ELSEVIER

Physica D 135 (2000) 345–368

PHYSICA D

www.elsevier.com/locate/physd

Modeling light bullets with the two-dimensional sine–Gordon equation

J.X. Xin

Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA

Received 16 November 1998; received in revised form 22 February 1999; accepted 26 April 1999
Communicated by C.K.R.T. Jones

Abstract

Light bullets are spatially localized ultra-short optical pulses in more than one space dimensions. They contain only a few electromagnetic oscillations under their envelopes and propagate long distances without essentially changing shapes. Light bullets of femtosecond durations have been observed in recent numerical simulation of the full Maxwell systems. The sine–Gordon (SG) equation comes as an asymptotic reduction of the two level dissipationless Maxwell–Bloch system. We derive a new and complete nonlinear Schrödinger (NLS) equation in two space dimensions for the SG pulse envelopes so that it is globally well-posed and has all the relevant higher order terms to regularize the collapse of the standard critical NLS (CNLS). We perform a modulation analysis and found that SG pulse envelopes undergo focusing–defocusing cycles. Numerical results are in qualitative agreement with asymptotics and reveal the SG light bullets, similar to the Maxwell light bullets. We achieve the understanding that the light bullets are manifestations of the persistence and robustness of the complete NLS asymptotics. ©2000 Elsevier Science B.V. All rights reserved.

Keywords: Light bullets; Sine–Gordon equation; Maxwell–Bloch system

1. Introduction

Ultra short optical pulses of femtosecond duration are of tremendous technological and fundamental interest, and have been the subject of many recent studies in nonlinear optics [4,10,15,16,20,21], among others. Such short light pulses have potential applications in time-domain spectroscopy of dielectrics, semiconductors and transient chemical processes, probing high-intensity plasmas, imaging and medical infrared tomography [20], light propagation through atmosphere, and near material interfaces [4].

Conventional nonlinear optics usually operate with almost harmonic EM (electromagnetic) oscillations modulated by an envelope much longer than a single cycle of the oscillation so that there are on the order of 100–1000 oscillations under the envelope. These pulses are hence referred to as long pulses. The time honored approach for long pulses is the slowly varying envelope approximation and the derivation of the nonlinear Schrödinger (NLS) equation for the envelope, see e.g [27].

In clear contrast, short pulses typically have only a few EM oscillation cycles under their envelopes. So there is *a lack of separation of scales* between the envelope and the underlying EM oscillations, which seems to make

* *E-mail address:* xin@math.arizona.edu (J.X. Xin)

PII: S0167-2789(99)00128-1

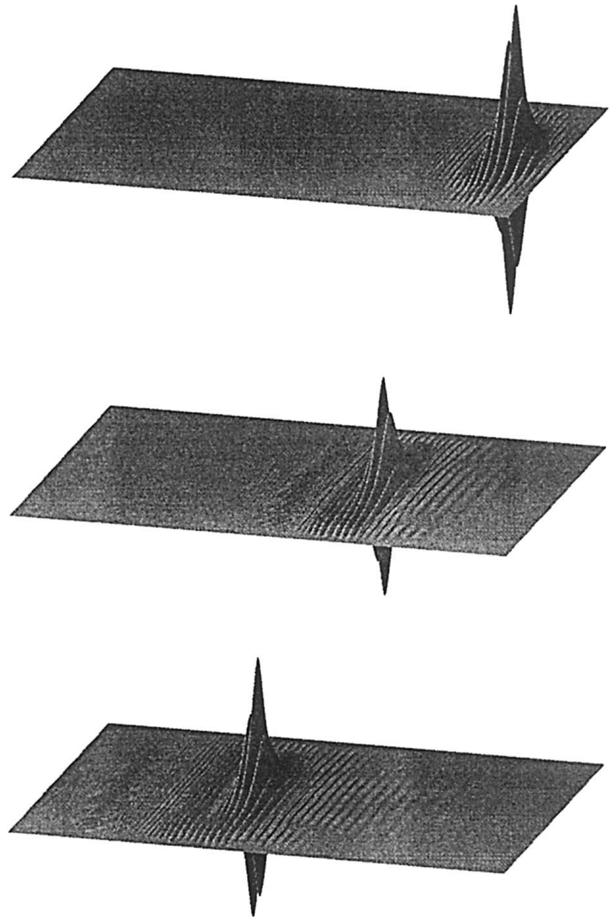

Fig. 1. Maxwell TE "light bullets" (courtesy of Paul Pax). Pulse parameters are: duration = 27 fs, width = 0.6 μm. Domain dimensions are 9.25 by 46.25 μm. Pulse group velocity is 0.2225c.

NLS much less applicable [15,20,21]. Kaplan and Shkolnikov [20,21] discussed the existence and propagation of one space dimensional unipolar pulses (light bubbles) with no EM oscillations for the dissipationless two level atom Maxwell–Bloch system, based on early work of Eilbeck et al. [9]. Goorjian and Silberberg [15] numerically simulated the full Maxwell system with instantaneous Kerr ($\chi^{(3)}$ or cubic) nonlinearity, and observed TE (transverse electric) light bullets in two space dimensions that are short femtosecond pulses and propagate without essentially changing shapes over a long distance. The light bullets have only a few EM oscillations under their envelopes. In Fig. 1, we show a three-dimensional plot of the propagating Maxwell light bullets, see also [15].

Unlike light bubbles that are constant speed traveling waves in one dimension, the light bullets are dynamic objects with distinct phase and group velocities. Even though direct numerical simulations on the full Maxwell system are convincing [4,5], there is an obvious lack of analytical understanding of their origin. It is the goal of this paper to provide a combined asymptotic and numerical approach to unravel the mystery behind.

Our approach starts with the much debated NLS approximation. In one space dimension, the cubic focusing NLS for $\chi^{(3)}$ medium is globally well-posed and its validity as an efficient approximation well-understood [22]. A comparison of Maxwell solutions and those of an extended NLS [18] also shows that the cubic NLS approximation works reasonably well on short stable one-dimensional pulses. Extending cubic NLS in this case is more of a technical improvement. In two space dimensions, making NLS approximation is much more delicate because the collapse of the critical NLS (CNLS) occurs in finite time and the envelope approximation with cubic NLS breaks down, see [11,12,30,32] among others. On the other hand, the Maxwell system itself typically behaves just fine at and beyond the NLS collapse time, due to the intrinsic smoothing physical mechanism or material response. One example is the two level dissipationless Maxwell–Bloch system, where smooth solutions persist forever [8]. It is however not clear how to modify the NLS in general to be sure that all correct physics have been captured. Many ad hoc procedures exist in the literature [1,5,6], etc. In spite of the difficulty, we shall examine a distinguished asymptotic limit of the two level dissipationless TE version Maxwell–Bloch system, which reduces to the scalar two space dimensional Sine–Gordon (SG) equation via a nonlinear change of variables. We carry out the envelope expansion, derive a new and complete perturbed CNLS equation by removing all resonance terms (*complete NLS approximation*), thanks to the explicit form of nonlinearity. The new envelope equation is second order in space–time, contains a nonparaxiality term, a mixed derivative term, and a novel nonlinear term that is saturating for large amplitude. This equation is globally well-posed and does not have finite time collapse. We perform a singular perturbation (modulation) analysis as in [11,12], and found that pulse peaks go up and down recurrently in time due to the competition between focusing and defocusing terms in the equation. The modulation analysis also shows the magnitude of solutions during the focusing–defocusing cycles, and allows us to extract the order of the error of the envelope approximation from a nonlinear Klein–Gordon equation.

Next we perform a direct numerical simulation of the two-dimensional SG, which is a much simpler task than simulating the full Maxwell. Our initial condition is just the leading term of the envelope approximation yet with the scale of separation much weakened so that there are only two or three oscillation cycles. The complete NLS approximation and the resulting asymptotics turn out to be extremely robust. The asymptotic behavior of the modulation analysis *persists into the regime with weak or almost no separation of scales*. We observe from our simulation the SG light bullets that propagate without essentially changing shapes over a long time, and that they look just like those Maxwell TE light bullets. Moreover, the pulse peaks oscillate in time with a slight dissipation due to radiation, a spectacular agreement with the asymptotics! We understand that light bullets are simply the manifestation of the persistence and robustness of the complete NLS approximation to a Maxwell type nonlinear wave equation (NLW)!

The rest of the paper is organized as follows. In Section 2, we give the physical motivation and properties of the two Maxwell systems used in direct simulations of the light bullets [3,4,5]. We derive the two-dimensional SG from the TE version of the semiclassical Maxwell–Bloch system, and also remark on bubble solutions and extensions. In Section 3, we derive the new and complete NLS approximation of SG, and the nonlinear Klein–Gordon error equation. In Section 4, we study properties of the new saturating nonlinearity, and contrast it with the existing ones. In Section 5, we carry out modulation analysis of the perturbed NLS and derive the reduced pulse dynamics. In Section 6, we illustrate SG light bullets numerically, and compare with the Maxwell light bullets. In Section 7, we summarize our findings and put forth a theory of light bullets.

In Appendix A, we prove the global well-posedness of the new NLS. It is an interesting and challenging task to rigorously prove the validity of the modulation analysis and the convergence of the complete NLS approximation to the SG beyond the collapse time of the standard CNLS.

2. Physical motivation and the SG limit

Light bullets have been recently observed in numerical simulations of both classical and semi-classical Maxwell systems. Goorjian and Silberberg [15] considered the classical model with Kerr-like $\chi^{(3)}$ cubic instantaneous nonlinearity:

$$\mathbf{D}_{tt} = -\frac{1}{\mu_0}\nabla \times (\nabla \times \mathbf{E}), \qquad \mathbf{D} = \epsilon_0[\epsilon_\infty \mathbf{E} + \chi^{(3)}(\mathbf{E}\cdot\mathbf{E})\mathbf{E} + \mathbf{P}], \qquad \mathbf{P}_{tt} + \omega_0^2 \mathbf{P} = \omega_0^2(\epsilon_s - \epsilon_\infty)\mathbf{E}, \quad (2.1)$$

where μ_0 and ϵ_0 are vaccum permeability and permitivity respectively; ϵ_s and ϵ_∞ are low and high frequency linear relative permittivities respectively; ω_0 is the medium resonance frequency. System (2.1) is three-dimensional in space, and is a coupled system of the electric field **E** and the polarization **P**. The electric displacement **D** is expressed a function of **E** and **P**. The first simplification is the restriction to TE (transverse electric) solutions of the form: $\mathbf{E} = (0, E(t, x, z), 0)$, $\mathbf{P} = (0, P(t, x, z), 0)$. For them, Eq. (2.1) reduces to a system of two scalar wave equations in two space dimensions:

$$(E + E^3)_{tt} - \Delta_\perp E = -P_{tt}, \qquad P_{tt} + \omega_0^2 P = \omega_0^2(\epsilon_s - \epsilon_\infty)E, \quad (2.2)$$

where $\Delta_\perp = \partial_{xx}^2 + \partial_{zz}^2$ is the transverse Laplacian operator.

Numerical solutions of Goorjian and Silberberg [15] illustrated the propagation of TE light bullets, i.e. spatially localized pulses that (1) are self-supporting without essentially changing shapes under the effects of dispersion and nonlinearity; (2) contain only a few oscillation cycles under their envelopes; (3) are short pulses of 100–600 fs duration. This is in contrast to the usual NLS envelope solutions which typically contain 100–1000 oscillation cycles and are therefore long pulses. In other words, there is *no separation of scale present* in light bullets between the envelope scale and the oscillation scale.

However, the instantaneous Kerr nonlinearity has an unpleasant drawback in that (2.1) or (2.2) is a quasilinear hyperbolic system, and solutions in general develop shocks in finite time [26]. This is later seen either through a direct numerical simulation [13], or analysis [31,34], and refined high order computation [19] of a limiting scalar equation:

$$(E + E^3)_{tt} - \Delta_\perp E = -E. \quad (2.3)$$

We derive (2.3) from (2.4) by taking the limit $\omega_0 \downarrow 0$, while keeping $\omega_0^2(\epsilon_s - \epsilon_\infty) = 1$. Shocks form if the initial data has enough spatial gradient [34], disappear and reappear recurrently in time due to the regularizing $-E$ term [19]. A hyperbolic system of similar nature arising in resonant nonlinear acoustics has been studied in [23].

