JOURNAL OF MODERN OPTICS, 1998, VOL. 45, NO. 6, 1111–1121

Non-paraxial solitons

P. CHAMORRO-POSADA†, G. S. McDONALD and G. H. C. NEW

Laser Optics and Spectroscopy Group, The Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, England

(*Received 15 September 1997*)

Abstract. In this paper, we propose the use of ultranarrow soliton beams in miniaturized nonlinear optical devices. We derive a nonparaxial nonlinear Schrödinger equation and show that it has an *exact* non-paraxial soliton solution from which the paraxial soliton is recovered in the appropriate limit. The physical and mathematical geometry of the non-paraxial soliton is explored through the consideration of dispersion relations, rotational transformations and approximate solutions. We highlight some of the unphysical aspects of the paraxial limit and report modifications to the soliton width, the soliton area and the soliton (phase) period which result from the breakdown of the slowly varying envelope approximation.

1. Introduction

The potential applications of nonlinear optical devices in the field of information technology (IT) has motivated enormous research effort in recent years. A great deal of attention has centred on the utilization of Kerr-type nonlinearities to generate soliton beams which do not undergo diffractive spreading. These beams are, strictly, only defined for two-dimensional (2D) Kerr media. The two dimensions concerned are spatial and define a plane which contains the propagation axis of the beam and an axis which is orthogonal to this. Beam propagation in 2D Kerr media is studied by means of the nonlinear Schrödinger equation (NSE) which is analytically solvable using inverse scattering techniques [1]. On the experimental side, soliton beams have been studied in different 2D Kerr-like materials where the light is allowed to develop structure in only one of the coordinates transverse to the direction of propagation [2–7]. Spatial solitons have an innate appeal as binary information elements since each element is not only self-trapped but also self-stabilizing. It is for this reason that they have been proposed for use in optical information processing and storage devices [8–10]. In this paper, we generalize the theory of spatial solitons in 2D Kerr media to include ultranarrow soliton beams and we propose the use of these beams in miniaturized nonlinear photonic devices.

To determine the maximum possible information density that can be stored or processed, one needs to understand fully the interplay of the nonlinearity of the

† Permanent address: Departamento de Teoría de la Señal y Comunicaciones e Ingenieria Telemática, Escuela Tećnica Superior de Ingenerieros de Telecomunicación, Universidad de Valladolid, 47011 Valladolid, Spain.

0950–0340/98 $12·00

device material and the diffraction of the propagating light. The balance point between these two effects defines a fundamental limit to the minimum possible size of spatial soliton information units. For light in 2D media, diffraction may occur in both the transverse and the longitudinal directions. The formulation of such 2D diffraction involves the coordinates of these axes in a symmetric way. Indeed, the coordinates are on an *equal* footing. Beam paraxiality is assumed when deriving the NSE and this assumption breaks the symmetry between the two spatial dimensions. In waveguide geometry, the paraxial approximation permits diffraction of light in only the transverse direction. The resulting effect can be thought of as one-dimensional diffraction.

However, in the context of the progressive miniaturization of IT devices, the paraxial approximation may be violated since the optical wavelength λ may not continue to be of vanishing magnitude in comparison with the width of the beam. There are also other, more general, situations where the paraxial approximation may break down. We find that one example is when a soliton beam propagates at a significant angle to the reference axis of the evolution equation or, equivalently, when two soliton beams interact at such an angle. Another example arises from the fact that the NSE supports higher-order soliton solutions of *arbitrary* order. While the launched beam for a higher-order soliton may be reasonably paraxial, subsequent evolution can involve stages of sufficiently strong focusing that non-paraxial effects become important.

Since solitons exist in real physical systems where (additional) higher-order effects often come into play, the role of such effects can be of central importance. Usually, the study of such effects results in only approximate solutions or recourse is taken to purely numerical investigations. In this paper, we consider the higher-order effect of non-paraxiality, which is quite general and fundamental to the light itself. Other workers have examined the consequences of the breakdown of the slowly varying envelope approximation in the context of nonlinear *pulse* propagation [11, 12]. Their approach was to solve a full set of nonlinear Maxwell's equations by numerical techniques and, in each case, soliton-like pulses were reported. However, since their approach results only in simulation data, a limited amount of physical insight can be gained. This paper deals with nonlinear *beam* propagation and our approach is quite complementary to previous numerical studies. We make a systematic generalization of well known results from spatial soliton theory and accommodate the breakdown of the slowly varying envelope approximation within this framework. Both exact and approximate analytical solutions are derived and the underlying geometry of these new solutions is discussed in detail.

