Self-trapped optical beams: Spatial solitons

ANDREY A SUKHORUKOV and YURI S KIVSHAR

Nonlinear Physics Group, Research School for Physical Sciences and Engineering, Australian National University, Canberra ACT 0200, Australia

Abstract. We present a brief overview of the basic concepts of the theory of *spatial optical solitons*, including the soliton stability in non-Kerr media, the instability-induced soliton dynamics, and collision of solitary waves in nonintegrable nonlinear models.

1. Introduction

Recent years have shown increased interest in the study of self-guided (or self-trapped) optical beams that propagate in slab waveguides or bulk nonlinear media without supporting waveguide structures [1]. Such optical beams are commonly referred to by physicists as *spatial optical solitons*. Simple physics explains the existence of spatial solitons in a generalized self-focusing nonlinear medium. First, we recall the physics of optical waveguides. Optical beams have an innate tendency to spread (diffract) as they propagate in a homogeneous unbounded medium. However, the beam's *diffraction* can be compensated for by beam *refraction* if the material refractive index is increased in the region of the beam. An optical waveguide is an important means to provide a balance between diffraction and refraction if the medium is uniform in the direction of propagation. The corresponding propagation of the light is confined in the transverse direction of the waveguide, and it is described by the so-called linear *guided modes*, spatially localized eigenmodes of the electric field in the waveguide.

As was discovered long time ago [2], a similar effect, i.e. suppression of diffraction through a local change of the refractive index, can be produced solely by nonlinearity (figure 1). As has already been established in many experiments, some materials can display considerable optical nonlinearities when their properties are modified by the light propagation. In particular, if the nonlinearity leads to a change of the refractive index of the medium in such a way that it generates an effective positive lens to the beam (figure 1), the beam can become *self-trapped*, and it propagates unchanged without any external waveguiding structure [2]. Such stationary self-guided beams are called *spatial optical solitons*, they exist with profiles of certain form allowing a local compensation of the beam diffraction by the nonlinearity-induced change in the material refractive index.
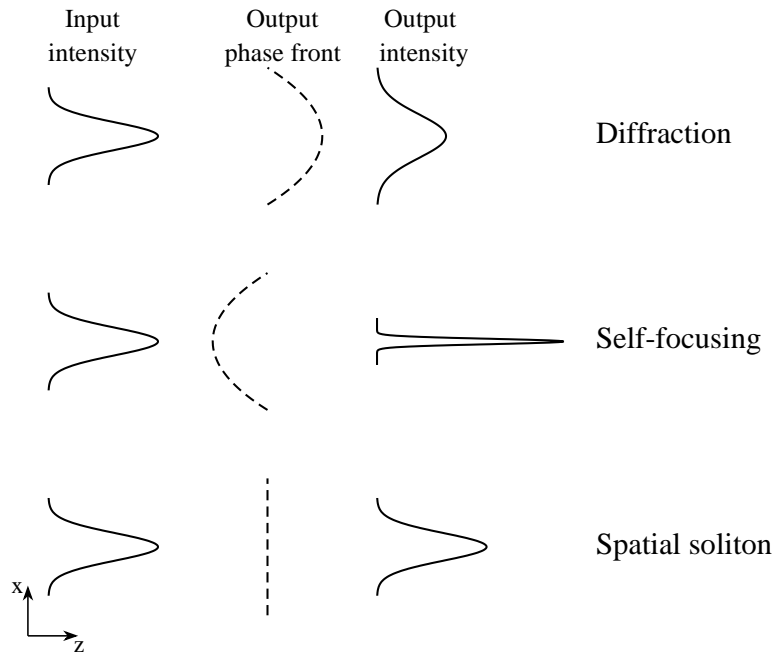
Figure 1. Qualitative picture of the beam self-trapping effect leading to the formation of a spatial optical soliton.

Until recently, the theory of spatial optical solitons has been based on the nonlinear Schrödinger (NLS) equation with a cubic nonlinearity, which is *exactly integrable* by means of the inverse scattering technique (IST) [3]. Generally speaking, integrability means that any localized input beam will be decomposed into *stable solitary waves* (or *solitons*) and *radiation*, and also that interaction of solitons is elastic. From the physical point of view, the integrable NLS equation describes (1+1)-dimensional (i.e. one transverse and one longitudinal dimensions) beams in a Kerr nonlinear medium in the framework of the so-called paraxial approximation. The cubic NLS equation is known to be a good model for *temporal optical solitons* propagating large distances along *existing* waveguides, optical fibers. In application to *spatial optical solitons*, the cubic NLS equation is not an adequate model. First, for spatial optical solitons much higher input powers are required to compensate for diffraction, meaning that the refractive index experiences large deviations from a linear (Kerr) dependence. Second, as was recognized long time ago [4], radially symmetric stationary localized solutions of the (2 + 1)-dimensional NLS equation are unstable and may collapse [5]. To deal with the realistic optical models, saturation had been suggested as a way to stabilize such a catastrophic self-focusing and produce stable solitary waves of higher dimensions. Accounting for this effect immediately leads to *nonintegrable models of generalized nonlinearities*, not possessing the properties of integrability and not allowing elastic soliton collisions. Another mechanism of non-Kerr nonlinearities and enhanced nonlinear properties of optical materials is a resonant, phase-matched interaction between the modes of different frequencies. In this latter case, *multi-component solitary waves* are created [6], and the mutual beam coupling can modify drastically the properties

of single beams, as it occurs in the case of the so-called *quadratic solitons of cascaded nonlinearities.*

In spite of the fact that, generally speaking, there exist no universal analytical tools for analyzing solitary waves and their interactions in nonintegrable models, recent advances of the theory suggest that many of the properties of optical solitons in non-Kerr media are similar, and they can be approached with the help of rather general physical concepts. Also, from this perspective we understand that *there is no simple mapping* between temporal and spatial optical solitons. Spatial solitons are a much richer and more complicated phenomenon, and this has already been demonstrated by a number of elegant experiments in this field.

In particular, it has been recently demonstrated experimentally, that self-guided beams can be observed in materials with strong photorefractive and photovoltaic effect [7], in vapours with a strong saturation of the intensity-dependent refractive index [8], and also as a result of the mutual self-focusing due to the phase-matched three-wave mixing in quadratic (or $\chi^{(2)}$) nonlinear crystals [9]. In all these cases, propagation of self-guided waves is observed in *non-Kerr materials* which are described by the models more general than the cubic NLS equation.