It is more physical to write down separate dynamic equations for the medium response instead of postulating an instantaneous nonlinear function such as the $\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{P})$ relation in (2.1). Indeed, this is the case in a well-known semiclassical model describing light propagation through a medium treated as a quantum system with two energy levels, see [29].

$$\epsilon_\infty \mathbf{E}_{tt} + c^2 \nabla \times \nabla \times \mathbf{E} = -\mathbf{P}_{tt}, \qquad \mathbf{P}_{tt} + \omega_0^2 \mathbf{P} = \omega_0^2(\epsilon_s - \epsilon_\infty)N_0^{-1} N \mathbf{E}, \qquad N_t = -4\mathbf{E}\cdot\mathbf{P}_t, \quad (2.4)$$

where **E** and **P** are the electric field and the medium polarization, respectively, and the scalar N is the difference between the number of electrons in the excited state and the ground state per unit volume. N_0 is the equilibrium value of N. If damping terms are added, (2.4) is also called a system of Maxwell–Bloch equations [8]. With or without damping, (2.4) is a semilinear hyperbolic system and is free from shocks. In fact, it even has persistence of smooth solutions for all time, see [8] for a proof. The reduced TE system can be written as:

$$\epsilon_\infty E_{tt} - c^2 \Delta_{x,y} E = -P_{tt}, \qquad P_{tt} + \omega_0^2 P = \omega_0^2(\epsilon_s - \epsilon_\infty)N_0^{-1} N E, \qquad N_t = -4EP_t. \quad (2.5)$$

The other interesting feature is that in one space dimension, (2.5) admits a family of closed-form solitary wave solutions:

$$E = \left(\frac{N_0(\epsilon_s v^2 - c^2)}{(\epsilon_s - \epsilon_\infty)(c^2 - \epsilon_\infty v^2)} \right)^{1/2} \operatorname{sech} \left(\frac{\omega_0}{v} \sqrt{\frac{\epsilon_s v^2 - c^2}{c^2 - \epsilon_\infty v^2}} (x - vt) \right), \qquad P = \frac{c^2 - \epsilon_\infty v^2}{v^2} E,$$
$$N = N_0 - 2 \frac{c^2 - \epsilon_\infty v^2}{v^2} E^2, \tag{2.6}$$

where the wave velocity $v \in (c/\sqrt{\epsilon_s}, c/\sqrt{\epsilon_\infty})$. The full model is however not integrable. These solitary solutions have no internal structures and are called nonoscillating EM pulses or light bubbles [20]. Numerics suggest that they are rather stable and interact like solitons [20], yet unstable to two-dimensional perturbations [14].

It is tempting to inquire whether there are any localized two-dimensional light bubbles of the form $(E, P, N) = (E, P, N)(x - vt, y)$. The system in the moving frame reduces to a single equation for E, provided (E, P) decays to zero along x direction. Let $G = G(\xi, y)$ solve the PDE:

$$(c^2 - \epsilon_\infty v^2) G_{\xi\xi} + G_{yy} + (c^2 v^{-2} - \epsilon_s) G + 2N_0^{-1}(\epsilon_s - \epsilon_\infty)(c^2 v^{-2} - \epsilon_\infty) G^3 + v^{-2} \int_{-\infty}^{\xi} \int_{-\infty}^{\xi} G_{yy}$$
$$+ 4N_0^{-1} v^{-2} (\epsilon_s - \epsilon_\infty) \int_{-\infty}^{\xi} G \int_{-\infty}^{\xi} G_{yy} = 0, \tag{2.7}$$

then $E(t, x, y) = G(\omega_0(x - vt), \omega_0 y/c)$. The other variables (P, N) are easily expressed in terms of E. Let us suppose that $v \in (c/\sqrt{\epsilon_s}, c/\sqrt{\epsilon_\infty})$ as in the one-dimensional case. Then looking at the linear part of (2.7) in Fourier space, we see that it is elliptic for small scale ξ but changes type to hyperbolic for large scale ξ. It is not clear whether (2.7) has any nontrivial localized two-dimensional solutions.

In order to gain insight into the problem, we consider the low frequency limit, $\omega_0 \to 0$, and $\omega_0^2(\epsilon_s - \epsilon_\infty)/N_0 = 1/4$, which is achieved by taking $\epsilon_\infty = N_0 = 1$, $\epsilon_s = \mathrm{O}(\omega_0^{-2})$. Scaling $N \to 4N$, we have as $\omega_0 \to 0$ the limiting system:

$$(\partial_{tt} - c^2 \Delta) E = -(P_t)_t, \tag{2.8}$$

$$(P_t)_t = NE, \tag{2.9}$$

$$N_t = -E\, P_t. \tag{2.10}$$

This is a regular limit, and convergence of solutions is straightforward. In particular, we see that the exact solitary waves (2.6) converge. Now, let us make the change of variables:

$$E = \Phi_t, \quad P_t = \sin \Phi, \quad N = \cos \Phi, \tag{2.11}$$

where Φ plays the role of a potential function. Eqs. (2.9) and (2.10) are automatically satisfied. Eq. (2.8) becomes the scalar equation:

$$(\partial_{tt} - c^2 \Delta) \Phi_t = -(\sin \Phi)_t. \tag{2.12}$$

In particular, if we choose Φ to be a solution to the two-dimensional SG:

$$(\partial_{tt} - c^2 \Delta) \Phi + \sin \Phi = 0, \tag{2.13}$$

then representation (2.11) solves the limiting two level atom Maxwell system!

We learn immediately from (2.13) and the Derrick's theorem [7,25] that there are no nontrivial localized bubble solutions of the form $\Phi(x - vt, y)$! The variational structure in the SG equation allowed a critical point argument [7] from the energy. It remains to find out whether nonexistence holds for the system (2.5), as Derrick originally proposed (to look at systems of wave equations). In the one-dimensional case, the exact solitary solutions are directly related via (2.11) to the SG kinks as $\omega_0 \to 0$. Hence, we can trace the stable solitary nature of the bubble solutions back to the integrable one-dimensional SG.

The direct numerical simulation by Glasgow and Moloney [14] suggested that light bubbles do not exist in two dimensions and observed instead light bullets with internal oscillations. The (semiclassical) light bullets are similar to those classical ones in [15] and they also appear to be quite dynamic [14].

In the coming sections, we shall see that the SG equation has very similar looking light bullets of its own. Yet SG, being the simplest nonlinear wave equation in two dimensions, allows us to derive a new and *complete* NLS approximation and understand the origin of light bullets!

3. Derivation of a perturbed CNLS from SG

Let us consider the two-dimensional sine–Gordon equation:

$$u_{tt} - c^2\Delta_{x,y}u + \sin u = 0, \tag{3.1}$$

and look for a modulated planar pulse solution of the form:

$$u = \epsilon A(\epsilon(x-\nu t), \epsilon y, \epsilon^2 t)\mathrm{e}^{\mathrm{i}(kx-\omega(k)t)} + \text{c.c.} + \epsilon^3 u_2, \quad \omega = \sqrt{1+c^2k^2}, \tag{3.2}$$

where $\omega = \omega(k) = \sqrt{1+c^2k^2}$, $\nu = \omega'(k)$ the group velocity, and c.c. refers to the complex conjugate of the previous term. Setting $X = \epsilon(x-\nu t)$, $Y = \epsilon y$, $T = \epsilon^2 t$, calculating derivatives and expressing the sine function in series, we obtain the resulting equations at different orders as:

$$\mathrm{O}(\epsilon): \quad -\omega^2(k) + c^2k^2 + 1 = 0,$$

$$\mathrm{O}(\epsilon^2): \quad 2\mathrm{i}\omega\nu + (2\mathrm{i}k)(-c^2) = 0 \quad \text{or} \quad \nu = \frac{c^2k}{\omega} = \omega'(k),$$

which are dispersion relation and the group velocity formula we adopted. At $\mathrm{O}(\epsilon^3)$ and beyond (setting $R \equiv \epsilon^3 u_2$), we have:

$$\begin{aligned} 0 = {} & R_{tt} - c^2\Delta_{x,y}R + R + \Big[\epsilon^3(-2\mathrm{i}\omega)A_T + \epsilon^3\nu^2 A_{XX} - 2\epsilon^4\nu A_{XT} + \epsilon^5 A_{TT} - c^2\epsilon^3 A_{YY} \\ & -c^2\epsilon^3 A_{XX}\Big]\mathrm{e}^{\mathrm{i}(kx-\omega t)} + \text{c.c.} + \sum_{j=1}^{\infty}(-1)^j[(2j+1)!]^{-1}\left(\epsilon A\mathrm{e}^{\mathrm{i}(kx-\omega t)} + \text{c.c.}\right)^{2j+1} \\ & + \sum_{j=1}^{\infty}(-1)^j[(2j+1)!]^{-1}\left\{\left[\epsilon A\mathrm{e}^{\mathrm{i}(kx-\omega t)} + \text{c.c.} + R\right]^{2j+1} - \left[\epsilon A\mathrm{e}^{\mathrm{i}(kx-\omega t)} + \text{c.c.}\right]^{2j+1}\right\}. \end{aligned} \tag{3.3}$$

Let us write (3.3) as:

$$\begin{aligned} 0 = {} & R_{tt} - c^2\Delta_{xy}R + R + \mathrm{e}^{\mathrm{i}(kx-\omega t)}\Big[\epsilon^3(-2\mathrm{i}\omega)A_T + \epsilon^3(\nu^2-c^2)A_{XX} - \epsilon^3c^2A_{YY} - 2\epsilon^4\nu A_{XT} + \epsilon^5 A_{TT} \\ & + \sum_{j=1}^{\infty}(-1)^j[(2j+1)!]^{-1}\cdot\epsilon^{2j+1}\cdot|A|^{2j}\cdot A\cdot\binom{2j+1}{j+1}\Big] + \text{c.c.} + F'_1\left(\epsilon A, \mathrm{e}^{\mathrm{i}(kx-\omega t)}\right) \\ & + R\cdot\sum_{j=1}^{\infty}(-1)^j[(2j)!]^{-1}\left(\epsilon A\mathrm{e}^{\mathrm{i}(kx-\omega t)} + \text{c.c.}\right)^{2j} + F_2\left(\epsilon A\cdot\mathrm{e}^{\mathrm{i}(kx-\omega t)}, R\right), \end{aligned} \tag{3.4}$$

where F'_1 contains cubic and higher powers of ϵA, multiplied by $e^{ij(kx-\omega t)}$, $j \neq \pm 1$, which are all nonresonant; F_2 contains quadratic and higher powers of R multiplied by powers of $\epsilon A e^{i(kx-\omega t)}$.

Now we remove all the resonant terms in (3.4) by setting the bracket to zero:

$$(-2i\omega)A_T + \epsilon^2 A_{TT} = (c^2-\nu^2)A_{XX} + c^2 A_{YY} + 2\epsilon\nu A_{XT} + \epsilon^{-3}\sum_{j=1}^{\infty}(-1)^{j+1}\frac{\epsilon^{2j+1}|A|^{2j}A}{(2j+1)!}\begin{pmatrix}2j+1\\ j+1\end{pmatrix}, \tag{3.5}$$

where

$$c^2-\nu^2 = c^2 - \frac{c^4k^2}{1+c^2k^2} = \frac{c^2}{1+c^2k^2} = \frac{c^2}{\omega^2}.$$

So (3.5) simplifies to:

$$(-2i\omega)A_T + \epsilon^2 A_{TT} = \frac{c^2}{\omega^2}A_{XX} + c^2 A_{YY} + 2\epsilon\nu A_{XT} + \epsilon^{-3}\sum_{j=1}^{\infty}(-1)^{j+1}[(j+1)!\,j!]^{-1}\cdot\epsilon^{2j+1}|A|^{2j}A. \tag{3.6}$$

We shall see that Eq. (3.6) is a perturbed critical NLS (CNLS), with both a saturating nonlinearity (the series) term and nonparaxial terms (the A_{TT} and A_{XT} terms). In the standard derivation of envelope equations, only the leading order cubic term of the saturating nonlinearity is kept, the remaining nonlinear terms and the nonparaxial terms are completely ignored. As a result, one has a focusing two-dimensional NLS with cubic nonlinearity and critical collapse occurs even though the original SG has no collapse. We shall analyze (3.6) as $\epsilon \downarrow 0$ with the modulation theory [11]. Under a condition on $\omega(k)$, we shall show that the critical collapse is arrested, and pulse peak oscillates in time as the focusing and defocusing mechanisms in the equation balance each other. During this time, $\epsilon|A| \sim O(\beta^{1/2})$ at most, where β is the small excessive power above critical, independent of ϵ.