2. Non-paraxial nonlinear Schrödinger equation

The 2D time-independent field envelope $\tilde{E}(x, z)$ of a continuous-wave monochromatic beam with carrier frequency ω is related to the total electric field $E(x, z, t)$ through

$$E(x, z, t) = \tfrac{1}{2}\tilde{E}(x, z) \exp(-\mathrm{i}\omega t) + \mathrm{cc}, \tag{1}$$

where cc denotes complex conjugate and x, z and t are the transverse, longitudinal and time coordinates respectively. In a medium which has a nonlinear refractive

index $n(\tilde{E})$, the field envelope obeys the nonlinear 2D Helmholtz equation

$$\frac{\partial^2 \tilde{E}}{\partial z^2} + \frac{\partial^2 \tilde{E}}{\partial x^2} + \frac{\omega^2}{c^2} n^2(\tilde{E})\tilde{E} = 0. \tag{2}$$

Introducing a normalization appropriate to a forward propagating beam, $\tilde{E}(x, z) = A(x, z)\exp(\mathrm{i}kz)$, a Kerr nonlinearity $n(\tilde{E}) = n_0 + n_2|\tilde{E}|^2$ and assuming that the approximation $n^2(\tilde{E}) \approx n_0^2 + 2n_0 n_2|\tilde{E}|^2$ is justified by the relatively low value of n_2 in the material, we derive the following non-paraxial nonlinear Schrödinger equation (NNSE):

$$\kappa \frac{\partial^2 u}{\partial \zeta^2} + \mathrm{i}\frac{\partial u}{\partial \zeta} + \frac{1}{2}\frac{\partial^2 u}{\partial \xi^2} + |u|^2 u = 0, \tag{3}$$

where we have employed the following normalizations:

$$\zeta = \frac{z}{L_\mathrm{D}}, \quad \xi = \frac{2^{1/2}x}{w_0}, \quad u(\xi, \zeta) = \left(\frac{kn_2 L_\mathrm{D}}{n_0}\right)^{1/2} A(\xi, \zeta). \tag{4}$$

w_0 is a transverse scale parameter that we shall later relate to the width of non-paraxial soliton beams. This scale parameter can also be considered as equivalent to the waist of a (reference) paraxial Gaussian beam, at $\zeta = 0$, which has a diffraction length $L_\mathrm{D} = kw_0^2/2$. This reference beam is a solution of equation (3) in the limit where both κ and $|u|^2 \to 0$. The propagation constant is defined as $k = n_0 k_0 = n_0\omega/c$, where c is the speed of light. Since non-paraxiality arises here from *linear* 2D diffraction, it is natural to find that the non-paraxial parameter κ of the NNSE depends only on linear variables: $\kappa = 1/k^2 w_0^2$. Two alternative expressions for κ lend their own particular insight into the character of the non-paraxial parameter:

$$\kappa = \frac{\tan^2 \Theta}{4} = \frac{1}{4\pi^2 n_0^2}\left(\frac{\lambda}{w_0}\right)^2. \tag{5}$$

In the first form, Θ is the far-field angle of the reference Gaussian beam and reflects the degree of spread that this beam would experience under purely linear paraxial propagation. The second expression gives κ directly in terms of the number of optical wavelengths that are present in w_0. In free space, $\kappa = 10^{-3}$ implies around ten wavelengths in the full width $2w_0$ of the reference Gaussian, while $\kappa = 10^{-4}$ and $\kappa = 10^{-5}$ give around 32 and 100 optical cycles respectively.

