Propagation of ultra-short optical pulses in nonlinear media

T. Schäfer [a] and G. E. Wayne [b]

[a] *Department of Mathematics*
University of North Carolina
Chapel Hill, NC 27599

[b] *Department of Mathematics and*
Center for BioDynamics
Boston University
Boston, MA 02215

Abstract

We derive a partial differential equation that approximates solutions of Maxwell's equations describing the propagation of ultra-short optical pulses in nonlinear media and which extends the prior analysis of Alterman and Rauch [1], [2]. We discuss (non-rigorously) conditions under which this approximation should be valid, but the main contributions of this paper are: (1) an emphasis on the fact that the model equation for short pulse propagation may depend on the details of the optical susceptibility in the wavelength regime under consideration, (2) a numerical comparison of solutions of this model equation with solutions of the full nonlinear partial differential equation, (3) a local well-posedness result for the model equation and (4) a proof that in contrast to the nonlinear Schrödinger equation which models slowing varying wavetrains this equation has no pulse solutions which propagate with fixed shape and speed.

Key words: Nonlinear guided waves, dynamics of nonlinear optical systems, fiber optics
PACS: 42.65.Tg, 42.65.Sf, 42.81

1 Introduction

Pulse propagation in optical fibers is usually modeled by the cubic nonlinear Schrödinger equation (NLSE) [3]. Hence, the NLSE forms the basis for optimizing existing fiber links and suggesting new fiber communication systems in attempts to achieve high bit-rate data transmission. Recently, much experimental progress has been made in creating ultra-short pulses that would allow very high data transmission in one channel [4]. However, for describing the propagation of these very narrow pulses, the validity of the NLSE as a slowly varying amplitude approximation of Maxwell's equations is questionable. The breakdown of the NLSE has been discussed, for instance, in the context of self-focusing of ultra-short pulses [5,6]. The reason for this breakdown is that the basic assumption that is made in the derivation of the NLSE as an approximation of Maxwell's equations is that the pulse's spectrum is localized around the carrier frequency [7]. This assumption is violated by short pulses.

One approach to describe the propagation of short pulses is to incorporate higher order terms in the cubic nonlinear Schrödinger equation, especially in order to account for Raman scattering [8,9]. A different way is to work in the Fourier domain [10] - but in this case the term arising from the nonlinear part of the polarization in Maxwell's equation will lead to convolution integrals that might be difficult to treat analytically and numerically [11].

Under some assumptions it is possible to derive directly from Maxwell's equations a "generic nonlinear envelope equation" [12], but this equation is formally an integral equation as it assumes the inversion of a differential operator. Still, it can be solved numerically very efficiently. Simultaneously, numerical schemes have been developed in order to compute the solution of the full Maxwell's equations [13,14].

Very recently, a new approach to study short pulses was developed by Alterman and Rauch, [1]. The basic idea is to make use of the fact that the pulse is broad in the Fourier domain. This approach leads to a different partial differential equation than the NLS. In the present work we follow this idea to study the propagation of very short pulses in Maxwell's equations. There are two significant differences between our work and that of [1]. The first is that whereas Alterman and Rauch assumed that there was no time-delay in the coefficients of their partial differential equation, we work with the experimentally determined optical susceptibility for silica. The presence of resonances in the susceptibility introduces additional time scales in the problem and as we show below these place significant restrictions on the applicability of the expansion method. At the end of section 2 we discuss the experimental conditions under which we expect our modulation equation to provide a good approximation to the true pulse evolution. The second difference is that since two time derivatives act on the nonlinear polarizability term in Maxwell's equation, the modulation equation we derive is quasi-linear as opposed to the semi-linear equation derived in [1]. As a consequence, the existence theory of the equation we derive is more complicated than that of [2] and section 4 is devoted to proving that our short pulse equation is locally well-posed.

One thing we do not prove in this paper are rigorous estimates relating the solutions of our short-pulse equation to the true solutions of Maxwell's equations. Our reason for this is that such estimates would be expected to hold only when the parameter ϵ describing the pulse length was extremely small. This would correspond to pulses whose Fourier spectrum was exceedingly broad. As we explain in the next section, our derivation of the short-pulse equation requires us to approximate the Fourier transform of the optical susceptibility by a polynomial in the wavelength. While we demonstrate that this approximation is accurate for infrared pulses of lengths accessible in current experiments, we do not expect it to remain valid as the pulse length tends to zero. Thus, while we believe that we could prove estimates relating the solutions of the short-pulse equation to the solutions of (8), similar to those proven in [2], such estimates would probably only be valid in a regime in which (8) was no longer an accurate approximation to Maxwell's equations in a silica fiber.