The remaining error equation is

$$R_{tt} - c^2\Delta_{x,y}R + \left(1 - \sum_{j=1}^{\infty}(-1)^j[(2j)!]^{-1}\left(\epsilon A e^{i(kx-\omega t)} + \text{c.c.}\right)^{2j}\right)R$$
$$= -F'_1(\epsilon A, e^{i(kx-\omega t)}) - F_2(\epsilon A\cdot e^{i(kx-\omega t)}, R), \tag{3.7}$$

where F' is nonresonant and F_2 is quadratic in R. For $\epsilon|A| \sim O(\beta^{1/2}) \ll 1$, over $t \in [0, T_0/\epsilon^2]$, and small initial value of R, we see that (3.7) is a basically the nonlinear Klein–Gordon equation with quadratic nonlinearity and nonresonant forcing. Recent results [28] showed that quadratic nonlinearities are irrelevant for small data in two-dimensional Klein–Gordon equation in the sense that solutions eventually converge to those of the linear Klein–Gordon. Hence, to leading order, the main contribution comes from the nonresonant forcing $F' \sim O(\beta^{3/2})$, and so

$$R(t,x,y) \sim O(\beta^{3/2}), \tag{3.8}$$

which is truly less than the leading order term $\epsilon A e^{i(kx-\omega t)} \sim O(\beta^{1/2})$. Certainly, the above argument is based on the presumption that $\epsilon|A| \sim O(\beta^{1/2}) \ll 1$, over $t \in [0, T_0/\epsilon^2]$ which will be substantiated by the modulation analysis for solutions with initial power slightly above critical. It is very interesting to give a rigorous mathematical proof of error estimates in the future.

So for initial data with power just above critical (but independent of ϵ) the envelope approximation (3.2) is valid uniformly in ϵ. We notice that since $A \sim O(\epsilon^{-1}\beta^{1/2})$, any resonant term in the first series on the right hand side of

(3.4) will induce a growth of order $O(t) \sim O(T_0/\epsilon^2)$, which is why we have removed all the resonant terms in Eq. (3.4). This is the major difference from the conventional envelope approximation where higher order (above order four) resonant terms do not matter to the dynamics on the timescale $t \sim O(\epsilon^{-2})$.

4. The saturating nonlinearity

Let us first study some properties of the nonlinearity in Eq. (3.6). Define the function:

$$f_0(x) \equiv x \sum_{n=1}^{\infty} \frac{(-1)^n (|x|^2)^n}{n!(n+1)!}, \tag{4.1}$$

and the nonlinearity of (3.6), denoted by $N(\epsilon, A)$, is written as:

$$N(\epsilon, A) = -\epsilon^{-3} f_0(\epsilon A). \tag{4.2}$$

Clearly for $|x| \ll 1$, $f_0(x) \sim -\frac{1}{2}|x|^2 x + O(|x|^5)$, so

$$N(\epsilon, A) = \tfrac{1}{2}|A|^2 A + O(\epsilon^2 |A|^5). \tag{4.3}$$

We now study the large $|x|$ behavior of $f_0(x)$, for which it is convenient to analyze the function:

$$f(y) = \sum_{n\geq 1} \frac{y^n}{n!(n+1)!}, \quad \text{as } y \to \infty. \tag{4.4}$$

Notice that

$$\begin{aligned} f(y) &= \sum_{n\geq 1} \frac{y^n}{n!n!} - \sum_{n\geq 1} \frac{y^n}{(n-1)!(n+1)!} = \left(\int_0^y \sum_{n\geq 1} \frac{y^{n-1}}{(n-1)!n!} \, dy \right) - \frac{y}{2!} - \sum_{n\geq 1} \frac{y^{n+1}}{n!(n+2)!} \\ &= \int_0^y \left(1 + \sum_{n\geq 1} \frac{y^n}{n!(n+1)!} \right) - \frac{y}{2} - \frac{1}{y} \int_0^y \sum_{n\geq 1} \frac{y^{n+1}}{n!(n+1)!} \\ &= \int_0^y (1 + f(y'))\, dy' - \frac{y}{2} - y^{-1} \int_0^y y' f(y')\, dy'. \end{aligned}$$

and so

$$f' = 1 + f(y) - \frac{1}{2} - f(y) + y^{-2} \int_0^y y' f(y')\, dy' = \frac{1}{2} + y^{-2} \int_0^y y' f(y')\, dy'. \tag{4.5}$$

It follows that:

$$\begin{aligned} &y^2 f' = \frac{y^2}{2} + \int_0^y y' f(y')\, dy', \\ &(y^2 f')' = y + y f(y), \\ &\text{or } y^2 f'' + 2y f' - y f = y, \text{ or : } y f'' + 2 f' - f = 1. \end{aligned} \tag{4.6}$$

Clearly, a special solution of (4.6) is $f = -1$ so $f = -1 + \tilde{f}(-y)$, where $\tilde{f} = \tilde{f}(\xi)$, $\xi = -y$, satisfies

$$-\xi \tilde{f}'' - 2\tilde{f}' - \tilde{f} = 0 \quad \text{or} \quad \tilde{f}'' + 2\xi^{-1}\tilde{f}' + \xi^{-1}\tilde{f} = 0, \tag{4.7}$$

where we are interested in the behavior as $\xi \to +\infty$. Making the change of variable $\tilde{f} = \mathrm{e}^{a(\xi)} g$, we find

$$g'' + (2a' + 2\xi^{-1})g' + (a'' + 2\xi^{-1}a' + \xi^{-1} + a^{2'})g = 0. \tag{4.8}$$

Letting $a' = -\xi^{-1}$, then

$$a'' + \frac{2a'}{\xi} + a'^2 + \frac{1}{\xi} = \frac{1}{\xi^2} - \frac{2}{\xi^2} + \frac{1}{\xi^2} + \frac{1}{\xi} = \frac{1}{\xi},$$

thus

$$g'' + g_1 g \equiv g'' + \frac{1}{\xi} g = 0. \tag{4.9}$$

The large ξ behavior of g is [17]:

$$g \sim g_1^{-1/4}(\xi) \exp\left\{ \pm \mathrm{i} \int^{\xi} g_1^{1/2}(\xi') \left(1 - \frac{g_1'^2}{16 g_1^3} \right)^{1/2} \mathrm{d}\xi' \right\}, \tag{4.10}$$

where $g_1(\xi) > 0$ is smooth and obeys

$$\int^{\infty} \left| \left(\frac{g'_1}{g_1^{3/2}} \right)' \right| < \infty, \quad \gamma = \lim_{\xi \to \infty} \frac{g'_1}{4 g_1^{3/2}}, \quad \gamma^2 \neq 1. \tag{4.11}$$

With $g_1(\xi) = 1/\xi$ in our particular case (4.9), it is easy to check that Eq. (4.11) is valid for $g_1(\xi) = 1/\xi$ and $\gamma = 0$. Hence (as $\xi \to +\infty$):

$$\begin{aligned} g(\xi) &\sim \xi^{1/4} \exp\left\{ \pm \mathrm{i} \int^{\xi} \frac{1}{\xi'^{1/2}} \left(1 - \frac{1}{16\xi'} \right)^{1/2} \mathrm{d}\xi' \right\} \\ &\sim \xi^{1/4} \exp\left\{ \pm \mathrm{i} \left(2\xi^{1/2} + \text{const.} + \frac{1}{16}\xi^{-1/2} + \text{h.o.t.} \right) \right\}, \end{aligned} \tag{4.12}$$

which implies that

$$\begin{aligned} \tilde{f}(\xi) &\sim \frac{1}{\xi} \cdot \xi^{1/4} \exp\left\{ \pm \mathrm{i} \left(2\xi^{1/2} + \text{const.} + \frac{1}{16}\xi^{-1/2} + \text{h.o.t.} \right) \right\} \\ &\sim \xi^{-3/4} \exp\left\{ \pm \mathrm{i} \left(2\xi^{1/2} + \text{const.} + \frac{1}{16}\xi^{-1/2} + \text{h.o.t.} \right) \right\}, \quad \xi \to +\infty. \end{aligned} \tag{4.13}$$

Finally,

$$f = f(y) \sim -1 + (-y)^{-3/4} \exp\left\{ \pm \mathrm{i} \left(2(-y)^{1/2} + \text{const.} + \frac{1}{16}(-y)^{-1/2} + \text{h.o.t.} \right) \right\}, \quad \text{as } y \to -\infty. \tag{4.14}$$

So by (4.1)–(4.4)

$$\begin{aligned} f_0(\epsilon A) &\sim \epsilon A \cdot f(-\epsilon^2 |A|^2) \\ &\sim \epsilon A \cdot \left\{ -1 + (\epsilon |A|)^{-3/2} \exp\left[\pm \mathrm{i} \left(2\epsilon |A| + \text{const.} + \frac{1}{16}(\epsilon |A|)^{-1} + \text{h.o.t.} \right) \right] \right\}, \quad \text{if} \quad \epsilon |A| \gg 1, \end{aligned} \tag{4.15}$$

so the nonlinearity $N(\epsilon, A)$ is asymptotic to linear (saturating) $\sim \epsilon^{-2} A$ as $\epsilon |A| \gg 1$.

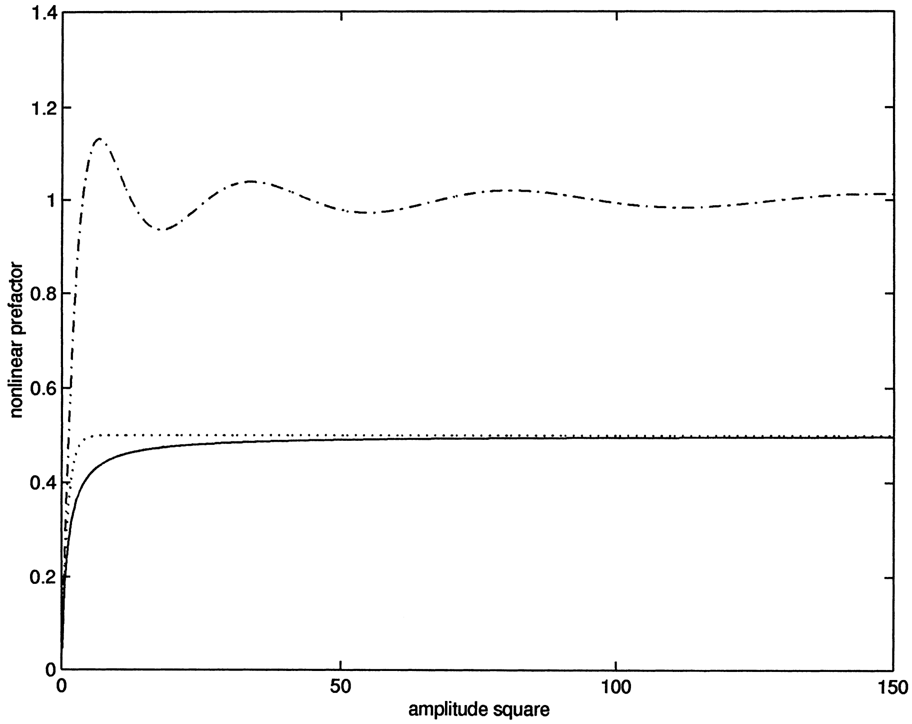

Fig. 2. Comparison of the new saturated nonlinear prefactor ($\epsilon = 1$, dot-dashed curve) with the exponential ($(1 - \mathrm{e}^{-x})/2$, dotted curve) and fractional ($x/(1+x)$, solid curve) saturating nonlinear prefactors (x = amplitude square).