3. The fundamental soliton

3.1. *Paraxial case*

The NSE can be recovered from equation (3) in the limit $\kappa \to 0$. This limit corresponds to the physical constraint that $w_0/\lambda \to \infty$, as opposed to specifying a magnitude of the longitudinal rate of change of the gradient of the amplitude or phase of the beam. The resulting (paraxial) NSE is

$$\mathrm{i}\frac{\partial u}{\partial \zeta} + \frac{1}{2}\frac{\partial^2 u}{\partial \xi^2} + |u|^2 u = 0. \tag{6}$$

Equation (6) has a fundamental paraxial soliton solution given by [1, 13]

$$u(\xi,\zeta) = \eta \,\mathrm{sech}\,[\eta(\xi + V\zeta)] \exp\left(-\mathrm{i}V\xi + \frac{\mathrm{i}}{2}(\eta^2 - V^2)\zeta\right), \tag{7}$$

where the amplitude (or energy) parameter η and the transverse velocity parameter V of the soliton correspond to the imaginary and real parts, respectively, of the eigenvalues of the associated scattering problem [13]. The solution is a bell-shaped beam that preserves its intensity profile with propagation and, in general, has a linear transverse phase profile.

The lateral width of this beam is fixed by the input amplitude η and is given, in normalized units, by $\xi_0 = 1/\eta$. Thus, the *soliton area*, which is proportional to $\eta\xi_0$, is conserved during propagation. The particular case of a beam which propagates along the ζ axis is found by setting $V = 0$, whereby the solution takes the simpler form

$$u(\xi,\zeta) = \eta \,\mathrm{sech}\,(\eta\xi) \exp\left(\frac{\mathrm{i}}{2}\eta^2\zeta\right). \tag{8}$$

Characteristic spatial lengths can be defined for the canonical forms of the fundamental and higher-order NSE solitons [13]. In contrast with the canonical fundamental soliton $(\eta, V) = (1, 0)$, which preserves its shape during propagation, each of the higher-order solitons are periodic in ζ, with a soliton (amplitude) period of $\zeta_\mathrm{A} = \pi/2$. On the other hand, both the canonical form of solution (8) and, at least, that of the first higher-order soliton have a *soliton* (*phase*) *period* which is given by $\zeta_\mathrm{P} = 4\pi$.

While the inverse and direct scattering problems predict the parameters of soliton formation from near-arbitrary input beam profiles, in the paraxial case, large propagation distances may be required to resolve the soliton components from the (untrapped) diffractive radiation modes. Thus it can be preferable to launch a beam into the nonlinear medium which has a spatial profile that closely matches the desired solution. For example, to generate an exact paraxial soliton of unit amplitude and finite velocity at $\zeta = 0$, the required input profile is

$$u(\xi, 0) = \mathrm{sech}\,(\xi) \exp\,(-\mathrm{i}V\xi). \tag{9}$$

The physical origin of the transverse phase profile in the above expression is the transverse variation in the phase of the carrier wave. This occurs quite naturally from the angle of the launched beam. However, the experimental generation of an exact $V = 0$ soliton requires extremely high interferomic accuracy. In terms of numerical work, the simulation of beams with $V \neq 0$ requires a sufficient number of transverse data points to sample *each* of the transverse spatial periods over which the phase changes by 2π. Hence, in the simulation of one or more paraxial solitons, for which the optical wavelength is negligible when compared with the soliton width, the requirement to have a reasonably finite number of transverse grid points restricts considerations to a very small number of configurations in which the angle involved is almost vanishingly small.

3.2. *Non-paraxial case*

Here, we report that the NNSE (3) has an *exact* non-paraxial soliton solution which is a three-parameter generalization of the two-parameter paraxial soliton

solution of the NSE. This can be found by seeking a solution of equation (3) of the form

$$u(\xi,\zeta)=\eta\,\mathrm{sech}\,[g(\xi+V\zeta)]\exp\,[\mathrm{i}\phi(\xi,\zeta)], \tag{10}$$

where g and ϕ are, initially, undetermined functions. Substitution of the *Ansatz* into equation (3), and solving the resulting equations for g and ϕ, subject to the constraint that the paraxial soliton is a particular solution, we find that

$$u(\xi,\zeta)=\eta\,\mathrm{sech}\left(\frac{\eta}{(1+2\kappa V^2)^{1/2}}(\xi+V\zeta)\right)\times\exp\left\{-\mathrm{i}V\xi\left(\frac{1+2\kappa\eta^2}{1+2\kappa V^2}\right)^{1/2}+\mathrm{i}\frac{1}{2\kappa}\zeta\left[-1+\left(\frac{1+2\kappa\eta^2}{1+2\kappa V^2}\right)^{1/2}\right]\right\}. \tag{11}$$