2 Derivation of the basic equation

2.1 The physical background

Although we believe that our results would be unchanged by considering more complicated and realistic geometries, in this paper we will limit ourselves to considering the propagation of linearly polarized light in a one-dimensional medium. In this case, if u represents the magnitude of the electric field, it satisfies

$$(\partial_x^2 - \frac{1}{c^2}\partial_t^2)u(x,t) = \partial_t^2 p, \tag{1}$$

where p is the polarization of the medium in response to the electric field.

The polarization can be split into two pieces, the linear, p_ℓ, and the nonlinear, $p_{n\ell}$ polarizability. Since the response of the medium to the electric field is not instantaneous, the polarization is a nonlocal (in time) function of the electric field. For the linear part of the polarizability we incorporate the retardation in the material response by writing

$$p_\ell(x,t) = \frac{1}{c^2}\int_{-\infty}^{\infty} \chi^{(1)}(t-\tau)u(x,\tau)d\tau \ , \tag{2}$$

where to enforce causality, the susceptibility must satisfy $\chi^{(1)}(\tau) = 0$ if $\tau < 0$.

If one assumes that the material can be modeled as a free atom interacting with an electromagnetic field, one can derive an expression for the Fourier transform of $\chi^{(1)}$ of the form [15]

$$\hat{\chi}^{(1)}(\omega) = c_\chi \sum_n |\mu_n|^2 \left\{ \frac{2\omega_{na}}{(\omega_{na}^2 - \omega^2) + \gamma_{na}^2 - 2i\gamma_{na}\omega} \right\} . \tag{3}$$

Here, the ω_{na}'s are the resonant frequencies of the medium and the γ_{na}'s are small, phenomenological damping coefficients added to insure that the susceptibility remains finite even at the resonant frequency.

Typically, for silica fibers and for light in the visible to mid-infrared range there are three resonances of importance which occur at wavelengths of $\lambda = 0.068\ldots$ μm, $\lambda = 0.116\ldots$ μm and $\lambda = 9.896\ldots$ μm.

If one restricts attention to wavelengths between 0.25 and 3.5 μm, approximate values for the various constants in (3) can be obtained by fitting experimental data for light propagation in silica [16], and over this range of wavelengths one can approximate $\hat{\chi}^{(1)}$ by

$$\hat{\chi}^{(1)}(\lambda) = \frac{0.696\lambda^2}{\lambda^2 - (0.0684)^2} + \frac{0.4079\lambda^2}{\lambda^2 - (0.116)^2} + \frac{0.8974\lambda^2}{\lambda^2 - (9.896)^2} , \tag{4}$$

where λ is the wavelength expressed in microns.

In the present paper we study the propagation of light in the infrared range with wavelengths of 1600-3000 nm. In this range (4) is well approximated by

$$\hat{\chi}^{(1)}(\lambda) \approx \hat{\chi}_0^{(1)} - \hat{\chi}_2^{(1)}\lambda^2 , \tag{5}$$

where we choose the constants $\hat{\chi}_j^{(1)}$ to have values

$$\hat{\chi}_0^{(1)} = 1.1104 , \quad \hat{\chi}_2^{(1)} = 0.01063 . \tag{6}$$

Comparing (4) and (5) in Figure 1 we see that over the wavelength range under discussion here, (5) approximates the nonconstant part of (4) with an error of less than 1%, and we will use the approximation (5) for the susceptibility throughout the remainder of this section of the paper. We note for comparison with the numerical section that follows that this approximation is the same as one would obtain by considering a susceptibility function with a single resonance at a wavelength much larger than the wavelengths of interest and then expanding the susceptibility in λ about $\lambda = 0$.