The nonlinearity $N(\epsilon, A)$ is similar to the familiar saturating nonlinearities in

$$\mathrm{i}\psi_t + \Delta\psi + (2\epsilon)^{-1}(1 - \mathrm{e}^{-2\epsilon|\psi|^2})\psi = 0, \tag{4.16}$$

and

$$\mathrm{i}\psi_t + \Delta\psi + \frac{|\psi|^2}{1 + \epsilon|\psi|^2}\psi = 0, \tag{4.17}$$

in that they are all asymptotic to linear $\sim \epsilon^{-1}\psi$ as $\epsilon|\psi| \gg 1$; all asymptotic to the focusing cubic $\sim |\psi|^2\psi$ for $\epsilon|\psi| \ll 1$. However, $N(\epsilon, A)$ is different in that its prefactor f_0 as a function of $|A|^2$ goes through a few oscillation cycles for the intermediate $\epsilon|A|$ values before approaching one, while the saturating prefactors in (4.16) and (4.17) are monotone increasing in $|\psi|^2$. In Fig. 2, we compare the nonlinear prefactors for the new nonlinearity, the exponential and the fractional nonlinearities.

5. Modulation analysis of a perturbed CNLS

We consider initial data slightly above the critical power of the unperturbed CNLS, and carry out a modulational analysis as in [11]. Let us scale (3.6) into a standard perturbed CNLS with $X = (c/\omega)\tilde{X}$, $Y = c\tilde{Y}$, $T = 2\omega\tilde{T}$, $\omega = \sqrt{1 + k^2 c^2}$

$$-\mathrm{i}A_{\tilde{T}} + \frac{\epsilon^2}{4\omega^2}A_{\tilde{T}\tilde{T}} = A_{\tilde{X}\tilde{X}} + A_{\tilde{Y}\tilde{Y}} + \frac{\epsilon\nu}{c}A_{\tilde{X}\tilde{T}} + \epsilon^{-3}\sum_{j=1}^{\infty}(-1)^{j+1}[(j+1)!j!]^{-1}\epsilon^{2j+1}|A|^{2j}A, \tag{5.1}$$

with initial data $A|_{t=0} = A_0(\tilde{X}, \tilde{Y}) \in H^1(R^2)$ such that

$$\left| \frac{1}{2\pi} \int |A_0|^2 \, \mathrm{d}\tilde{X} \, \mathrm{d}\tilde{Y} - N_\mathrm{c} \right| = \mathrm{O}(\beta_0) \ll 1, \tag{5.2}$$

with β_0 measuring the amount of excessive power above critical power

$$N_\mathrm{c} = \int_0^\infty R^2(\varrho)\varrho \, \mathrm{d}\varrho \simeq 1.86,$$

$R(\varrho)$ being the well-known profile of CNLS ground state [11]:

$$R(0) = 0, \qquad R(\varrho) \sim A_R \varrho^{-1/2} \mathrm{e}^{-\varrho}, \quad \varrho \gg 1, \quad A_R \simeq 3.52.$$

Omitting the tildes, we put (5.1) into the form:

$$0 = \mathrm{i}A_T + A_{XX} + A_{YY} + |A|^2 A + \underbrace{\epsilon^{-3} \sum_{j=2}^{\infty} \frac{(-1)^{j+1}|A|^{2j} A \cdot \epsilon^{2j+1}}{(j+1)!\, j!} - \frac{\epsilon^2}{4\omega^2} A_{TT} + \frac{\epsilon \nu}{c} A_{XT}}_{F}, \tag{5.3}$$

where F is treated as a perturbation.

The modulation theory of [11] is based on the following two assumptions:

- the focusing part of the solution is close to the asymptotic profile

$$\psi_s(T, X, Y, \cdot) \sim \frac{1}{L(T, \cdot)} V(\zeta, \xi, \eta, \cdot) \exp\left[\mathrm{i}\zeta(T, \cdot) + \mathrm{i} \frac{L_T}{L} \frac{r^2}{4} \right], \tag{5.4}$$

  where $\xi = X/L$, $\eta = Y/L$, $V = R + \mathrm{O}(\beta, \epsilon)$, $r^2 = X^2 + Y^2$,

$$\zeta_T = \frac{1}{L^2}, \qquad \beta = -L^3 L_{TT}, \tag{5.5}$$

  and "·" referring to dependence on ϵ;
- $|F| \ll |\Delta A|$ and $|A|^2 A$.

The first term in F is a saturating nonlinearity as discussed in the previous section, the second term in F with double time derivative turns out to promote focusing, and the third term is neutral, up to the leading order asymptotics we shall perform.

The reduced modulation system is

$$\beta_T + \frac{\nu(\beta)}{L^2} = \frac{(f_1)_T}{2M} - \frac{2f_2}{M}, \qquad L_{TT} = -\frac{\beta}{L^3},$$

where

$$\nu(\beta) \sim \frac{2A_R^2}{M} \mathrm{e}^{-\pi/\sqrt{\beta}}, \quad M = \frac{1}{4} \int_0^\infty R(r) r^3 \, \mathrm{d}r \simeq 0.55, \tag{5.6}$$

$$f_1(T, \cdot) = 2L(T, \cdot) \mathrm{Re} \left[\frac{1}{2\pi} \int F(\psi_R) \mathrm{e}^{-\mathrm{i}S} [R(\varrho) + \varrho R_\varrho(\varrho)] \, \mathrm{d}X \, \mathrm{d}Y \right], \tag{5.7}$$

$$f_2(T, \cdot) = \mathrm{Im} \left[\frac{1}{2\pi} \int \psi_R^* F(\psi_R) \, \mathrm{d}X \, \mathrm{d}Y \right], \tag{5.8}$$

and

$$\psi_R = \frac{1}{L} R(\varrho) \mathrm{e}^{\mathrm{i}S}, \quad S = \zeta(T, \cdot) + \frac{L_T}{L} \cdot \frac{r^2}{4}, \ \zeta_T = \frac{1}{L^2}, \ \varrho = \frac{r}{L}. \tag{5.9}$$

During self-focusing, L, on the order of the pulse width, is a small quantity; so is ν. During the adiabatic stage, β changes much less than L. Note that if $f_2 \neq 0$, f_2 is much larger than f_1, also that F is additive in terms of all perturbation terms. We calculate their contributions below with (5.9) and integration by parts.

Let us begin with the perturbation

$$\epsilon^{-3} \sum_{j=2}^{\infty} \frac{(-1)^{j+1} \epsilon^{2j+1} |A|^{2j} A}{(j+1)! j!}$$

from $N(\epsilon, A)$

$$f_2 = 0 \ \ (\text{conservative type}), \tag{5.10}$$

$$\begin{aligned} f_1 &= 2L \,\mathrm{Re} \left[\frac{\epsilon^{-3}}{2\pi} \int \mathrm{e}^{-\mathrm{i}S} \sum_{j=2}^{\infty} \frac{(-1)^{j+1} \epsilon^{2j+1} |\psi_R|^{2j} \psi_R}{(j+1)! j!} (R(\varrho) + \varrho R_\varrho) \,\mathrm{d}X \,\mathrm{d}Y \right] \\ &= 2L \mathrm{Re} \left[\frac{\epsilon^{-3}}{2\pi} \sum_{j=2}^{\infty} \frac{(-1)^{j+1} \epsilon^{2j+1}}{(j+1)! j!} \int |\psi_R|^{2j+1} (R(\varrho) + \varrho R_\varrho) \,\mathrm{d}X \,\mathrm{d}Y \right] \\ &= 2 \sum_{j=2}^{\infty} \frac{(-1)^{j+1}}{(j+1)! j!} \left(\frac{\epsilon}{L} \right)^{2j-2} \int_0^\infty \left(R^{2j+2} \varrho - \frac{1}{j+1} \varrho R^{2j+2} \right) \mathrm{d}\varrho \\ &= - \left(\frac{\epsilon}{L} \right)^2 \cdot \frac{2}{3! \cdot 3} \int_0^\infty R^6 \varrho \,\mathrm{d}\varrho + \mathrm{O}\left(\left(\frac{\epsilon}{L} \right)^4 \right) = - \frac{\alpha_0 N_\mathrm{c}}{9} \left(\frac{\epsilon}{L} \right)^2 + \mathrm{O}\left(\left(\frac{\epsilon}{L} \right)^4 \right), \end{aligned}$$

where we have used $\int_0^\infty R^6 \varrho \,\mathrm{d}\varrho = \alpha_0 N_\mathrm{c}$, with constant $\alpha_0 \in (5, 6)$. So if $\epsilon \ll L$, or $\epsilon |A| \ll 1$, then to leading order

$$RHS_1 = \frac{(f_1)_T}{2M} = - \frac{\epsilon^2 \alpha_0 N_\mathrm{c}}{18M} \left(\frac{1}{L^2} \right)_T . \tag{5.12}$$

Next for the nonparaxiality term $-(\epsilon^2 / 4\omega^2) A_{TT}$:

$$\begin{aligned} f_2 &= \mathrm{Im} \left[\frac{1}{2\pi} \int \psi_R^* (\psi_R)_{TT} \,\mathrm{d}X \,\mathrm{d}Y \right] \cdot \left(- \frac{\epsilon^2}{4\omega^2} \right) = - \frac{\epsilon^2}{4\omega^2} \mathrm{Im} \left[\frac{1}{2\pi} \int \psi_R^* \cdot R(\varrho) \left(\frac{1}{L} \mathrm{e}^{\mathrm{i}S} \right)_{TT} \mathrm{d}X \,\mathrm{d}Y \right] \\ &= - \frac{\epsilon^2}{4\omega^2} \int_0^\infty R^2(\varrho) \left(S_{TT} - \frac{2 S_T L_T}{L} \right) \varrho \,\mathrm{d}\varrho \\ &= - \frac{\epsilon^2}{4\omega^2} \int_0^\infty R^2(\varrho) \left[\left(\frac{1}{L^2} \right)_T + \left(\frac{L_{TT} L - L_T^2}{L^2} \right)_T \frac{r^2}{4} - \frac{2}{L} \left(\frac{1}{L^2} + \left(\frac{L_T}{L} \right)_T \frac{r^2}{4} \right) L_T \right] \varrho \,\mathrm{d}\varrho \\ &= - \frac{\epsilon^2}{4\omega^2} \int_0^\infty R^2(\varrho) \left[2 \left(\frac{1}{L^2} \right)_T + \varrho^2 \cdot \frac{L_T^3}{L} + \frac{L^2 \varrho^2}{4} \left(\frac{-4 L_T L_{TT}}{L^2} + \left(\frac{L_{TT}}{L} \right)_T \right) \right] \varrho \,\mathrm{d}\varrho. \end{aligned} \tag{5.13}$$

The last term in the bracket of the last line of (5.13) contains L_{TT}, L_{TTT}, which are higher order terms since they are proportional to $\beta^{1/2} \ll 1$. The second term in the same bracket is $\mathrm{O}(L_T^3 / L) \ll \mathrm{O}((1/L^2)_T)$, either a posteri or as suggested by CNLS focusing law [24]:

$$L(T) \sim \sqrt{2\sqrt{\beta}(T_c - T)}, \tag{5.14}$$

where β is a slow variable compared with T. It follows that to leading order:

$$f_2(T,\cdot) \sim \frac{-\epsilon^2 N_c}{2\omega^2}\left(\frac{1}{L^2}\right)_T, \tag{5.15}$$

implying that

$$RHS_2 \sim -\frac{2\epsilon^2}{M}\left(-\frac{N_c}{2\omega^2}\right)\left(\frac{1}{L^2}\right)_T = \frac{\epsilon^2 N_c}{M\omega^2}\left(\frac{1}{L^2}\right)_T. \tag{5.16}$$

Lastly, for the mixed derivative term $\epsilon\nu/cA_{XT}$:

$$\begin{aligned} f_2 &= \mathrm{Im}\frac{1}{2\pi}\int \psi_R^*\left[R(\varrho)\mathrm{e}^{\mathrm{i}S}/L\right]_{XT}\,\mathrm{d}X\,\mathrm{d}Y = -\mathrm{Im}\frac{1}{2\pi}\int(\psi_R^*)_X\left[R(\varrho)\mathrm{e}^{\mathrm{i}S}/L\right]_T\,\mathrm{d}X\,\mathrm{d}Y \\ &= -\frac{1}{2\pi}\int\left(\frac{L_T R^2 L_T\varrho}{2L}\cos\theta + \frac{S_T R R'}{L}\cos\theta\right)\varrho\,\mathrm{d}\varrho\,\mathrm{d}\theta = 0; \end{aligned} \tag{5.17}$$

and with $G(\varrho) = R + \varrho R_\varrho$:

$$\begin{aligned} \frac{f_1}{2L} &= \mathrm{Re}\frac{1}{2\pi}\int\left[R(\varrho)\mathrm{e}^{\mathrm{i}S}/L\right]_{XT}\mathrm{e}^{-\mathrm{i}S}\left[R + \varrho R_\varrho\right]\mathrm{d}X\,\mathrm{d}Y \\ &= -\mathrm{Re}\frac{1}{2\pi}\int\left[R(\varrho)\mathrm{e}^{\mathrm{i}S}/L\right]_T\left[\mathrm{e}^{-\mathrm{i}S}(R+\varrho R_\varrho)\right]_X\mathrm{d}X\,\mathrm{d}Y \\ &= -\frac{1}{2\pi}\int\left(-\frac{L_T}{L}RG'\cos\theta + \frac{LS_T R L_T G\varrho\cos\theta}{2}\right)\varrho\,\mathrm{d}\varrho\,\mathrm{d}\theta = 0. \end{aligned} \tag{5.18}$$

Hence the mixed derivative term makes no contribution to the modulation system.