Because the phenomenon of the long-distance propagation of temporal optical solitons in optical fibers is known to a much broader community of researchers in optics and non-linear physics, first we emphasize the difference between spatial and temporal solitons. Indeed, stationary beam propagation in planar waveguides has been considered somewhat similar to the pulse propagation in fibers. This is a direct manifestation of the so-called *spatio-temporal analogy* in wave propagation [10], meaning that the propagation coordinate z is treated as the evolution variable and the spatial beam profile along the transverse direction, for the case of waveguides, is similar to the temporal pulse profile, for the case of fibers. This analogy has been employed for many years, and it is based on a simple notion that both beam evolution and pulse propagation can be described by the cubic NLS equation. However, contrary to this widely accepted opinion, we point out *a crucial difference between these two phenomena.* Indeed, in the case of the nonstationary pulse propagation in fibers, the operation wavelength is usually selected near the zero of the group-velocity dispersion. This means that the absolute value of the fiber dispersion is small enough to be compensated by a weak nonlinearity such as that produced by the (very weak) Kerr effect in optical fibers which leads to a relative nonlinearity-induced change in the refractive index of the order of 10^{-10}. Therefore, nonlinearity in such systems is *always weak* and it should be well modeled by the same form of the cubic NLS equation, which is known to be integrable by means of the IST technique. However, for very short (fs) pulses the cubic NLS equation describing the long-distance propagation of pulses should be corrected to include some additional effects such as higher-order dispersion, Raman scattering, etc. [11]. All these corrections can be taken into account by a perturbation theory [12]. Thus, in fibers nonlinear effects are weak and they become important only when dispersion is small (near the zero-dispersion point) affecting the pulse propagation over large distances (of order of hundred of meters or even kilometers).

In contrary to pulse propagation in optical fibers, the physics underlying stationary beam propagation in planar waveguides and bulk media is different. In this case the nonlinear change in the refractive index should compensate for the beam spreading caused by diffraction *which is not a small effect.* That is why to observe spatial solitons, much larger nonlinearities are usually required, and very often such nonlinearities are not of the Kerr type (e.g.

they saturate at higher intensities). This leads to the models of generalized nonlinearities with the properties of solitary waves different from those described by the integrable cubic NLS equation. For example, unlike the solitons of the cubic NLS equation, solitary waves of generalized nonlinearities may be unstable, they also show some interesting features, such as fusion due to collision, inelastic interactions and spiraling in a bulk, wobbling, amplitude oscillation, etc. Propagation distances usually involved in the phenomenon of beam self-focusing and spatial soliton propagation are of the order of millimeters or centimeters. As a result, the physics of spatial solitary waves is rich, and it should be understood in the framework of the theory of nonintegrable nonlinear models.

2. Basic equations

To describe optical spatial solitons in the framework of the *simplest scalar model of nonresonant nonlinearities*, we consider the propagation of a monochromatic scalar electric field E in a bulk optical medium with an intensity-dependent refractive index, $n = n_0 + n_{nl}(I)$, where n_0 is the linear refractive index, and $n_{nl}(I)$ describes the variation in the index due to the field with the intensity $I = |E|^2$. The function $n_{nl}(I)$ is assumed to be dependent on the light intensity only, and it may be introduced phenomenologically.

Solutions of the governing Maxwell's equation can be presented in the form

$$E(\mathbf{R}_\perp, Z; t) = \mathscr{E}(\mathbf{R}_\perp, Z)e^{ik_0 Z - i\omega t} + \text{c.c.}, \tag{1}$$

where c.c. denotes complex conjugate, ω is the source frequency, and $k_0 = 2\pi/\lambda$ is the linear wave number, or the plane-wave propagation constant for the uniform background medium, $\lambda = 2\pi c/(\omega n_0)$ is the linear wavelength, with c being the free-space speed of light. Usually, the spatial solitons are discussed for two geometries. For the beam propagation in a bulk, we assume a (2+1)-dimensional model, so that the Z-axis is parallel to the direction of propagation, and the X- and Y-axis are two transverse directions. For the beam propagation in a planar waveguide, the effective field is found by averaging the Maxwell's equations over the transverse structure defined by the waveguide confinement, and therefore the model becomes (1+1)-dimensional.

The function $\mathscr{E}(\mathbf{R}_\perp, Z)$ describes the wave envelope which in the absence of nonlinear and diffraction effects $\mathscr{E}$ would be a constant. If we substitute eq. (1) into the two-dimensional, scalar wave equation, we obtain the generalized nonlinear parabolic equation,

$$2ik_0\frac{\partial \mathscr{E}}{\partial Z} + \left(\frac{\partial^2 \mathscr{E}}{\partial X^2} + \frac{\partial^2 \mathscr{E}}{\partial Y^2}\right) + 2k_0^2 n_0^{-1} n_{nl}(I)\mathscr{E} = 0. \tag{2}$$

In dimensionless variables, eq. (2) becomes the well-known *generalized NLS equation*, where local nonlinearity is introduced by the function $n_{nl}(I)$.

For the case of the Kerr (or cubic) nonlinearity we have $n_{nl}(I) = n_2 I$, n_2 being the coefficient of the Kerr effect of an optical material. Now, introducing the dimensionless variables, i.e., measuring the field amplitude in the units of $\sqrt{2n_0/|n_2|}$, and the transverse coordinates and propagation distance in the units of $(2k_0)^{-1}$, we obtain the (2+1)-dimensional NLS equation in the standard form,

$$i\frac{\partial \Psi}{\partial z} + \left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2}\right) \pm |\Psi|^2 \Psi = 0, \tag{3}$$

where the complex Ψ stands for a dimensionless envelope, and the sign ($\pm$) is defined by the type of nonlinearity, *self-defocusing* ('minus', for $n_2 < 0$) or *self-focusing* ('plus', for $n_2 > 0$).

For propagation in a slab waveguide, the field structure in one of the directions, say Y, is defined by the linear guided mode of the waveguide. Then, the solution of the governing Maxwell's equation has the structure

$$E(\mathbf{R}_\perp, Z; t) = \mathscr{E}(X, Z) A_n(Y) e^{ik_n^{(0)}z - i\omega t} + \ldots, \tag{4}$$

where the function $A_n(Y)$ describes the corresponding fundamental mode of the slab waveguide, and $k_n^{(0)}$ is the corresponding linear propagation constant. Similarly, substituting this ansatz into Maxwell's equations and averaging over Y, we come again to the renormalized equation of the form (3) with the Y-derivative *omitted*, which in the dimensionless form becomes the standard cubic NLS equation

$$i\frac{\partial \Psi}{\partial z} + \frac{\partial^2 \Psi}{\partial x^2} \pm |\Psi|^2 \Psi = 0. \tag{5}$$

Equation (5) coincides formally with the equation for the pulse propagation in dispersive nonlinear optical fibers, and it is known to be integrable by means of the inverse scattering transform (IST) technique [3].

For the case of nonlinearities more general then the cubic one, we should consider the generalized NLS equation,

$$i\frac{\partial \Psi}{\partial z} + \frac{\partial^2 \Psi}{\partial x^2} + \mathscr{F}(|\Psi|^2)\Psi = 0, \tag{6}$$

where the function $\mathscr{F}(I)$ describes a nonlinearity-induced change of the refractive index, usually $\mathscr{F}(I) \propto I$ for small I.

The generalized NLS equation (2) [or eq. (6)] has been considered in many papers for analyzing the beam self-focusing and properties of spatial bright and dark solitons. All types of non-Kerr nonlinearities in optics can be divided, generally speaking, into *three main classes*: (i) *competing nonlinearities*, e.g., focusing (defocusing) cubic and defocusing (focusing) quintic nonlinearity, (ii) *saturable nonlinearities*, and (iii) *transiting nonlinearities*. Many references can be found in the recent review paper [13].