The usual amplitude parameter η and transverse velocity parameter V are now supplemented with the non-paraxial parameter κ. As required, the paraxial soliton is a particular case of the more general solution and is recovered when $\kappa\to 0$. The longitudinal phase factor (proportional to ζ) can also be expressed in a form which is a direct generalization of the corresponding dependence of the paraxial soliton solution. Writing this factor as

$$\exp\left(\frac{\mathrm{i}}{2}(\eta^2-V^2)\beta\zeta\right), \tag{12}$$

we find that the new parameter β, introduced in expression (12), satisfies the following quadratic equation:

$$\kappa(\eta^2-V^2)\beta^2+2\beta=\frac{2}{1+2\kappa V^2}. \tag{13}$$

From equation (11) it can be seen that the width of the non-paraxial soliton is given by

$$\xi_0=\frac{(1+2\kappa V^2)^{1/2}}{\eta}. \tag{14}$$

Thus the non-paraxial *soliton area* is proportional to $\xi_0\eta=(1+2\kappa V^2)^{1/2}$ and depends on both the transverse velocity and the size of the beam (through κ). In common with the paraxial case, this area is defined by the initial conditions and remains conserved during subsequent propagation.

For a beam with zero transverse velocity, solution (11) takes the form

$$u(\xi,\zeta)=\eta\,\mathrm{sech}\,(\eta\xi)\exp\left(\mathrm{i}\frac{\zeta}{2\kappa}(1+2\kappa\eta^2)^{1/2}\right)\exp\left(-\mathrm{i}\frac{\zeta}{2\kappa}\right). \tag{15}$$

Comparing with the corresponding paraxial solution (8), one sees that the non-paraxial term introduces a correction in the phase of the $V=0$ soliton. For a unit amplitude beam $\eta=1$, zero transverse velocity implies that $\xi_0=1$ in normalized units. Using equation (4), this width can be equated to $w_0/2^{1/2}$ in unscaled units. Hence, while κ only explicitly appears in the phase of this solution, the relationship between κ and w_0, given in equation (5), fixes the absolute width of the non-paraxial soliton in units of the optical wavelength.

Expressing the axial solution using the β parametrization results in a particu-

larly simple form of the $V = 0$ solution:

$$u(\xi, \zeta) = \eta \operatorname{sech}(\eta\xi) \exp\left(\frac{\mathrm{i}}{2}\eta^2\beta\zeta\right), \tag{16}$$

where

$$\kappa\eta^2\beta^2 + 2\beta = 2. \tag{17}$$

In the paraxial limit, $\kappa \to 0$ and $\beta \to 1$, while finiteness of the optical wavelength ($\kappa > 0$) leads to $\beta < 1$. Consequently, we find that the *soliton* (*phase*) *period* of the canonical non-paraxial solution is given by $\zeta_0 = 4\pi/\beta$ and that it is larger than the period of the paraxial solution.

Considering the experimental generation of non-paraxial solitons, it is likely that the precise spatial profile of the launched beam will not be as critical as it is in the paraxial case. This is because much shorter propagation distances will be required for the solitonic components to be resolved, since the low-amplitude (quasilinear) radiation modes will be on a shorter transverse length scale and will, consequently, be more rapidly dispersed across the transverse plane. It is also worthwhile to note that the simulation of ultranarrow non-paraxial beams is not as severely restricted, in requiring that the angles involved are vanishingly small, as computations in the paraxial regime. This is because the relaxation of the constraint $w_0/\lambda \to \infty$ can accommodate the consideration of large angles with only a moderate number of transverse data points.