It follows from the above calculations that the modulation system to leading order is

$$\beta_T = \frac{\epsilon^2 N_c}{M}\left[-\frac{\alpha_0}{18} + \frac{1}{\omega^2}\right]\left(\frac{1}{L^2}\right)_T \equiv \frac{\epsilon^2}{2M}\left(-\frac{C_1}{L^2}\right)_T, \quad C_1 = (2N_c)\left(\frac{1}{\omega^2} - \frac{\alpha_0}{18}\right), \tag{5.19}$$

which upon integration gives

$$-L^3 L_{TT} = \beta_0 + \frac{\epsilon^2}{2M}\left(-C_1/L^2\right), \quad \beta_0 = \beta(0,\cdot) + \frac{\epsilon^2 C_1}{2ML^2(0)}. \tag{5.20}$$

(5.20) is put into the following form upon integrating in T once

$$(y_T)^2 = -\frac{4H_0}{M}(y_M - y)(y - y_m)/y, \quad y = L^2, \tag{5.21}$$

where

$$H_0 \sim H(0) + \frac{\epsilon^2 C_1}{4L^4(0)}, \tag{5.22}$$

$$y_M = \frac{\sqrt{\frac{\beta_0^2 + \epsilon^2 C_1 H_0}{M^2}} + \beta_0}{\frac{-2H_0}{M}} = \frac{M\beta_0}{-H_0}\left(1 + \mathrm{O}\left(\frac{\epsilon^2 H_0}{\beta_0^2}\right)\right), \tag{5.23}$$

$$y_m = \frac{\epsilon^2 C_1}{2M} \frac{1}{\sqrt{\beta_0^2 + \frac{\epsilon^2 C_1 H_0}{M^2}} + \beta_0} = \frac{\epsilon^2 C_1}{4M\beta_0}\left[1 + \mathrm{O}\left(\frac{\epsilon^2 H_0}{\beta_0^2}\right)\right]. \quad (5.24)$$

We conclude from (5.21)–(5.24) that:

I. If $C_1 > 0$ or $\omega^2 < 18/\alpha_0$ ($\alpha_0 \in (5, 6)$), the perturbation arrests blow up, L remains positive for all T finite.

(1) If $\beta_0 > 0$, $H_0 < 0$, then $0 < y_m < y_M$, L goes through oscillation between $\sqrt{y_m}$ and $\sqrt{y_M}$, with period:

$$2\sqrt{M y_M/(-H_0)} \int_0^{\pi/2} \left[1 - \left(1 - \frac{y_m}{y_M}\right) \sin^2\theta\right]^{1/2} \mathrm{d}\theta. \quad (5.25)$$

(2) If $\beta_0 > 0$ and $H_0 > 0$, then when $L_T(0) < 0$, self-focusing is arrested when $L = \sqrt{y_m}$, and L is monotonically defocusing to ∞ afterward; when $L_T(0) > 0$, L is monotonically increasing to ∞.

II. If $C_1 < 0$ ($\omega^2 > 18/\alpha_0$), $\beta_0 > 0$, and either $H_0 > 0$ and $L_T(0) < 0$, or $H_0 < 0$, there exists a finite time T^* such that $L(T^*) = 0$. Solutions collapse according to (5.21).

However, before L goes to zero, when $L \ll \epsilon$, higher order terms in Eq. (5.11) must be taken into account, and this saturates the growth of solution, as suggested by the nature of nonlinearity $N(\epsilon, A)$, see also Appendix A for the global in time estimates precluding finite time collapse. This stage of evolution is outside the range of modulation theory.

Remark 5.1. *When collapse is arrested,* $L \geq \sqrt{y_m} = \mathrm{O}\left(\epsilon/\sqrt{\beta}\right)$*, and so* $\epsilon|A| \sim \epsilon(1/L) \sim \mathrm{O}(\beta^{1/2}) \leq \mathrm{O}(\beta_0^{1/2}) \ll 1$*, validating the approximation hypotheses made earlier. Following the discussion at the end of Section* 3*, we have that the error of envelope approximation is* $\mathrm{O}(\beta_0^{3/2}) \ll 1$.

The regime in I(1) with $\omega^2 < 18/\alpha_0$ *is the most interesting and will be compared with numerical solutions of SG.*

Remark 5.2. *The nonparaxiality term with* A_{TT} *promotes focusing and competes with the saturating nonlinearity. This is different from the role of smoothing it plays in the case of the Helmhotz equation* [11]. *For a review of the effects of nonparaxial terms on NLS-type systems such as on well-posedness and wave collapse, see* [2].

6. Numerical simulation of SG bullets

In this section, we simulate the SG equation:

$$u_{tt} - \Delta_{x,y} u + \sin u = 0, \quad (6.1)$$

on a square domain $[-16, 16]^2$. We discretize SG by an explicit second order finite difference scheme [33] with zero Neumann boundary condition at the edges of $[-16, 16]^2$. The initial data is chosen as:

$$u(x, y, 0) = A_0 \mathrm{e}^{-0.04(x^2+y^2)} \sin kx, \quad (6.2)$$

$$u_t(x, y, 0) = -\omega A_0 \mathrm{e}^{-0.04(x^2+y^2)} \cos kx, \quad (6.3)$$

where $\omega = \sqrt{1 + k^2}$, and is properly truncated as input to the numerical scheme. Such initial data is motivated by the NLS asymptotics, and contains only a few cycles under the envelope, like the light bullets. As an initial step, we calculate u_j^1 by Taylor expanding solution using (6.1)–(6.3), as in [33].

For all the runs we present here, $\mathrm{d}x = 0.1$, $\mathrm{d}t = 0.1 * \mathrm{d}x$, and $A_0 \in (0.75, 0.85)$. We refined the mesh by halving $\mathrm{d}x$, and observed no substantial improvement in the numerical solutions. If the value of A_0 is smaller than 0.75, the pulses quickly radiate and spread out. This corresponds to the lower threshold regime of the CNLS.

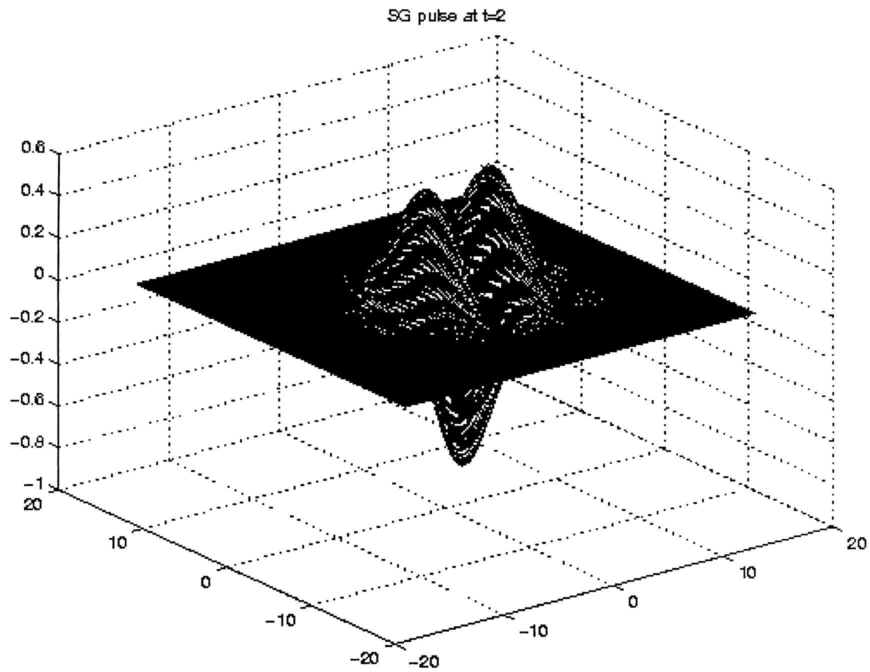

Fig. 3. SG light bullet at $t = 2$, with $k = 1$, $A_0 = 0.8$, sh $= 0.0$ (no phase shift).

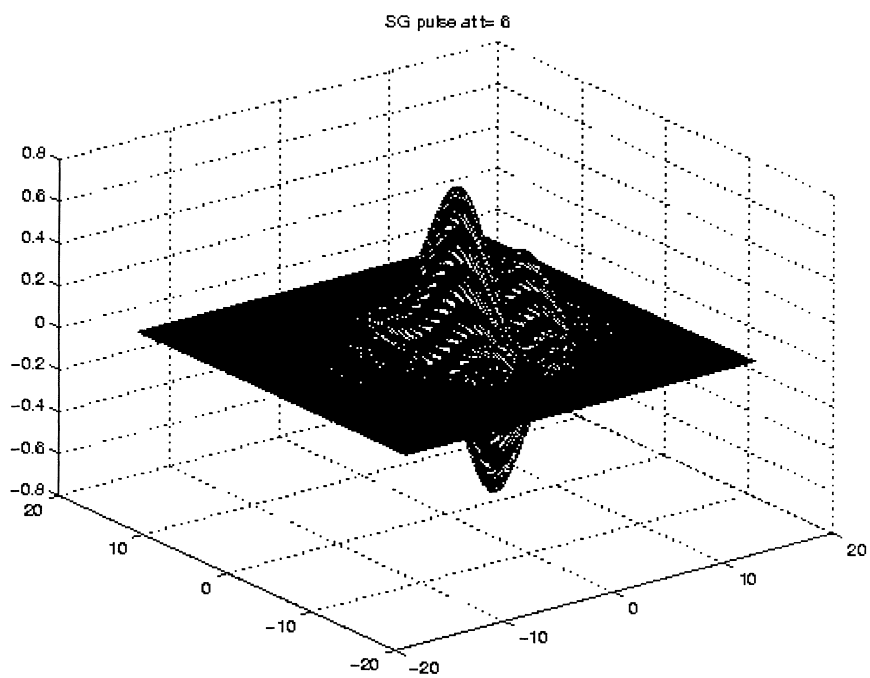

Fig. 4. SG light bullet at $t = 6$, with $k = 1$, $A_0 = 0.8$, sh $= 0.0$.