Usually, the nonlinear refractive index of an optical material deviates from the linear (Kerr) dependence for larger light intensities. Nonideality of the nonlinear optical response is known for semiconductor (e.g., AlGaAs, CdS, $CdS_{1-x}Se_x$) waveguides and semiconductor-doped glasses. In the case of small intensities, this effect can be modeled by *competing*, cubic-quintic nonlinearities, $n_{nl}(I) = n_2 I + n_3 I^2$. This model describes a competition between self-focusing ($n_2 > 0$), at smaller intensities, and self-defocusing ($n_3 < 0$), at larger intensities. Similar models are usually employed to describe the stabilization of wave collapse in the (2+1)-dimensional NLS equation.

Models with *saturable nonlinearities* are the most typical ones in nonlinear optics. The effective generalized NLS equation with saturable nonlinearity is also the basic model to describe the recently discovered (1+1)-dimensional photovoltaic spatial solitons in photovoltaic-photorefractive materials such as $LiNbO_3$. Unlike the phenomenological models usually used to describe saturation of nonlinearity, in the case of photovoltaic solitons this model can be justified rigorously.

There exist several different models for saturating nonlinearities. In particular, the phenomenological model $n_{nl}(I) = n_\infty[1 - (1 + I/I_{sat})^{-1}]$ is used more frequently, and it corresponds to the well-known expression derived from the two-level model.

Finally, *bistable solitons* introduced by Kaplan [14] usually require a special type of the intensity-dependent refractive index which changes from one type to another one, e.g., it varies from one kind of the Kerr nonlinearity, for small intensities, to another kind with different value of n_2, for larger intensities. Unfortunately, examples of nonlinear optical materials with such dependencies are not yet known, but the bistable solitons possess attractive properties useful for their possible future applications in all-optical logic and switching devices.

At last, we would like to mention the model of logarithmic nonlinearity, $n^2(I) = n_0^2 + \varepsilon \ln(I/I_0)$, that allows close-form exact expressions not only for stationary Gaussian beams (or *Gaussons*, as they were introduced in [15]), but also for periodic and quasi-periodic regimes of the beam evolution [16]. The main features of this model are the following: (i) the stationary solutions do not depend on the maximum intensity (quasi-linearization) and (ii) radiation from the periodic solitons is absent (the linearized problem has purely discrete spectrum). Such exotic properties persist in any dimension.

3. Stability of spatial solitons

Spatial optical solitons are of both fundamental and technological importance if they are stable under propagation. Existence of stationary solutions of the different models of non-Kerr nonlinearities does not guarantee their stability. Therefore, stability is *a key issue in the theory of spatial optical solitons.*

For temporal solitons in optical fibers, nonlinear effects are usually weak and the model based on the cubic NLS equation and its deformations is valid in most of the cases [11]. Therefore, being described by integrable or nearly integrable models, solitons are always stable, or their dynamics can be affected by (generally small) external perturbations which can be treated in the framework of the soliton perturbation theory [12].

As has been discussed above, much higher powers are usually required for spatial solitons in waveguides or a bulk, so that real optical materials demonstrate essentially non-Kerr change of the nonlinear refractive index with the increase of the light intensity. Generally speaking, the nonlinear refractive index always deviates from Kerr for larger input powers, e.g., it saturates. Therefore, models with a more general intensity-dependent refractive index are employed to analyze spatial solitons and, as we discuss below, solitary waves in such non-Kerr materials can become unstable. Importantly, in many cases the stability criteria for solitary waves can be formulated in a rather general form using the system invariants.

3.1 *Linear eigenvalue problem*

To discuss the stability properties of spatial optical solitons, we consider the nonintegrable generalized NLS equation that describes the $(1+1)$-dimensional beam self-focusing in a waveguide geometry,

$$i\frac{\partial \psi}{\partial z} + \frac{\partial^2 \psi}{\partial x^2} + \mathscr{F}(I)\psi = 0, \tag{7}$$

where $\psi(x,z)$ is the dimensionless complex envelope of the electric field, x is the transverse spatial coordinate, z is the propagation distance, $I = |\psi(x,z)|^2$ is the beam intensity, and the real function $\mathscr{F}(I)$ characterizes nonlinear properties of a dielectric medium, for which we assume with no lack of generality that $\mathscr{F}(0) = 0$. Stationary spatially localized solutions of the model (7) have the standard form, $\psi(x,z) = \Phi(x)e^{i\beta z}$, where β is the soliton propagation constant ($\beta > 0$) and the real function $\Phi(x)$ vanishes for $|x| \to \infty$. An important conserved quantity of the soliton in the model (7) is its *power* defined as

$$P(\beta) = \int_{-\infty}^{+\infty} |\psi(x,z)|^2 dx. \tag{8}$$

To find the linear stability conditions, we consider the evolution of a small-amplitude perturbation of the soliton presenting the solution in the form

$$\psi(x,z) = \left\{ \Phi(x) + [v(x) - w(x)]\,e^{i\gamma z} + [v^*(x) + w^*(x)]e^{-i\gamma^* t} \right\} e^{i\beta z}, \tag{9}$$

where the star stands for a complex conjugation, and obtain the linear eigenvalue problem for $v(x)$ and $w(x)$,

$$L_0 w = \gamma v, \quad L_1 v = \gamma w$$

$$L_j = -\frac{d^2}{dx^2} + \beta - U_j, \tag{10}$$

where $U_0 = \mathscr{F}(I)$ and $U_1 = \mathscr{F}(I) + 2I(\partial \mathscr{F}(I)/\partial I)$.

A stationary solution of the model (7) is stable if all the eigenmodes of the corresponding linear problem (10) do not have exponentially growing amplitudes. It can be demonstrated that the continuum part of the linear spectrum of the problem (10) consists of *two symmetric branches* corresponding to real eigenvalues, and the eigenstates in the gap $|\gamma| < \beta$ are responsible for the stability properties of localized solutions. Then, several cases of such discrete eigenvalues should be distinguished:

- *internal modes* with real eigenvalues describe periodic oscillations;
- *instability modes* correspond to purely imaginary eigenvalues;
- *oscillatory instabilities* can develop if the eigenvalues are complex.

In what follows, we present several approaches that allow to study analytically and numerically the structure of the discrete spectrum in order to determine the linear stability properties of solitary waves. We also analyze the nonlinear evolution of unstable solitons.

3.2 *Soliton internal modes and stability*

Since the soliton instabilities always occur in nonintegrable models, we wonder what kind of novel features can be found for solitary waves in nonintegrable models which might be responsible for the instabilities of their localized solutions. It is commonly believed

that solitary waves of nonintegrable nonlinear models differ from solitons of integrable models only in the character of the soliton interactions: unlike proper solitons, interaction of solitary waves is accompanied by radiation [12]. However, the soliton instabilities are associated with nontrivial effects of different nature, generic for localized waves in nearly integrable and nonintegrable models. In particular, a small perturbation to an integrable model may create *an internal mode* of a solitary wave [17] that is indeed responsible for the appearance of instabilities. This effect is beyond a regular perturbation theory, because solitons of integrable models do not possess internal modes. But in nonintegrable models such modes may introduce *qualitatively new features* into the system dynamics and, in particular, lead to instabilities.