4. Geometry of the non-paraxial soliton

4.1. *Dispersion relations*

To gain insight into the form of the non-paraxial soliton solution (11), in this section we consider the linear limits of the Helmholtz equation and the NNSE and the resulting dispersion relations. Neglecting the intensity-dependent part of the refractive index in equation (2), one has the linear 2D Helmholtz equation. Now assuming that $\tilde{E}$ varies as $\exp[\mathrm{i}(k_x x + k_z z)]$, one arrives at a dispersion relation for a plane wave propagating at an angle θ to the z direction:

$$k^2 = k_x^2 + k_z^2, \tag{18}$$

where $\theta = \tan^{-1}(k_x/k_z)$. This rather straightforward relationship is shown in figure 1. Assuming now that u varies as $\exp[\mathrm{i}(k_\xi \xi + k_\zeta \zeta)]$, the dispersion relation for the NNSE in the linear limit is found to be

$$\kappa k_\zeta^2 + k_\zeta + \frac{k_\xi^2}{2} = 0. \tag{19}$$

One can solve equation (19) to find the longitudinal component k_ζ of a propagating field envelope:

$$k_\zeta = -\frac{1}{2\kappa} \pm \frac{1}{2\kappa}(1 - 2\kappa k_\xi^2)^{1/2} = -\frac{1}{2\kappa} \pm \Delta k_\zeta. \tag{20}$$

Δk_ζ is defined in the above expression for notational convenience. In the derivation of either the paraxial or the non-paraxial normalized evolution equation, one eliminates the fast longitudinal phase variation of a forward propagating field by

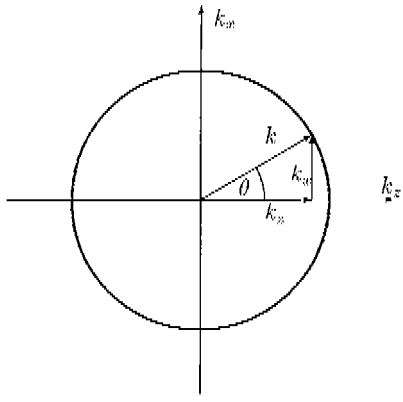

Figure 1. The dispersion relation for the linear 2D Helmholtz equation. k_z and k_x are the longitudinal and transverse components respectively of the spatial wave-vector.

introducing the factor $\exp(\mathrm{i}kz) = \exp(\mathrm{i}\zeta/2\kappa)$. However, it is clear that solutions corresponding to both forward- and backward-propagating waves still exist. These are given by the plus and minus signs respectively in equation (20). For visualization purposes, it is instructive to write the linearized NNSE dispersion relation as

$$\left(\frac{k_\zeta + \gamma}{\gamma}\right)^2 + \left(\frac{k_\xi}{\gamma^{1/2}}\right)^2 = 1, \tag{21}$$

where $\gamma = 1/2\kappa$. In figure 2 we plot this ellipse and highlight where the solution corresponds to forward and backward components. The geometry of the propagation vector, as defined in equation (20), is shown explicitly. The Helmholtz dispersion relation (18), which is a circle of radius k around the origin in (k_x, k_z) space, is now mapped onto an ellipse which is centred at $(k_\xi, k_\zeta) = (0, -1/2\kappa)$ and

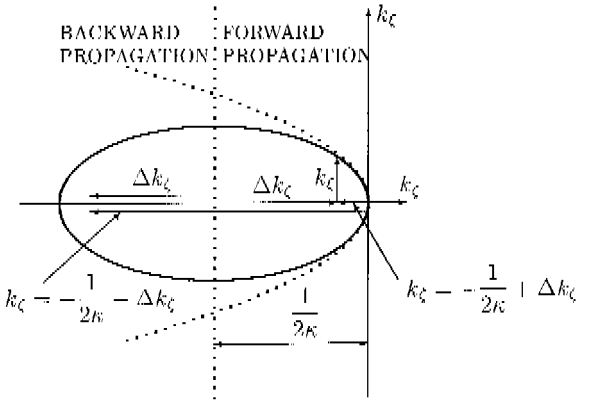

Figure 2. The dispersion relation for the linearized NNSE. Solutions reside on an ellipse (shown as a solid curve). This ellipse is partitioned into regions of forward and backward propagating waves (vertical dotted line). The construction of the longitudinal components of a pair of forward and backward solutions, $k_\zeta = -1/2\kappa \pm \Delta k_\zeta$, is shown. Also plotted is the paraxial dispersion relation which, in these coordinates, appears as a parabola (shown as a dotted curve).