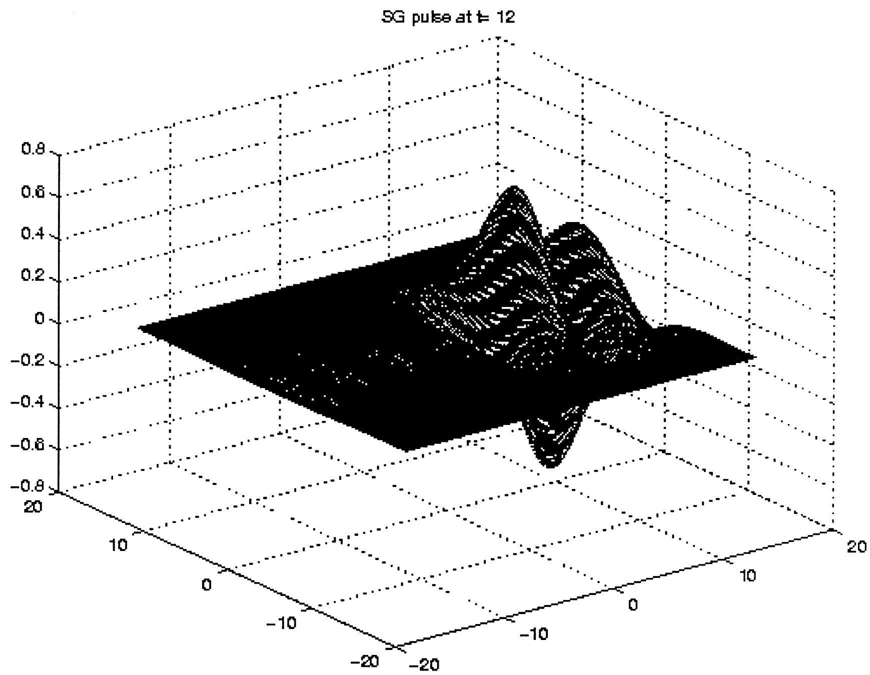

Fig. 5. SG light bullet at $t = 12$, with $k = 1$, $A_0 = 0.8$, sh $= 0.0$.

In Figs. 3–5, $A_0 = 0.8$, $k = 1$, we plot the solutions at three time slices $t = 2$ (early time), $t = 6$ (intermediate time), and $t = 12$ (long time). We see that the pulse has three major peaks and maintains its shape as it travels to the right. The three peaks are dynamic, the left one is lower than the other two at $t = 2$, yet becomes highest at $t = 6$. At $t = 12$, the pulse almost reaches the numerical boundary with essentially the same shape. There is slight radiation near the pulse, especially at the wake. In Fig. 6, $k = 2$, we have a similar pulse with six peaks at $t = 6$. These SG pulses are qualitatively the same as the Maxwell light bullets, compare with Fig. 1. We also triple the size of domains ($[-48, 48]^2$, and 960^2 grid points) with the same resolution in order to follow the pulses longer. In Figs. 7 and 8, we plot the pulse cross section along the propagation direction for $t = 0.5$ and $t = 18$. Comparing them, we see that they are almost exactly the same except that there is a little radiation in the wake of the $t = 18$ pulse.

In Fig. 9, we plot the entire history of the largest pulse peak (the maximum norm of the pulse). The peak goes up over the time $[0, 2.5]$ due to the focusing mechanism in the equation, then goes down over $t \in [2.5, 3.5]$, goes up again near $t = 3.5$, then down over $t \in [4, 5]$, then goes up once more (a larger rise), then down and up repeatedly in time. This is just like what the modulation analysis has predicted. There is a slight depletion of pulse energy as seen from the small negative slope of the peak envelop which is approximately equal to -0.01. Considering that we are so far away from the asymptotic regime of the modulation analysis on the perturbed CNLS, the agreement between the numerics and the asymptotics on the pulse oscillation and focusing–defocusing phenomenon is remarkable! The SG light bullets are very robust. We perturb the initial data above with a phase shift equal to sh $* (x^2 + y^2)$. In Fig. 10, we plot the maximum pulse history with sh$= -0.05$. The pulse has more momentum to bounce up than before, and the slow radiation decay is not apparent even at $t = 18$. It is interesting to find the optimal initial conditions so as to increase the life-span of the SG bullets and delay the radiation.

Finally, Figs. 11 and 12, are top views of the $t = 6$ SG bullets (of Figs. 4 and 6).

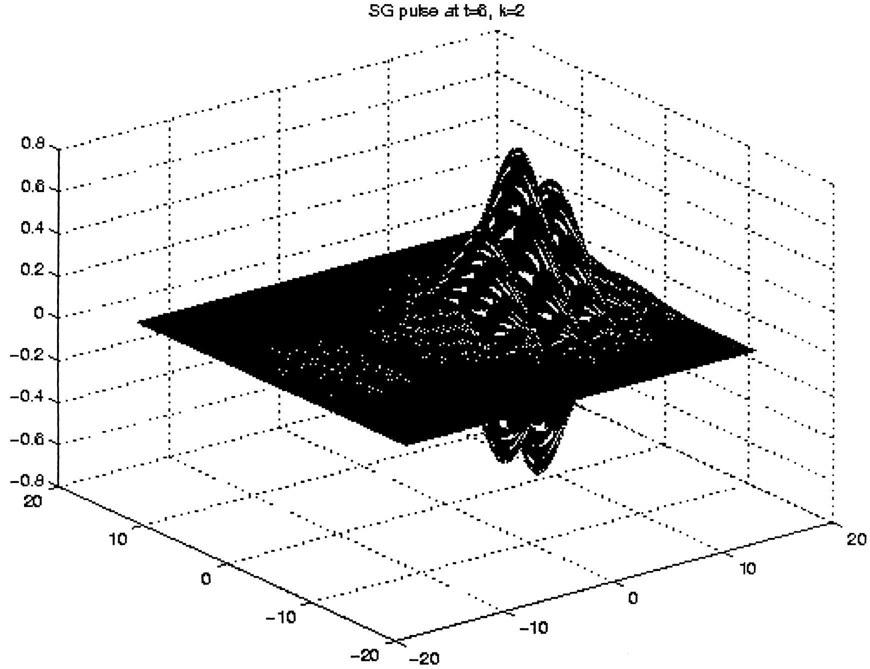

Fig. 6. SG light bullet at $t = 6$, with $k = 2$, $A_0 = 0.8$, sh $= 0.0$.

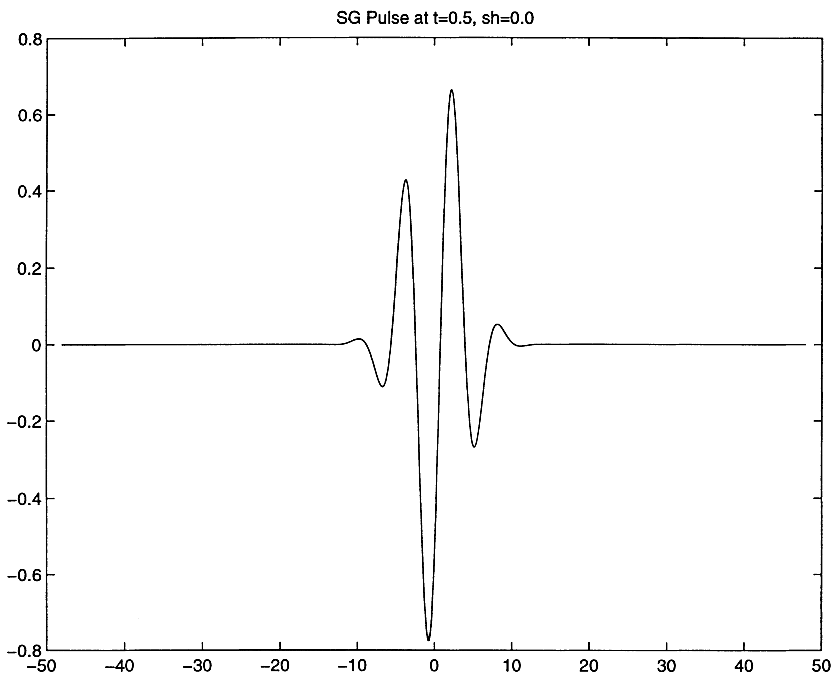

Fig. 7. SG light bullet cross-section along the propagation direction at $t = 0.5$, on $[-48, 48]^2$ square domain, $k = 1$, $A_0 = 0.8$, sh $= 0.0$.

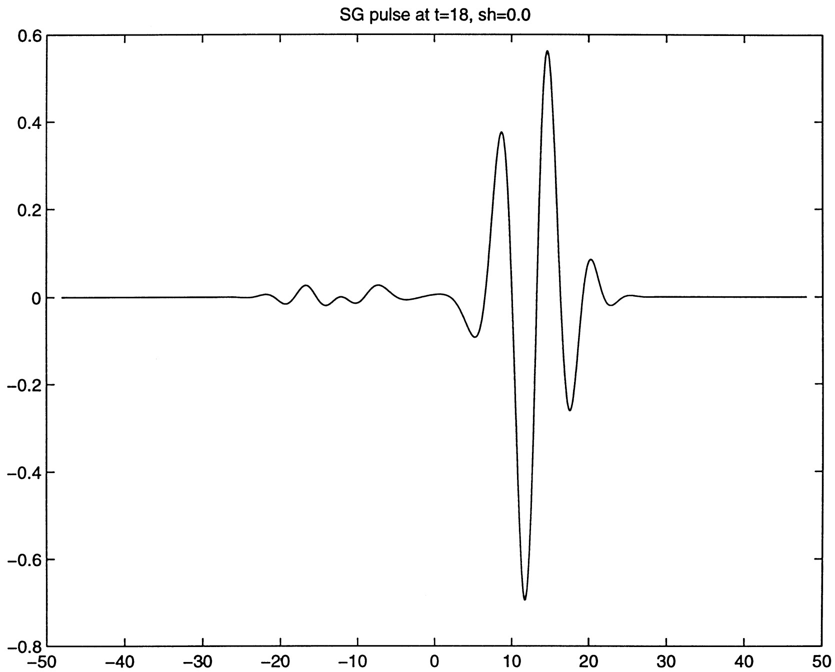

Fig. 8. SG light bullet cross section along the propagation direction at $t = 18$, parameters same as in Fig. 7.

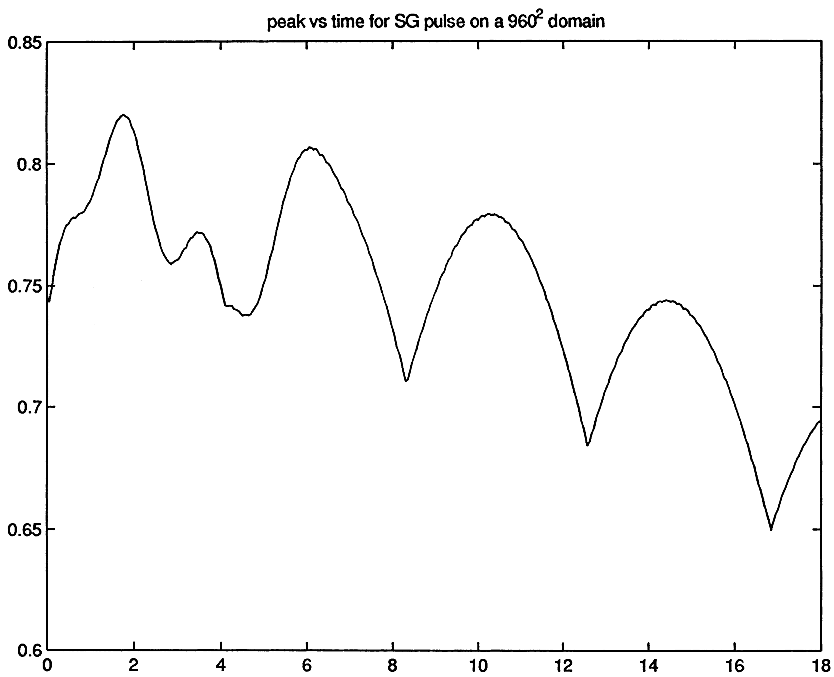

Fig. 9. SG light bullet peak history, parameters same as in Fig. 7.