To demonstrate that the internal modes are generic for nonintegrable models, we consider a weakly-perturbed cubic NLS equation with the nonlinear term,

$$\mathscr{F}(I) = I + \varepsilon f(I), \tag{11}$$

where $f(I)$ describes a deviation from the Kerr nonlinear response, and ε is a small parameter. Then, the stationary solution can be expressed asymptotically as $\Phi(x) = \Phi_0(x) + \varepsilon\Phi_1(x) + \mathrm{O}(\varepsilon^2)$, where $\Phi_0(x) = \sqrt{2\beta}\,\mathrm{sech}(\sqrt{\beta}x)$ is the soliton of the cubic NLS equation, and $\Phi_1(x)$ is a localized correction defined from eqs (7) and (11). Neglecting the second-order corrections, we find the results for the effective potentials of the linearized eigenvalue problem (10), $U_0 = \Phi_0^2 + \varepsilon\widetilde{U}_0$ and $U_1 = 3\Phi_0^2 + \varepsilon\widetilde{U}_1$, where $\widetilde{U}_0 = f(\Phi_0^2) + 2\Phi_0\Phi_1$ and $\widetilde{U}_1 = f(\Phi_0^2) + 2\Phi_0^2 f'(\Phi_0^2) + 6\Phi_0\Phi_1$, and the prime denotes differentiation with respect to the argument.

The linear eigenvalue problem (10) and (11) can be solved exactly at $\varepsilon = 0$ ([18]). Its discrete spectrum contains only the degenerated eigenvalue at the origin, $\gamma = 0$, corresponding to the so-called *soliton neutral mode* (see figure 2a). We may show that a small perturbation can lead to *the creation of an internal mode*, which bifurcates from the continuum spectrum band, as shown in figure 2b. For definiteness, we consider the upper branch of the spectrum and suppose that the cut-off frequencies $\gamma_\mathrm{m} = \pm\beta$ are not shifted by the perturbation. Then, the internal mode frequency can be presented in the form $\gamma = \beta - \varepsilon^2\kappa^2$, where κ is defined by the following result [17]

$$|\kappa| = \frac{1}{4}\mathrm{sign}(\varepsilon)\int_{-\infty}^{\infty}\left\{V(x,\beta)\widetilde{U}_1 V(x,\beta) + W(x,\beta)\widetilde{U}_0 W(x,\beta)\right\}dx. \tag{12}$$

Here $V(x,\gamma)$ and $W(x,\gamma)$ are the eigenfunctions of the cubic NLS equation calculated at the edge of the continuum spectrum, $V(x,\beta) = 1 - 2\mathrm{sech}^2(\sqrt{\beta}x)$ and $W(x,\beta) = 1$. A soliton internal mode appears if the right-hand side of eq. (12) is positive.

As an important example, we consider the case of the NLS equation (7) and (11) perturbed by a higher-order power-law nonlinear term,

$$f(I) = I^3. \tag{13}$$

The first-order correction to the soliton profile can be found in the form

$$\Phi_1(x) = -\frac{\sqrt{2}\beta^{5/2}\left[2\cosh(2\sqrt{\beta}x) + \cosh(4\sqrt{\beta}x)\right]}{3\cosh^5(\sqrt{\beta}x)}.$$
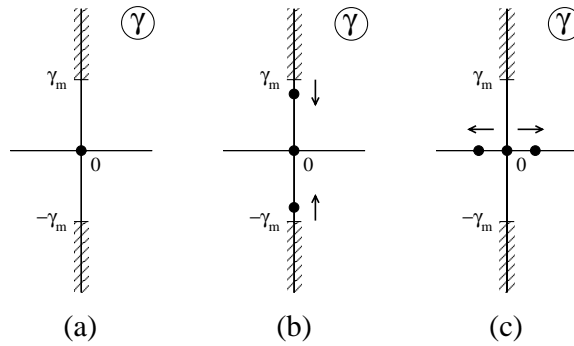
Figure 2. Schematic presentation of the origin of the bifurcation-induced soliton instabilities: (**a**) spectrum of the integrable cubic NLS model, (**b**) bifurcation of the soliton internal mode, (**c**) collision of the internal mode with the neutral mode resulting in the soliton instability.

With the help of eq. (12), it is easy to show that for $\varepsilon > 0$ a perturbed NLS soliton possesses an internal mode that can be found analytically near the continuum spectrum edge,

$$\gamma = \beta \left[1 - \left(\frac{64\varepsilon}{15} \right)^2 \beta^4 + O(\beta^6) \right]. \tag{14}$$

For high intensities, the additional nonlinear term (13) is no longer small, and the soliton solutions, together with the associated linear spectrum, should be calculated numerically. Power dependence $P(\beta)$ calculated with the help of eq. (8) for the soliton of the model (7), (11), and (13) is presented in figure 3a, and it corresponds to the discrete eigenvalue of the linearized problem (10) shown in figure 3b. First of all, we notice that the asymptotic theory provides accurate results for small intensity solitons, at $\beta < 0.1$ (shown by the dotted curves in the inserts). Second, the slope of the power dependence changes its sign at the point $\beta = \beta_{\mathrm{cr}}$ where the soliton internal mode vanishes colliding with the soliton neutral mode, as depicted in figures 2a–c. At that point, the soliton stability changes due to the appearance of unstable (*purely imaginary*) eigenvalues (dashed curve in figure 3b). In §4, we demonstrate rigorously a link between the soliton stability and the slope of the dependence $P(\beta)$.