which has half-axis lengths of $1/2\kappa$ and $1/(2\kappa)^{1/2}$. A plane wave propagating close to the forward z direction can be seen to have a small negative value of k_ζ, whereas a backward-propagating wave has a value of k_ζ which has an absolute value greater than $1/2\kappa$. The physical origin of the final (longitudinal) phase factor, which appears in equation (15) for example, is attributed to the normalization, $\exp(\mathrm{i}kz)$, to a forward-propagating field which is parallel to the ζ axis.

Since the boundary between forward and backward propagation can be given as $k_\xi^2 = 1/2\kappa$, any numerical algorithm which determines the forward beam will be required to filter out any spectral components which have $k_\xi^2 > 1/2\kappa$ [14]. More specifically, these high-spatial-frequency components experience a level of loss which is given by $\exp(\mathrm{i}k_\zeta\zeta) = \exp[-(2\kappa k_\xi^2 - 1)^{1/2}\zeta/2\kappa]\exp(-\mathrm{i}\zeta/2\kappa)$. The normalized paraxial dispersion relation is $k_\zeta = -\frac{1}{2}k_\xi^2$ and can be used to represent a plane wave which is propagating very close to the ζ axis. This relation, which is a parabolic approximation of equation (19) in the vicinity of the point $(k_\xi, k_\zeta) = (0,0)$, is plotted (as a dotted curve) in figure 2. However, when a beam becomes sufficiently non-paraxial, the paraxial wave equation will erroneously treat arbitrary values of k_ξ as forward-propagating components. In particular, it can be seen that this paraxial relation supports solutions for which the unphysical condition $k_x^2 > k^2$ is satisfied.

4.2. *Transverse velocity and soliton width*

In this section we show that the paraxial NSE fails to describe accurately soliton beam profiles which are not, almost exactly, parallel to the ζ axis. Furthermore, we find that this particular shortcoming does not appear in the corresponding non-paraxial representation. The intensity profile of a soliton beam, propagating in the ζ direction, plots out a stripe of light which has a width of $1/\eta$. However, when this beam propagates at a finite angle to the ζ direction, the cross-sections of the resulting intensity pattern (at each ζ) should give a width larger than $1/\eta$. The degree of this widening can be deduced from the geometrical considerations presented in figure 3. This expected enlargement of the $V \neq 0$ soliton width does not appear in the solution of the paraxial NSE.

Such broadening is, however, built into the non-paraxial solution through the dependence of the width of the soliton on the transverse velocity parameter V. We find that one can introduce, or remove, the velocity dependence of the soliton width by implementing an appropriate rotational transformation in the transverse plane. To do this, the non-paraxial solution and the NNSE are firstly written in terms of unscaled coordinates. Then, one implements a rotational transformation of magnitude θ (as shown in figure 3). Finally, one transforms the solution and the equation back to normalized units. These three operations can be combined into a single transformation of the normalized (NNSE) units:

$$\begin{bmatrix} \xi \\ \zeta \end{bmatrix} = \begin{bmatrix} \cos\theta & \dfrac{1}{(2\kappa)^{1/2}}\sin\theta \\ -(2\kappa)^{1/2}\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \xi' \\ \zeta' \end{bmatrix}. \tag{22}$$

For the NNSE under this transformation of units, $(\xi', \zeta') \rightarrow (\xi, \zeta)$, one finds that

$$\kappa\frac{\partial^2 u}{\partial\zeta^2} + \mathrm{i}\cos\theta\left(\frac{\partial u}{\partial\zeta} + V\frac{\partial u}{\partial\xi}\right) + \frac{1}{2}\frac{\partial^2 u}{\partial\xi^2} + |u|^2 u = 0. \tag{23}$$

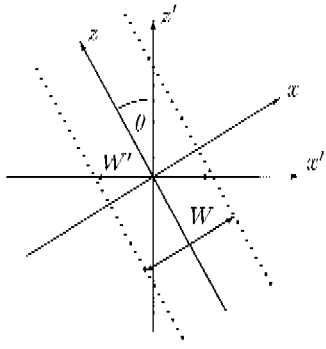

Figure 3. Schematic diagram of a propagating soliton beam, of width W, which makes an angle of θ to the (unscaled) coordinates of the evolution equation (x', z'). The projected width in the original coordinate system should be $W' = W/\cos\theta$.