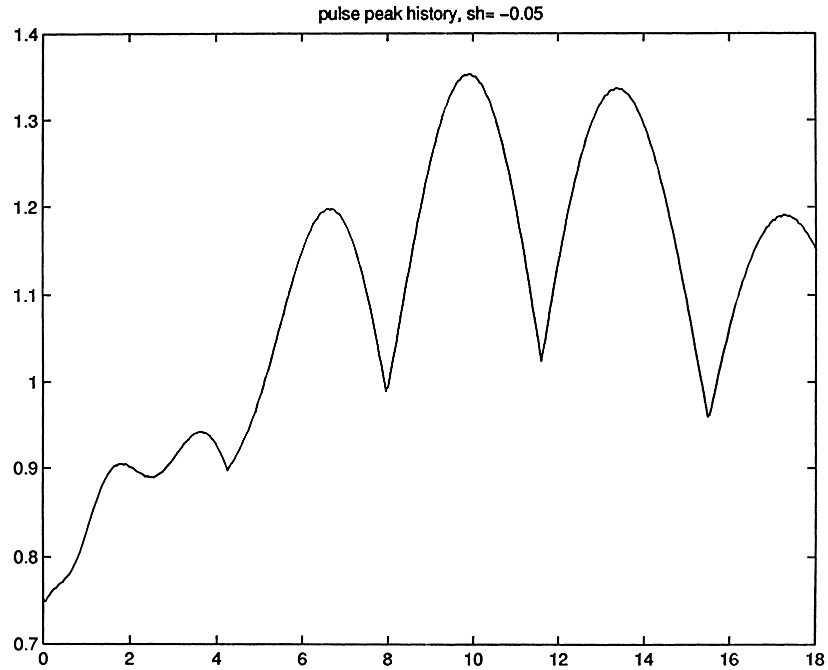

Fig. 10. SG light bullet peak history, phase shift constant sh= -0.05, other parameters same as in Fig. 7.

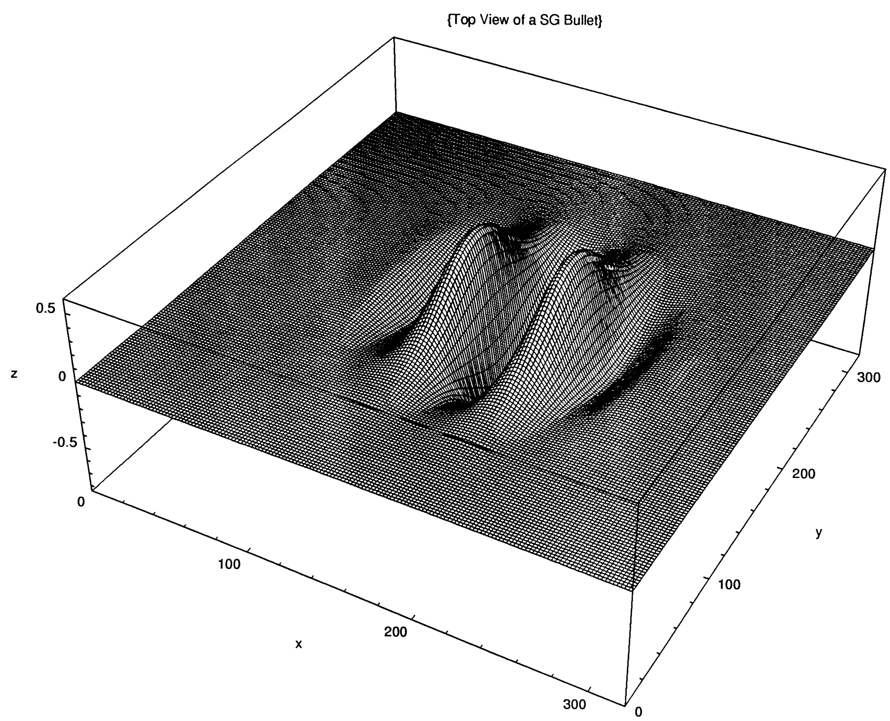

Fig. 11. Top view of SG light bullet in Fig. 4.

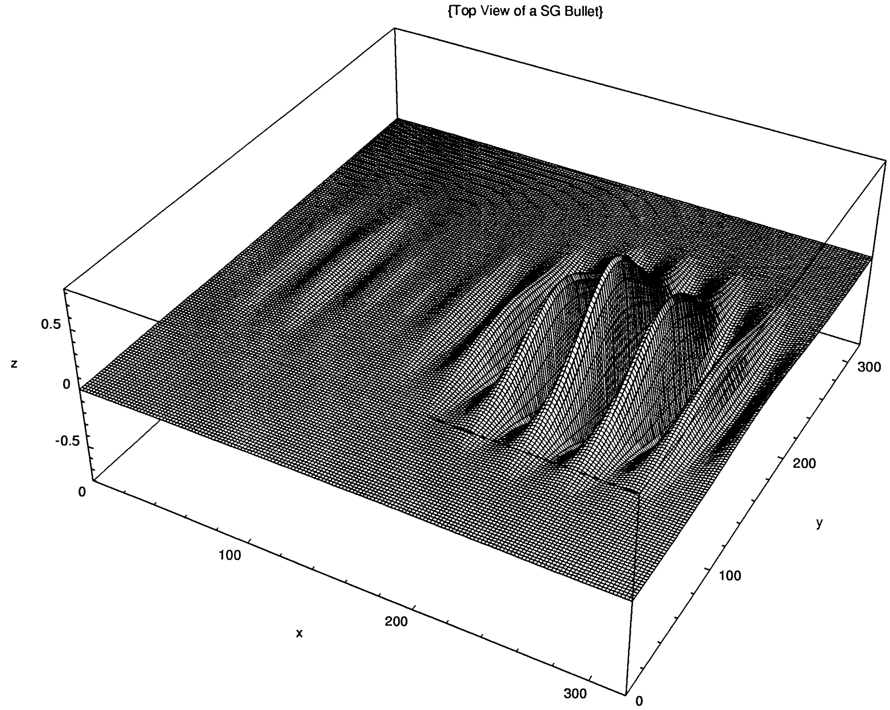

Fig. 12. Top view of SG light bullet in Fig. 6.

7. A theory of light bullets

In view of both the asymptotic and numerical results on the SG bullets, we reached the following understanding of their origin.

- Light bullets are manifestations of the robustness of *the complete NLS approximation* to Maxwell type nonlinear wave equation (NLW) in more than one space dimensions. They are dynamic objects with internal oscillation cycles, not single hump structures like EM bubbles. The complete NLS approximation refers to a modified NLS such that all resonances incurred in the slowly varying envelope approximation are removed.
- The asymptotic behavior as a consequence of the focusing and defocusing competition in the complete NLS with strong scale separation persists in NLW even when scale separation is significantly weakened or lost.
- Light bullets only exist over a long enough time period but not forever. Radiation of energy in two and higher dimensions eventually leads to their decay. However, their ability to maintain essentially the same shapes and long distance propagation with little diffraction makes them attractive for practical purposes [4].

Acknowledgements

This work was partially supported by NSF grant DMS-9625680 and by AFOSR through the Arizona Center for Math Sciences (ACMS). I would like to thank Prof. J. Moloney for introducing me to the light bullet problem, and for making available numerical results on the vector Maxwell system at ACMS prior to publication.

Much of the work was initiated during my sabbatical leave in the Spring semester of 1998. I would like to thank my hosts at various institutions for their warm hospitality and support, all the stimulating and helpful discussions, and the pleasant working environment. In particular, let me express my appreciation to Profs. D. McLaughlin and J. Shatah at Courant Institute; Prof. K. Kirchgässner and his research group at the Universität Stuttgart; Prof. T. Hoffmann-Ostenhof at ESI (Schrödinger Institute); Profs. T. Passot, C. Sulem, P.-L. Sulem at the Observatoire de Nice and Profs. R. Caflisch and S. Osher at UCLA.

I would also like to thank Dr. A. Nachman for his interest, Dr. L. Bergé and Dr. M. Keel for useful discussions. Last but not the least, I thank the anonymous referees for their constructive comments on the early manuscript.

Appendix A. Global well-posedness of the perturbed CNLS

Let us write the perturbed CNLS (5.3) as:

$$\frac{\epsilon^2}{4\omega^2}A_{TT} - \frac{\epsilon\nu}{c}A_{XT} - \Delta_{X,Y}A - \mathrm{i}A_T = f_\epsilon(|A|^2)A, \tag{A.1}$$

where $|f_\epsilon(|A|^2)| \le C\epsilon^{-2}$, uniformly in A for an ϵ independent constant C. The linear part of (A.1) is hyperbolic and we state its properties in Lemma A.1

Lemma A.1. *Let* $(A_0, A_1) \in H^2(R^2) \times H^1(R^2)$, *then the initial value problem*:

$$\frac{\epsilon^2}{4\omega^2}A_{TT} - \frac{\epsilon\nu}{c}A_{XT} - \Delta_{X,Y}A - \mathrm{i}A_T = 0, \tag{A.2}$$

$$A|_{T=0} = A_0, \qquad A_T|_{T=0} = A_1, \tag{A.3}$$

is well-posed in $C([0,\infty), H^2(R^2))$ *and satisfies the bound* (C *independent of* ϵ):

$$\|A\|_{H^2}(T) \le C(\|A_0\|_{H^2} + \epsilon\,\|A_1\|_{H^1}), \quad \forall\, T \ge 0. \tag{A.4}$$

Proof. It is easy to check that the second order terms in (A.2) form a hyperbolic operator with characteristic speeds -1, $\mu_\pm$, where

$$2\mu_\pm = -\left(1 - \frac{\epsilon^2}{4\omega^2}\right) \pm \sqrt{\left(1 - \frac{\epsilon^2}{4\omega^2}\right)^2 + \frac{\epsilon^2\nu^2}{c^2}},$$

and can be converted into $A_{TT} - \Delta_{X,Y}A$ via an orthogonal linear transform. Taking Fourier transform of (A.2) shows

$$\frac{\epsilon^2}{4\omega^2}\hat{A}_{TT} + \mathrm{i}\left(1 - \frac{\epsilon\nu\xi_1}{c}\right)\hat{A}_T + |\xi|^2\hat{A} = 0, \tag{A.5}$$

giving upon using the initial data:

$$\hat{A}(\xi, T) = \frac{1}{4}\omega^{-2}\epsilon^2\mathrm{i}\hat{A}_0\frac{(\lambda_2\exp\{\lambda_1 T\} - \lambda_1\exp\{\lambda_2 T\})}{\sqrt{(1-\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2\omega^{-2}}} + \frac{1}{4}\omega^{-2}\epsilon^2\mathrm{i}\hat{A}_1\frac{\exp\{\lambda_2 T\} - \exp\{\lambda_1 T\}}{\sqrt{(1-\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2\omega^{-2}}}, \tag{A.6}$$

where $\lambda_{1,2}$ are pure imaginary:

$$\lambda_{1,2} = 2\omega^2\mathrm{i}\epsilon^{-2}\left[-(1-\epsilon\nu\xi_1 c^{-1}) \pm \sqrt{(1-\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2\omega^{-2}}\right].$$

Let $r_0 \in (0,1)$, $r_0 \nu c^{-1} \le 1/4$, then if $\epsilon|\xi| \ge r_0$:

$$\sqrt{(1-\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2\omega^{-2}} \ge \epsilon|\xi|\omega^{-1} \ge r_0\omega^{-1}. \tag{A.7}$$

If $\epsilon|\xi| \le r_0$:

$$\sqrt{(1-\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2\omega^{-2}} \ge 1 - r_0\nu c^{-1}.$$

Hence the uniform lower bound $\min(r_0\omega^{-1}, 1-r_0\nu c^{-1})$ follows. If $|\epsilon\xi| \gg 1$, $\lambda_{1,2} \sim C(\omega,\nu,c)\epsilon^{-2}|\epsilon\xi|$, so:

$$|\lambda_{1,2}|/\sqrt{(1+\epsilon\nu\xi_1 c^{-1})^2 + \epsilon^2|\xi|^2 c^{-2}} \le C(\omega,\nu,c)\epsilon^{-2}, \tag{A.8}$$

while if $|\epsilon\xi|$ is bounded by a constant, $\lambda_{1,2} \sim C(\omega,\nu,c)\epsilon^{-2}$, (A.8) again follows with the lower bound on the denominator. The estimate (A.4) follows from (A.6)–(A.8).