For $\beta \approx \beta_{\mathrm{cr}}$, the instability-induced dynamics of an unstable soliton can be described by approximate equations for the soliton parameters derived by the multi-scale asymptotic technique (see §3.3.4) but, in general, we should perform numerical simulations in order to study the evolution of linearly unstable solitons. In figures 4a, b, we show two different types of the soliton evolution in our model. In the first case, a small perturbation that effectively *increases* the soliton power results in an unbounded growth of the soliton amplitude and subsequent beam collapse (see figure 4a). In the second case, a small *decrease* of the soliton power leads to a switching of a soliton of an unstable (dashed line, figure 3a) branch to a stable (solid line, figure 3a) one, as is shown in figure 4b. This latter scenario becomes possible because, in the model under consideration, all small-amplitude solitons are stable. However, if the small-amplitude solitons are unstable, the soliton beam does not converge to a stable state but, instead, it diffracts. Thus, in the NLS-type nonlinear models there exist *three distinct types* of the instability-induced solitons dynamics [19].
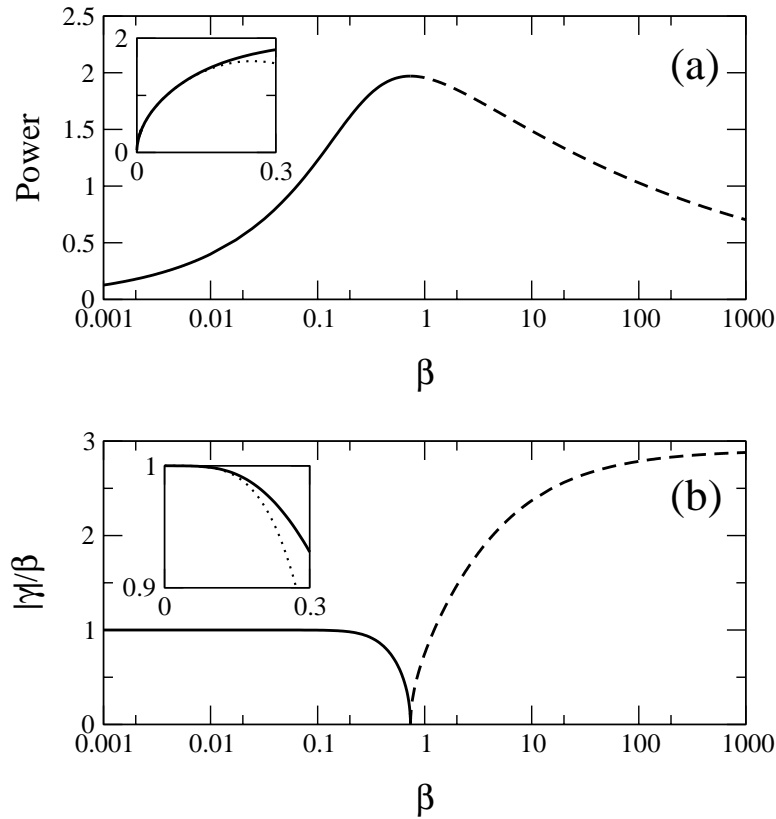
Figure 3. Soliton instability in the model defined by eqs (7), (11), (13) for $\varepsilon = 1$, presented through (**a**) the power dependence $P(\beta)$ and (**b**) the evolution of the discrete eigenvalue of the problem (3) that defines the soliton internal mode (solid) and an instability mode (dashed). Solitons for $\beta > \beta_{cr}$, i.e. for $dP/d\beta > 0$ (dashed curves in (a) and (b)) are *linearly unstable*. Dotted lines in the insets show the asymptotic dependences calculated analytically.

3.3 *Stability criterion for fundamental solitons*

Direct investigation of the eigenvalue problem (10) is a complicated task which, in general, does not have a complete analytical solution. However, for a class of 'fundamental' solitary waves (i.e. solitons with no nodes), the analysis can be greatly simplified. First, we reduce the system (10) to a single equation:

$$L_0 L_1 v = \gamma^2 v, \tag{15}$$

and the stability condition requires all eigenvalues γ^2 to be positive. It is obvious to show that $L_0 = L^+ L^-$, where $L^{\pm} = \pm d/dx + \Phi^{-1}(d\Phi/dx)$, and thus one can consider instead an auxiliary eigenvalue problem [20],

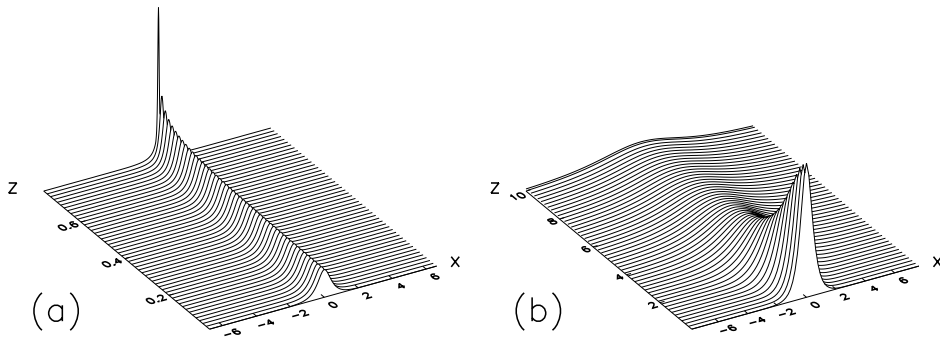$$L^- L_1 L^+ \tilde{v} = \gamma^2 \tilde{v}, \tag{16}$$

Figure 4. Evolution of an unstable NLS soliton in the model (7),(11),(13) for $\varepsilon = 1$, and $\beta = 2$, in the case of (**a**) increased power (collapse) and (**b**) decreased power (switching to a low-amplitude stable state). The soliton power was changed by 1%.

which reduces to eq. (15) after the substitution $v = L^+ \tilde{v}$. Since the operator $L^- L_1 L^+$ is Hermitian, all eigenvalues γ^2 of eqs (15) and (16) are real, and in this case oscillatory instabilities are not possible.

Operators L_j ($j = 0, 1$) are well-studied in the literature, in particular, as a characteristic example of the spectral theory of the second-order differential operators [21]. For our problem, we use two *general mathematical results* about the spectrum of the linear eigenvalue problem $L_j \varphi_n^{(j)} = \gamma_n^{(j)} \varphi_n^{(j)}$:

- the eigenvalues can be ordered as $\gamma_{n+1}^{(j)} > \gamma_n^{(j)}$ where $n \geq 0$ defines the number of zeros in the corresponding eigenfunction $\varphi_n^{(j)}$;
- for a 'deeper' potential well, $\widetilde{U}_j(x) \geq U_j(x)$, the corresponding set of the eigenvalues is shifted 'down', i.e., $\widetilde{\gamma}_n^{(j)} \leq \gamma_n^{(j)}$.

Let us first discuss the properties of the operator L_0, for which the soliton neutral mode is an eigenstate, i.e., $L_0 \Phi(x) = 0$. As we have assumed earlier, $\Phi(x) > 0$ is the ground state solution with no nodes and, therefore, $\gamma_n^{(0)} > \gamma_0^{(0)} = 0$ for $n > 0$. This means that the operator L_0 is positive definite on the subspace of the functions orthogonal to $\Phi(x)$, which allows to use several general theorems [22–26] in order to link the soliton stability properties to the number of negative eigenvalues of the operator L_1. Specifically, in a homogeneous medium the fundamental soliton stability depends on the slope of the function $P(\beta)$, according to the Vakhitov-Kolokolov stability criterion [22], i.e., a soliton is *stable* if $\partial P / \partial \beta > 0$, and it is *unstable*, otherwise.