From the properties of the solution transformation and the transformed NNSE, one can define the propagation angle θ in terms of the transverse velocity and non-paraxial parameters:

$$\tan\theta = (2\kappa)^{1/2} V. \tag{24}$$

Equivalently, the apparent soliton beam broadening factor, due to the rotation of the coordinate frame, is found to be

$$(1 + 2\kappa V^2)^{1/2} = \frac{1}{\cos\theta}. \tag{25}$$

Thus, for a finite value of κ, $\theta \to \pi/2$ requires $V \to \infty$, as expected for a beam propagating at a right angle to the longitudinal axis. For an intermediate angle of $\theta = \pi/4$, one requires that $2\kappa V^2 = 1$ and, in the cases of $\kappa = 10^{-3}$, 10^{-4} and 10^{-5}, one finds that the parameter V assumes values of around 22, 71 and 224 respectively. In addition to establishing the physical relationship between V, θ and κ, equations (24) and (25) permit θ to be eliminated from both equation (22) and equation (23), whereby the inverse transformation $(\xi, \zeta) \to (\xi', \zeta')$ and the correspondingly transformed NNSE are found by simply allowing $V \to -V$.

4.3. *Approximate solutions*

Since for most situations $\kappa \ll 1$, there are likely to be approximate expressions for the non-paraxial soliton solution which involve only the leading-order corrections to paraxial theory. For definiteness, we examine the case where η is of the order of unity and κ is orders of magnitude less than this. With these assumptions, the non-paraxial soliton takes the following form:

$$u(\xi, \zeta) \approx \eta \operatorname{sech}\left[\eta(1 - \kappa V^2)(\xi + V\zeta)\right] \times \exp\left[-\mathrm{i}V\xi[1 + \kappa(\eta^2 - V^2)] + \frac{\mathrm{i}}{2}(\eta^2 - V^2)\zeta\left(1 - \frac{\kappa}{2}(\eta^2 + 3V^2)\right)\right]. \tag{26}$$

Each argument of the approximate solution can be seen to be composed of two parts: the corresponding argument of the paraxial soliton (7), and a factor which is a correction to account for a small degree of non-paraxiality. The assumption for the validity of the above solution is that κ is sufficiently small that $(\kappa V^2)^2 \ll 1$ for

$\eta \approx 1$. $V \approx 1$ is clearly a particular range of validity but this is not a necessary requirement. For example, when $\kappa \approx 10^{-3}$, then V may be around 10 and, when $\kappa \approx 10^{-4}$ or 10^{-5}, then V may be around 30 or 100 respectively. Thus, $V \gg 1$ can be acceptable for this approximate solution. In this case, the non-paraxial soliton takes the following, even simpler, form:

$$u(\xi, \zeta) \approx \eta \operatorname{sech}\left[\eta(1 - \kappa V^2)(\xi + V\zeta)\right] \exp\left[-\mathrm{i}V\xi(1 - \kappa V^2) - \frac{\mathrm{i}}{2} V^2 \zeta (1 - \tfrac{3}{2}\kappa V^2)\right]. \tag{27}$$

We now consider the experimental or computational input requirements to generate a non-paraxial soliton. The general case, when no constraints are placed on η and V, is

$$u(\xi, 0) = \eta \operatorname{sech}\left(\frac{\eta\xi}{(1 + 2\kappa V^2)^{1/2}}\right) \exp\left[-\mathrm{i}V\xi\left(\frac{1 + 2\kappa\eta^2}{1 + 2\kappa V^2}\right)^{1/2}\right]. \tag{28}$$

