u\partial^2 u^*/\partial\xi^2$. We developed a computer program based on Eq. (13) by the SSFM.[12] In the program we insert σ_0 and w_0 at $z = 0$ for σ and w to relate the initial real physical parameters to the initial normalized parameters. The validity of our program is verified by repeated testing with different transverse grid and longitudinal step length.

To compare the results of our model and those of the two different scalar nonparaxial models, we use the same input beam parameters for numerical simulations. Figure 1 gives the on-axis intensity obtained by the three models for the beam that initially has a symmetric Gaussian shape $u(\xi, \eta, \zeta = 0) = u_0 \exp[-(\xi^2 + \eta^2)/2]$ when $u_0 = 1.4575$ and $\sigma_0 = 0.0697$. The paraxial approximation result for the same parameters is also given. The figure shows that, although all three nonparaxial models give the noncatastrophic results, the quantitative difference between them is obvious.

The most surprising feature is that for an initial transverse symmetric shape the beam described by Eq. (10) will become transverse asymmetrical after some propagation distance. This feature can be directly observed from Eq. (10) or (12), where the last three terms show an anisotropiclike property. Figure 2 gives the initial normalized transverse intensity distribution and the distribution at $\zeta_{\max}$, where the on-axis intensity reaches a maximum. The figure shows that for the initial symmetric beam linearly polarized along the x axis its beam width w_y along the y axis is smaller than its beam width along the x axis w_x at $\zeta_{\max}$. It is evident that σ will reach $\sigma_{\max}$ at the same point $\zeta_{\max}$, since w_x and w_y are minimum at this point. For the parameter given in Fig. 2, $\sigma_{\max} = (\sigma_y)_{\max} = w_0\sigma_0/(w_y)_{\min} \approx 0.27$. In fact, for different initial values, it is shown that $\sigma_{\max}$ is less than 0.3. As pointed out above, the neglected terms in Eq. (10) are of the order of σ^4, and the next higher-order terms kept are of the order of σ^2. This indicates that the maxima of the dropped terms are approximately 0.01 in order of magnitude and they are at least 1 order smaller than terms kept from the beginning to the end. Therefore our vector nonparaxial model is self-consistent.

In conclusion, it is shown that the scalar theory on the self-focusing of the optical beam is no longer valid for a very narrow optical beam, and the vectorial analysis is needed to describe the beam behavior correctly in this case. The new model for this purpose is established directly from the vector wave equation by means of the order-of-magnitude analysis method. Our numerical results also show noncatastrophic self-focusing but with less maximum on-axis intensity than that of Feit and Fleck and Akhmediev and co-workers; in addition, our results show that the linearly polarized circular input beam will become elliptic in the self-focusing process.

This research was supported by the National Science Council of China under contract NSC 84-2215-E009-016. The authors thank P. Yeh for illuminating discussions.

*Permanent address, Institute of Quantum Electronics, South China Normal University, Guangzhou, China.

References

1. R. Y. Chiao, E. Garmire, and C. H. Townes, Phys. Rev. Lett. **13,** 479 (1964).
2. P. L. Kelley, Phys. Rev. Lett. **15,** 1005 (1965).
3. W. G. Wagner, H. A. Haus, and J. H. Marburger, Phys. Rev. **175,** 256 (1968).
4. For a comprehensive review of experimental work on self-focusing see Y. R. Shen, Prog. Quantum Electron. **4,** 1 (1975).
5. For a comprehensive review of theoretical work on self-focusing see J. H. Marburger, Prog. Quantum Electron. **4,** 35 (1975); see also N. C. Kothari and S. C. Abbi, Prog. Theor. Phys. **83,** 414 (1990).
6. M. D. Feit and J. A. Fleck, Jr., J. Opt. Soc. Am. B **5,** 633 (1988).
7. N. Akhmediev, A. Ankiewicz, and J. M. Soto-Crespo, Opt. Lett. **18,** 411 (1993).
8. J. M. Soto-Crespo and N. Akhmediev, Opt. Commun. **101,** 223 (1993).
9. D. Marcuse, *Light Transmission Optics*, 2nd ed. (Van Nostrand Reinhold, New York, 1982), pp. 9–11.
10. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984), pp. 1–5, 25–29.
11. R. W. Boyd, *Nonlinear Optics* (Academic, New York, 1992), Chaps. 4.2 and 4.3.
12. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, Boston, Mass., 1989), Chap. 2.