To demonstrate the validity of the Vakhitov-Kolokolov criterion, we follow the standard procedure [22]. First, we note that, for the fundamental solitons with no nodes, both the direct, L_0, and inverse, L_0^{-1}, operators exist, and they are positively definite for any function orthogonal to $\Phi(x)$, which is not an eigenmode of eq. (15), and thus can be ignored. Thus, by applying the inverse operator L_0^{-1} to eq. (15), we obtain another linear problem with the same spectrum:

$$L_1 v = \gamma^2 L_0^{-1} v, \tag{17}$$

where $v(x)$ now satisfies the orthogonality condition:

$$\langle v | \Phi \rangle \equiv \int_{-\infty}^{+\infty} v^*(x) \Phi(x) dx = 0. \tag{18}$$

Second, we multiply both sides of eq. (17) by the function $v^*(x)$, integrate over x, and obtain the following result,

$$\gamma^2 = \frac{\langle v | L_1 v \rangle}{\langle v | L_0^{-1} v \rangle}. \tag{19}$$

Because the denominator in eq. (19) is positively definite for v satisfying eq. (18), the sign of this ratio depends only on the numerator. The instability will appear if there exists an eigenvalue $\gamma^2 < 0$ (so that γ is imaginary), and this can be possible only if

$$\min \langle v | L_1 v \rangle \ < 0, \tag{20}$$

where we normalize the function $v(x)$ to make the expression in eq. (20) finite, $\langle v | v \rangle = 1$.

In order to find the minimum in eq. (20) under the constraints given by eqs (18) and (20), we use the method of Lagrange multipliers, and look for a minimum of the following functional:

$$\mathscr{L} = \langle v | L_1 v \rangle - \nu \langle v | v \rangle - \mu \langle v | \Phi \rangle,$$

where real parameters ν and μ are unknown. With no lack of generality we assume that $\mu \geq 0$, as otherwise the sign of the function $v(x)$ can be inverted. The extrema point of the functional $\mathscr{L}$ can be then found from the condition $\delta \mathscr{L} / \delta v = 0$, where δ denotes the variational derivative. As a result, we obtain the following equation:

$$L_1 v = \nu v + \mu \Phi, \tag{21}$$

where the values of ν and μ should be chosen in such a way that conditions in eqs (18) and (20) are satisfied. Then, it immediately follows that $\langle v | L_1 v \rangle = \nu \langle v | v \rangle$ and, according to eq. (20), the stationary state is unstable if and only if there exists a solution with $\nu < 0$.

Operator L_1 has a full set of orthogonal eigenfunctions $\varphi_n^{(1)}$ [21], i.e., $\left\langle \varphi_n^{(1)} | \varphi_m^{(1)} \right\rangle = 0$ if $n \neq m$. The spectrum of L_1 consists of discrete ($\gamma_n^{(1)} < \beta$) and continuous ($\gamma_n^{(1)} \geq \beta$) parts, and we scale the eigenmodes norms $\left\langle \varphi_n^{(1)} | \varphi_n^{(1)} \right\rangle$ to unity or a delta-function, respectively. Then, we can decompose the function $v(x)$ in the following way:

$$v(x) = \sum_n D_n \varphi_n^{(1)}(x) + \int_{\beta}^{+\infty} D_n \varphi_n^{(1)}(x) d\gamma_n^{(1)}, \tag{22}$$

where the sum goes only over the eigenvalues of the discrete spectrum of L_1. Coefficients in eq. (22) can be found as $D_n = \left\langle \varphi_n^{(1)} | v \right\rangle$. Function $\Phi(x)$ can be then decomposed in a similar way, with the coefficients $C_n = \left\langle \varphi_n^{(1)} | \Phi \right\rangle$. Then, eq. (21) can be reduced to:

$$D_n = \begin{cases} \mu\, C_n / (\gamma_n^{(1)} - v), & \text{if } C_n \neq 0 \text{ and } \mu > 0, \\ 1, & \text{if } C_n = 0, \ \mu = 0, \text{ and } v = \gamma_n^{(1)}, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

In order to find the Lagrange multiplier v, we substitute eqs (22) and (23) into the orthogonality condition (18), and obtain the following equation for the parameter v:

$$Q(v) \equiv \langle v | \Phi \rangle = \sum_n C_n D_n^* + \int_\beta^{+\infty} C_n D_n^* d\gamma_n^{(1)} = 0. \tag{24}$$

As has been mentioned above, instability appears if there exists a root $v < 0$, and thus we should determine the sign of minimal v solving eq. (24). Because the lowest-order modes of the operators L_0 and L_1, $\Phi(x)$ and $v_0(x)$ respectively, do not contain zeros, the coefficient $C_0 \neq 0$. Then, from the structure of eq. (24) it follows that $Q(v < \gamma_0^{(1)}) > 0$, and thus the solutions are only possible for $v > \gamma_0^{(1)}$.

Since $\mathscr{F}(I)$ does not depend on x, a fundamental soliton has a symmetric profile with a single maximum, and $d\Phi/dx$ is the first-order neutral mode of the operator L_1, i.e., $\gamma_1^{(1)} = 0$. Therefore, the modes with $v = \gamma_m^{(1)}$ do not give rise to instability, and we search for the solutions with $\mu > 0$. We notice that the function $Q(v)$ is monotonic in the interval $(-\infty, +\infty)$ for $\gamma_0^{(1)} < v < \gamma_m^{(1)}$, where $m > 1$ corresponds to the smallest eigenvalue with $C_m \neq 0$. Then, because $\gamma_m^{(1)} > \gamma_1^{(1)} = 0$, the sign of the solution v is determined by the value of $Q(0)$. Indeed, if $Q(0) > 0$, the function $Q(v)$ vanishes at some $v < 0$ which indicates *instability*, and vice versa. From eqs (21) and (24), it follows that $Q(v = 0) = \langle L_1^{-1} \mu \Phi | \Phi \rangle$. To calculate this value, we differentiate the equality $L_0 \Phi = 0$ with respect to the propagation constant and obtain

$$L_1 \frac{\partial \Phi}{\partial \beta} = -\Phi, \tag{25}$$

that finally determines the sign, $\text{sign}[Q(0)] = -\text{sign}(dP/d\beta)$. These results provide a proof of the stability condition outlined above.

Linear stability discussed above should be compared with a more *general Lyapunov stability theorem* which states that in a conservative system a stable solution (in the Lyapunov sense) corresponds to an extrema point of an invariant such as the system Hamiltonian, provided it is bounded from below (or above). For the cubic NLS equation, this means that a soliton solution is a stationary point of the Hamiltonian H for a fixed power P, and it is found from the variational problem $\delta(H + \beta P) = 0$. In order to prove the Lyapunov stability, one needs to demonstrate that, for a class of localized solutions, the soliton solution realizes a minimum of the Hamiltonian when P is fixed. This fact can be shown rigorously for a homogeneous medium with cubic nonlinearity [23], i.e., for $\mathscr{F}(I,x) = I$, and it follows from the integral inequality,

$$H > H_s + (P^{1/2} - P_s^{1/2}), \tag{26}$$

where the subscript 's' defines the integral values calculated for the NLS soliton. Condition (26) demonstrates the soliton stability for both small and finite-amplitude perturbations, and a similar relation can be obtained for generalized nonlinearity, being consistent with the Vakhitov-Kolokolov criterion.

3.4 *Marginal stability point: Asymptotic analysis*

As was shown above, even a small perturbation of the cubic NLS equation may lead to novel features in the soliton dynamics associated with the soliton internal modes. As follows from the linear stability analysis, the solitons in a homogeneous medium are unstable when the slope of the power dependence is negative, i.e., for $dP/d\beta < 0$.