It is standard to apply the contraction mapping theorem to establish a local solution. □

Lemma A.2. *For initial data* $(A_0, A_1) \in H^2(R^2) \times H^1(R^2)$, *there is a time* $T^* = T^*(\epsilon)$ *such that the initial value problem for* (A.1) *has a unique solution* $A \in C([0,T^*); H^2(R^2))$, $A_T \in C([0,T^*); H^1(R^2))$, *and* $A_{TT} \in C([0,T^*); L^2(R^2))$.

To extend the above local solution to a global one, we make use of the energy conservation and the saturating nonlinearity. The energy identity is

$$E_T \equiv \frac{\mathrm{d}}{\mathrm{d}T}\int \frac{\epsilon^2}{4\omega^2}|A_T|^2 + |\nabla A|^2 - F_\epsilon(|A|^2) = 0, \tag{A.9}$$

where $F_\epsilon(u) = \int_0^u f_\epsilon(s)\,\mathrm{d}s$. We verify (A.9):

$$\begin{aligned} E_T &= 2\,\mathrm{Re}\int \frac{\epsilon^2}{4\omega^2} A_{TT}A_T^* + A_{XT}A_X^* + A_{YT}A_Y^* - f_\epsilon(|A|^2)A_T^*A \\ &= 2\,\mathrm{Re}\int (\mathrm{i}A_T + \Delta A + f_\epsilon(|A|^2)A + \epsilon\nu c^{-1}A_{XT})A_T^* + A_{XT}A_X^* + A_{YT}A_Y^* - f_\epsilon(|A|^2)A_T^*A \\ &= \epsilon\nu c^{-1}\,\mathrm{Re}\int (|A_T|^2)_X = 0. \end{aligned} \tag{A.10}$$

The mass balance identity is

$$\frac{\mathrm{d}}{\mathrm{d}T}\left(\int |A|^2 - \frac{\epsilon^2}{2\omega^2}\,\mathrm{Im}\int A_T A^*\right) = 2\epsilon\nu c^{-1}\,\mathrm{Im}\int A_X A_T^*. \tag{A.11}$$

We note that the mixed derivative term A_{XT} does not change the energy identity yet does affect the conservation of mass. Without A_{XT}, (A.1) reduces to the model studied by Bergé and Colin [3], where uniform bounds on total mass independent of ϵ are derived. With A_{XT} the global well-posedness remains true, however the bounds below depend on ϵ.

Theorem A.1. *For initial data* $(A_0, A_1) \in H^2(R^2) \times H^1(R^2)$, *the initial value problem of* (A.1) *has a unique global solution* $A \in C([0,\infty); H^2(R^2))$, $A_T \in C([0,\infty); H^1(R^2))$, *and* $A_{TT} \in C([0,\infty); L^2(R^2))$.

Proof. Multiply (A.1) by A^*, integrate and take real part to get:

$$0 = \mathrm{Re}\left(\mathrm{i}\int A^* A_T + \int A^*\Delta A + \int |A|^2 f_\epsilon(|A|^2) - \frac{\epsilon^2}{4\omega^2}\int A^* A_{TT} + \frac{\epsilon\nu}{c}\int A^* A_{XT}\right)$$
$$= -\mathrm{Im}\int A^* A_T - \frac{1}{2}\int |\nabla A|^2 + \int |A|^2 f_\epsilon(|A|^2) - \epsilon\nu c^{-1}\mathrm{Re}\int A^*_X A_T - \frac{\epsilon^2}{4\omega^2}\,\mathrm{Re}\int (A^* A_T)_T - |A^*_T|^2$$
$$= -\mathrm{Im}\int A^* A_T - \frac{1}{2}\int |\nabla A|^2 + \int |A|^2 f_\epsilon(|A|^2) - \frac{\epsilon\nu}{c}\mathrm{Re}\int A^*_X A_T + \frac{\epsilon^2}{4\omega^2}\int |A^*_T|^2 - \frac{\epsilon^2}{8\omega^2}\frac{\mathrm{d}^2}{\mathrm{d}T^2}\int |A|^2.$$

It follows with Cauchy–Schwartz inequality and energy identity that

$$\frac{\epsilon^2}{8\omega^2}\frac{\mathrm{d}^2}{\mathrm{d}T^2}\int |A|^2 \le \frac{-1}{2}\int |\nabla A|^2 + \frac{\epsilon^2}{4\omega^2}\int |A^*_T|^2 + C\epsilon^{-2}\int |A|^2 + \frac{\epsilon^2}{2}\int |A_T|^2$$
$$+\frac{1}{2\epsilon^2}\int |A^*|^2 + \frac{1}{4}\int |A_X|^2 + \frac{\epsilon^2\nu^2}{c^2}\int |A_T|^2$$
$$\le -\frac{1}{4}\int |\nabla A|^2 + \left(\frac{1}{4\omega^2} + \frac{\nu^2}{c^2} + \frac{1}{2}\right)(4\omega^2)\left(E_0 + \int F_\epsilon(|A|^2) - \int |\nabla A|^2\right) + \left(C + \frac{1}{2}\right)\epsilon^{-2}\int |A|^2$$
$$= -\left(\frac{5}{4} + \left(\frac{\nu^2}{c^2} + \frac{1}{2}\right)(4\omega^2)\right)\int |\nabla A|^2 + \left(1 + \left(\frac{\nu^2}{c^2} + \frac{1}{2}\right)(4\omega^2)\right)E_0 + \left(C + \frac{1}{2}\right)\epsilon^{-2}\int |A|^2. \tag{A.12}$$

Integrating (A.12) once in time ($M(T) \equiv \int |A|^2$):

$$\frac{\epsilon^2}{8\omega^2}M_T \le \frac{\epsilon^2}{8\omega^2}M_T(0) + C_1 T + C_2\epsilon^{-2}\int_0^T M(s)\,\mathrm{d}s. \tag{A.13}$$

Multiplying M on both sides of (A.13) and integrating once more, we find

$$\frac{\epsilon^2}{16\omega^2}M^2(T) \le \frac{\epsilon^2}{16\omega^2}M^2(0) + C_3(T^2+1) + C_2\frac{\epsilon^{-2}}{2}\left(\int_0^T M(s)\,\mathrm{d}s\right)^2$$
$$\le \left((1+T)C_4 + C_2\frac{\epsilon^{-1}}{2}\int_0^T M(s)\,\mathrm{d}s\right)^2,$$

implying

$$\frac{\epsilon}{4\omega}M(t) \le C_4(1+T) + \frac{C_2\epsilon^{-1}}{2}\int_0^T M(s)\,\mathrm{d}s. \tag{A.14}$$

and by Gronwall inequality:

$$M(T) \le 4\omega\,\epsilon^{-1}C_4(1+T)\mathrm{e}^{2\omega C_2\epsilon^{-2}T}. \tag{A.15}$$

Energy identity then implies that $\|(\nabla A, A_T)\|_{L^2}$ are similarly bounded. It is standard to obtain bounds on higher derivatives. These bounds allow the global continuation of local solutions. The proof is complete. □

References

[1] N. Akhmediev, A. Ankiewicz, Does the nonlinear Schrödinger equation correctly describe beam propagation? Optics Lett. 18 (6) (1993) 411–413.

[2] L. Bergé, wave collapse in physics: principles and applications to light and plasma waves, Phys. Reports 303(5–6) (1998) 259–372.
[3] L. Bergé, T. Colin, A singular perturbation problem for an envelope equation in plasma physics, Physica D 84 (1995) 437–459.
[4] N. Bloembergen, M. Mlejnek, J. Moloney, E. Wright, Femtosecond pulses at the boundary of a nonlinear dielectric, preprint, 1998.
[5] T. Brabec, F. Krausz, Nonlinear optical pulse propagation in the single-cycle regime, Phys. Rev. Lett. 78 (17) (1997) 3282–3286.
[6] B. Crosignani, P. Di Porto, A. Yariv, Nonparaxial equation for linear and nonlinear optical propagation, Optics Lett. 22 (11) (1997) 778–780.
[7] G.H. Derrick, Comments on nonlinear wave equations as models for elementary particles, J. Math. Phys. 5 (1964) 1252–1254.
[8] P. Donnat, J. Rauch, Global solvability of the Maxwell–Bloch equations from nonlinear optics, Arch. Rat. Mech. Anal. 136 (1996) 291–303.
[9] J. Eilbeck, J. Gibbon, P. Caudrey, R. Bullough, J. Phys. A 6 (1973) 1337.
[10] M.E. Fermann, M.J. Andrejco, Y. Silberberg, A.M. Weiner, Generation of pulses shorter than 200 fs from a passively mode-locked Er fiber laser, Optics Lett. 18(1) (1993) 48–50.
[11] G. Fibich, G. Papanicolaou, Self-focusing in the perturbed and unperturbed nonlinear Schrödinger equation in critical dimension, SIAM J. Appl. Math, to appear.
[12] G. Fibich, G. Papanicolaou, A modulational method for self-focusing in the perturbed critical nonlinear Schrödinger equation, Phys. Lett. A 239 (1998) 167–173.
[13] S. Glasgow, J. Moloney, unpublished simulation, 1997.
[14] S. Glasgow, J. Moloney, Instability of electromagnetic bubbles to transverse perturbations: the transition to bullets and shocks in 2-D, unpublished notes, 1997.
[15] P. Goorjian, Y. Silberberg, Numerical simulations of light bullets using the full-vector time-dependent nonlinear Maxwell equations, J. Optical Soc. Am. B (Optical Physics) 14(11) (1997) 3253–3260.
[16] P.M. Goorjian, G.P. Agrawal, Computational modeling of ultrashort pulse propagation in semiconductor materials, Quantum Optoelectronics, 1997 Technical Digest Series, vol. 9, Postconference Edition, 1997, pp. 95–97.
[17] P. Hartman, Ordinary Differential Equations, Wiley, New York, 1973.
[18] C. Hile, Comparisons between Maxwell's equations and an extended nonlinear Schrödinger equation, Wave Motion 24 (1996) 1–12.
[19] J.M. Hyman, G. Mercado, Simulation of a quasilinear wave equation with lower order dispersion using a high resolution shock capturing method, unpublished notes, 1997.
[20] A.E. Kaplan, P. Shkolnikov, Electromagnetic "bubbles" and shock waves: unipolar, nonoscillating EM solitons, Phys. Rev. Lett. 75(12) (1995) 2316–2319.
[21] A.E. Kaplan, P. Shkolnikov, Subfemtosecond high-intensity unipolar electromagnetic solitons and shock Waves, J. Nonlinear Optical Phys. Mater. 4(4) (1993) 831–841.
[22] P. Kirrmann, G. Schneider, A. Mielke, The validity of modulation equations for the extended systems with cubic nonlinearities, Proc. Roy. Soc. Edin. 122A (1992) 85–91.
[23] A. Majda, R. Rosales, M. Schonbek, A canonical system of integral-differential equations arising in resonant nonlinear acoustics, Stud. Appl. Math 79 (1988) 205–262.
[24] V. Malkin, Dynamics of wave collapse in the critical case, Phys. Lett. A 151 (1990) 285–288.
[25] D. McLaughlin, J. Shatah, private communication, 1998.
[26] P. Lax, Hyperbolic systems of conservation laws and the mathematical theory of shock waves, SIAM Regional Conf. Ser. in Appl. Math 11 (1973).
[27] A. Newell, J. Moloney, Nonlinear Optics, Addison-Wesley, Redwood City, CA, 1991.
[28] T. Ozawa, K. Tsutaya, Y. Tsutsumi, Global existence and asymptotic behavior of solutions for the Klein–Gordon equations with quadratic nonlinearity in two space dimensions, Math. Z. 222 (1996) 341–362.
[29] R. Pantell, H. Puthoff, Fundamentals of Quantum Electronics, Wiley, New York, 1969.
[30] T. Passot, C. Sulem, P. Sulem, private communication, 1998.
[31] P. Roseneau, private communication, 1997.
[32] Y. Silberberg, Collapse of optical pulses, Optics Lett. 15(22) (1990) 1282–1284.
[33] J. Strikwerda, Finite Difference Schemes and PDE, Wadsworth and Brooks, Cole Adv. Books and Software, 1989.
[34] J. Xin, Shock formation in a quasilinear wave equation with lower order dispersion, unpublished notes, 1997.