Generally, $\kappa \ll 1$ and, when $\eta \approx 1$ and $(\kappa V^2)^2 \ll 1$, the above input condition reduces to

$$u(\xi, 0) \approx \eta \operatorname{sech}\left[\eta(1 - \kappa V^2)\xi\right] \exp\left\{-\mathrm{i}V\xi[1 + \kappa(\eta^2 - V^2)]\right\}. \tag{29}$$

Again, this approximation encompasses parameter regimes in which $V^2 \gg 1$ is permitted. In this case, the launched beam can be written as

$$u(\xi, 0) \approx \eta \operatorname{sech}\left[\eta(1 - \kappa V^2)\xi\right] \exp\left[-\mathrm{i}V\xi(1 - \kappa V^2)\right]. \tag{30}$$

Here, the modifications due to non-paraxiality can be parametrized by the single quantity κV^2 which, from equation (24), can be associated directly with the angular variable θ.

5. Conclusions

In this paper, we have considered the consequences of the miniaturization of (solitonic) nonlinear optical devices through the incorporation of non-paraxial effects in the nonlinear evolution equation. A NNSE was derived and an *exact* solution for the fundamental non-paraxial soliton was presented. A detailed comparison with the paraxial soliton has been made. In particular, the breakdown of the slowly varying envelope approximation has been shown to lead to modifications of the soliton width, the soliton area and the soliton (phase) period. The underlying physical and mathematical geometry of the non-paraxial soliton has been explored through considerations of dispersion relations, rotational transformations in the transverse plane, approximate solutions and the physical parameters involved (such as the width of the beam in terms of the number of optical wavelengths).

In a forthcoming paper we shall present a complete and self-contained account of a substantial amount of numerical work that we have under way in this field. There, we shall detail our non-paraxial beam propagation method, which we use to solve the NNSE numerically, and present results dealing with the formation, propagation and interaction of non-paraxial solitons. Further considerations, such as the fission of higher-order paraxial solitons, will also be presented.

Acknowledgments

We thank Dr S. Chávez-Cerda for useful discussions. This research was supported in part by UK Engineering and Physical Sciences Research Council grant No. GR/K 54748. One of us (P. C.-P.) wishes to acknowledge support from the Spanish Ministry of Education and Culture and Fellowship No. EX95 35457819.

References

[1] Zaharov, V., and Shabat, A., 1972, *Soviet Phys. JETP*, **34,** 62.
[2] Maneuf, S., and Reynaud, F., 1988, *Optics Commun.*, **66,** 325.
[3] Maneuf, S., Desailly, R., and Froehly, C., 1988, *Optics Commun.*, **65,** 193.
[4] Aitchison, F., Weiner, A., Silberberg, Y., Oliver, M., Jackel, J., Leaird, D. Vogel, E., and Smith, P., 1990, *Optics Lett.*, **15,** 471.
[5] Aitchison, F., Weiner, A., Silberberg, Y., Oliver, M., Jackel, J., Leaird, D., Vogel, E., and Smith, P., 1991, *J. opt. Soc. Am.* B, **8,** 1290.
[6] Aitchison, F., Weiner, A., Silberberg, Y., Oliver, M., Jackel, J., Leaird, D., Vogel, E., and Smith, P., 1991, *Optics Lett.*, **16,** 15.
[7] Shalaby, M., Reynaud, F., and Barthelemy, A., 1988, *Optics Commun.*, **65,** 193.
[8] McDonald, G. S., and Firth, W. J., 1990, *J. mod. Optics*, **37,** 613.
[9] McDonald, G. S., and Firth, W. J., 1990, *J. opt. Soc. Am.* B, **7,** 1328.
[10] McDonald, G. S., and Firth, W. J., 1993, *J. opt. Soc. Am.* B, **10,** 1081.
[11] Goorjian, P. M., Taflove, A., Joseph, R. M., and Hagness, S. C., 1992, *IEEE Jl quant. Electron.*, **28,** 2416.
[12] Hile, C. V., and Kath, W. L., 1996, *J. opt. Soc. Am.* B, **13,** 1135.
[13] Satsuma, J., and Yajima, N., 1974, *Prog. Theor. Phys., Osaha, Suppl.*, **55,** 284.
[14] Feit, M., and Fleck, J., 1988, *J. opt. Soc. Am.* B, **5,** 75.