Near the marginal stability point $\beta = \beta_{cr}$, such that $(dP/d\beta)_{\beta=\beta_{rmcr}} = 0$, where the instability growth rate is small, we can derive a general analytical asymptotic model which describes not only linear instabilities but also the nonlinear long-term evolution of unstable solitons. Such an approach is based on a nontrivial modification of the soliton perturbation theory [12] near the marginal stability point $\beta = \beta_{cr}$.

The standard soliton perturbation theory [12] is usually applied to analyse the soliton dynamics under the action of external perturbations. Here we should deal with *a qualitatively different physical problem* when an unstable bright soliton evolves under the action of its 'own' perturbations. As a result of the development of the instability, the soliton propagation constant varies slowly along the propagation direction, i.e., $\beta = \beta(z)$. Near the instability threshold, when the derivative $dP/d\beta$ vanishes at $\beta = \beta_{cr}$, the instability growth rate is small, and we can assume that the instability-induced evolution of the perturbed soliton is slowly varying in z and the soliton evolves almost adiabatically (i.e., it remains self-similar). Therefore, we can develop *an asymptotic theory* representing the solution to the original model (7) in the form $\psi = \phi(x; \beta; Z) \exp[i\beta_0 z + i\varepsilon \int_0^Z \beta(Z')dZ']$, where $\beta = \beta_0 + \varepsilon^2 \Omega(Z)$, $Z = \varepsilon z$, and $\varepsilon \ll 1$. Here the constant value β_0 is chosen in the vicinity of the critical point β_{cr}. Then, using the asymptotic multi-scale expansion in the form $\phi = \Phi(x; \beta) + \varepsilon^3 \phi_3(x; \beta; Z) + O(\varepsilon^4)$, we obtain the following equation for the soliton propagation constant β (details can be found in [19,27]),

$$M(\beta_{cr})\frac{d^2\Omega}{dZ^2} + \frac{1}{\varepsilon^2}\frac{dP}{d\beta}\bigg|_{\beta=\beta_0}\Omega + \frac{1}{2}\frac{d^2P}{d\beta^2}\bigg|_{\beta=\beta_{cr}}\Omega^2 = 0, \tag{27}$$

where $P(\beta)$ and $M(\beta)$ are calculated through the stationary soliton solution, and

$$M(\beta) = \int_{-\infty}^{+\infty}\left[\frac{1}{\Phi(x;\beta)}\int_0^x \Phi(x';\beta)\frac{\partial\Phi(x';\beta)}{\partial\beta}dx'\right]^2 dx > 0.$$

A remarkable result which follows from this asymptotic analysis is the following. In the generalized NLS equation (7) the dynamics of solitary waves near the instability threshold can be described by a simple model (27) which is equivalent to the equation for motion of an effective *inertial and conservative* particle of the mass $M(\beta_{cr})$ with the coordinate Ω moving under the action of a potential force which is proportional to the difference $P_0 - P(\beta)$, where $P_0 = P(\beta_0)$.

The first two terms in eq. (27) give the result of the linear stability analysis, according to which the soliton is *linearly unstable* provided $dP/d\beta < 0$. Nonlinear term in eq. (27) allows to describe not only linear but also long-term (nonlinear) dynamics of an unstable soliton and, moreover, to identify qualitatively different scenarios of the instability-induced soliton dynamics near the marginal stability point (see details in [19,27]).

4. Interaction of spatial solitons

In the case of optical solitons generated in a waveguide geometry, i.e., those described by the (1+1)-dimensional NLS equation, the special case of the cubic nonlinearity is known to be integrable by means of the IST transform. This is in a sharp contrast with the (2+1)-dimensional models where integrability is very rare, and usually it is not associated with the optical models. One of the many properties of an integrable model is the existence of exact analytical solutions describing *elastic interaction of any number of solitons*, that is, when a complicated nonlinear interaction of spatially localized waves reduces to simple nonlinear superposition resulting in phase shifts. In the framework of the generalized NLS equation the soliton interaction is not elastic because the model is non-integrable. Non-integrability produces a number of interesting effects which are closely associated with soliton interactions in realistic physical models corresponding to those observed in experiment. Many of the effects produced by the absence of integrability can be already seen in a simple example of the cubic-quintic NLS equation (11). When the parameter ε is small, the inelastic effects associated with non-integrability of the model can be divided into *two classes*. The first class of the effects is characterized by the order of ε^2 produced by radiation (of the amplitude $\sim \varepsilon$) that appears as a result of collision between two solitons in the framework of eq. (11). Radiation is also responsible for a change of the soliton amplitude after the collision. The second class of effects is associated with the order of ε. An important example is the so-called *multi-soliton inelastic effects* observed in collision of three (or more) solitons as an energy exchange between the colliding solitary waves not involving radiation (in the order of ε) [28].

Similar inelastic effects of the first order of ε are produced by the excitation of the soliton internal modes [17], that can appear for $\varepsilon \neq 0$, being associated with discrete eigenvalues splitting off the soliton continuum spectrum (see above). No soliton internal modes are known to exist in integrable soliton-supporting models.

When the perturbation parameter ε is not small, the soliton interaction becomes more complicated, and it is usually accompanied by strong radiation and energy exchange. However, for two colliding solitons of equal amplitudes, the interaction crucially depends on the relative phase θ between the colliding solitons and their mutual coherence. These properties are very common for different types of solitons (including those of the integrable models), and they are explained by simple physics. Two mutually coherent solitons attract each other when they are in phase [$\theta = 0$, figure 5a] or repel when they are out-of-phase [$\theta = \pi$, figure 5c]. In the intermediate case, i.e., for $0 < \theta < \pi$ the soliton interaction is accompanied by energy exchange, as shown in figure 5b.

The basic attractive and repulsive interaction involves the following. For in-phase beams the intensity in the overlapping region between the beams increases. This, in turn, results in a local increase of the refractive index which effectively attracts both beams. Exactly the opposite situation arises when the solitons are out-of-phase. Then, the light intensity drops in the overlap region and so does the refractive index change. This results in the beams moving away from each other, which is interpreted as a repulsive force.

5. Conclusions

During the last 10 years, the field of spatial optical solitons has expanded dramatically being driven by a number of impressive experimental demonstrations of the soliton effects
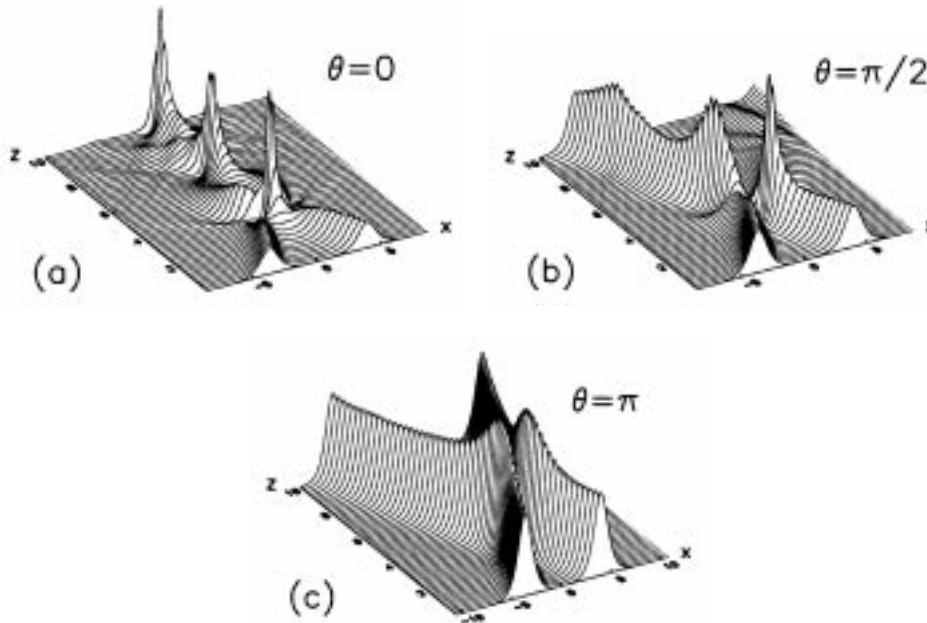
Figure 5. Result of the collision of two solitons with equal amplitudes in the nonintegrable model defined by eqs (7),(11) for $f(I) = I^2$ and $\varepsilon = 0.2$. The initial soliton relative phase θ is (a) 0, (b) $\pi/2$, and (c) π.

in different types of nonlinear bulk media, waveguide structures, and waveguide arrays. In this (largely tutorial) paper, we have presented only a small fraction of the fundamental problems associated with this field. Unlike temporal solitons in fiber optics, the spatial solitons are described by higher-dimensional nonintegrable nonlinear models, they may become unstable even in the Kerr medium, and their analysis requires the development of novel analytical techniques for studying of localized solutions of nonintegrable nonlinear equations.

Spatial optical solitons bring many novel concepts into the nonlinear physics, they can carry a topological charge analogous to 'spin', they can consist of several components, they can spiral, fuse, repel each other in a bulk medium, etc. We believe that novel nonlinear materials such as nonlinear photonic crystals may expand the possibility of realistic applications of spatial optical solitons towards creating a novel generation of nonlinear optical devices operating entirely with light.

Acknowledgements

indebted to all of them for helping us in getting a deeper insight into the field of spatial optical solitons.

References

[1] M Segev and G I Stegeman, *Phys. Today* **51**(8), 42 (1998)
    G I Stegeman and M Segev, *Science* **286**, 1518 (1999)
[2] R Y Chiao, E Garmire, and C H Townes, *Phys. Rev. Lett.* **13**, 479 (1964)
[3] V E Zakharov and A B Shabat, *Zh. Éksp. Teor. Fiz.* **61**, 118 (1971) [*Sov. Phys. JETP* **34**, 62 (1972)]
[4] P L Kelley, *Phys. Rev. Lett.* **15**, 1005 (1965)
[5] For a comprehensive overview of the wave collapse phenomenon, see J J Rasmussen and K Rypdal, *Phys. Scr.* **33**, 481 (1986); and also a more recent review
    L Bergé, *Phys. Rep.* **303**, 259 (1998)
[6] Yu S Kivshar, In *Advanced photonics with second-order optically nonlinear processes* edited by A D Boardman, L Pavlov and S Tanev (Kluwer, Dordretch, 1998) pp. 451-475
[7] M Segev, G C Valley, B Crosignani, P DiPorto and A Yariv, *Phys. Rev. Lett.* **73**, 3211 (1994)
    M Shih, P Leach, M Segev, M H Garrett, G Salamo and G C Valley, *Opt. Lett.* **21**, 324 (1996)
    M D Iturbe-Castillo, P A Marquez Aguilar, J J Sanchez-Mondragon, S Stepanov and V Vysloukh, *Appl. Phys. Lett.* **64**, 408 (1994)
[8] V Tikhonenko, J Christou and B Luther-Davies, *J. Opt. Soc. Am.* **B12**, 2046 (1995)
    V Tikhonenko, J Christou and B Luther-Davies, *Phys. Rev. Lett.* **76**, 2698 (1996)
[9] W E Torruellas, Z Wang, D J Hagan, E W VanStryland, G I Stegeman, L Torner and C R Menyuk, *Phys. Rev. Lett.* **74**, 5036 (1995)
[10] S A Akhmanov, A P Sukhorukov and R V Khokhlov, *Usp. Fiz. Nauk* **93**, 19 (1967) [*Sov. Phys. Usp.* **10**, 609 (1968)]
[11] A Hasegawa and Y Kodama, *Solitons in Optical Communications* (Oxford University Press, Oxford, 1995)
[12] For a comprehensive overview of the soliton perturbation theory, see Yu S Kivshar and B A Malomed, *Rev. Mod. Phys.* **61**, 763 (1989)
[13] Yu S Kivshar and B Luther-Davies, *Phys. Rep.* **298**, 81 (1998)
[14] A E Kaplan, *Phys. Rev. Lett.* **78**, 1291 (1985)
[15] I Bialynicki-Birula and J Mycielski, *Phys. Scr.* **20**, 539 (1979)
[16] A W Snyder and D J Mitchell, *Opt. Lett.* **22**, 16 (1997)
[17] Yu S Kivshar, D E Pelinovsky, T Cretegny and M Peyrard, *Phys. Rev. Lett.* **80**, 5032 (1998)
[18] D J Kaup, *Phys. Rev.* **A42**, 5689 (1990)
[19] D E Pelinovsky, A V Buryak and Yu S Kivshar, *Phys. Rev. Lett.* **75**, 591 (1995)
    D E Pelinovsky, V V Afanasjev and Yu S Kivshar, *Phys. Rev.* **E53**, 1940 (1996)
[20] M M Bogdan, A S Kovalev and I V Gerasimchuk, *Fiz. Nizk. Temp.* **23**, 197 (1997) [*Low Temp. Phys.* **23**, 145 (1997)]
[21] E C Titchmarsh, *Eigenfunction expansions associated with second-order differential equations* (Oxford University Press, London, 1958)
[22] N G Vakhitov and A A Kolokolov, *Radiophys. Quantum Electron.* **16**, 783 (1973)
    A A Kolokolov, *Radiophys. Quantum Electron.* **17**, 1016 (1976)
[23] E A Kuznetsov, A M Rubenchik and V E Zakharov, *Phys. Rep.* **142**, 103 (1986)
[24] C K R T Jones, *J. Diff. Eq.* **71**, 34 (1988); *Ergod. Theor. Dynam. Sys.* **8***, 119 (1988)
[25] M Grillakis, *Comm. Pure Appl. Math.* **41**, 747 (1988); **43**, 299 (1990)
[26] M Grillakis, J Shatah and W Strauss, *J. Funct. Anal.* **74**, 160 (1987); **94**, 308 (1990)