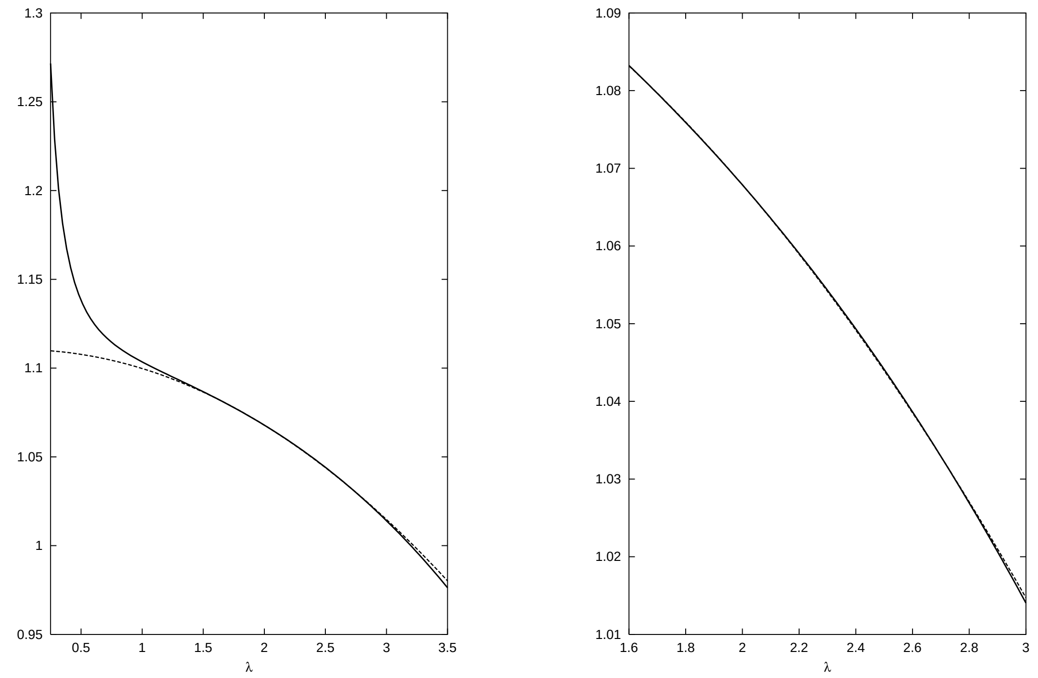

Fig. 1. The approximation of the optical susceptibility by (5). The figure on the left side shows the approximation over the range of wavelength from 0.25 to 3.5 μm. As it can be seen from the figure on the right side, the approximation (5) is very good if the spectrum of the pulse is between 1.6 and 3 μm.

While the exact form of $\chi^{(1)}$ is not important for our results it is important that we can approximate $\hat{\chi}^{(1)}$ by a polynomial in λ. In particular, we do not see how to construct such an approximation for λ in the blue-green wavelength range and thus we expect that short pulses in that wavelength region will be governed by a different equation than the one we derive below.

We next proceed to non-dimensionalize the equation. The natural time and length scales in the model are determined by the principal resonance which in the regime under consideration is the resonance at 9.896 μm and the phase velocity of light in the medium (again in our wavelength regime) is about $c_{\text{eff}} = 2.06\times 10^8$ m/s. Thus, we rescale all lengths by $L = 10^{-5}$ m and $T = 10^{-14}$ s. Note that in these units the speed of light (in vacuum) is $c = .299$, while the angular frequency of the principal resonance is 1.2 and the wavelength of the resonance is .986.

In these units, if we use approximation (5) for the susceptibility, the linear part of equation (1) (in Fourier transformed variables) becomes

$$\partial_x^2 \hat{u} + \frac{1+\hat{\chi}_0^{(1)}}{c^2}\omega^2 \hat{u} - (2\pi)^2 \hat{\chi}_2^{(1)} \hat{u} = 0 \ . \tag{7}$$

One must now consider the nonlinear term in the polarizability. In general this will also include a time delay as well, however, on the basis of our preliminary calculations we expect that only the instantaneous contribution will affect the propagation of short, small amplitude pulses to the order of approximation we consider. Thus, we model $p_{n\ell} = \chi^{(3)} u^3$, with $\chi^{(3)}$ a constant. We note that for silica fibers there is no quadratic term in the susceptibility [15].

Thus, we will study equations of the form

$$\partial_x^2 u = \frac{1}{c_1^2} \partial_t^2 u + \frac{1}{c_2^2} u + \chi^{(3)} \partial_t^2 u^3 \ , \tag{8}$$

which is an accurate approximation of the Maxwell's equation (1), in the 1600-3000 nm wavelength regime.

We begin our analysis of the solutions of this equation by noting that if we consider only the principal part of the linearized equation, i.e. $\partial_x^2 u = \frac{1}{c_1^2} \partial_t^2 u$, the solution splits into two wave packets, one moving to the left and one moving to the right, both with speed c_1 that can be set to 1 after renormalization. In general, the nonlinear term will generate interactions between the left and right moving wave trains. However, because the pulses are short, we expect that the left and right moving waves pass through each other so quickly that the effects of the interaction would only appear in a higher order approximation than the one we are making here. (This fact has been proven rigorously in some related contexts such as the propagation of long-waves on a fluid surface [17].) For this reason we will concentrate on a right-moving wave packet and ignore the left-moving part of the solution in what follows.

To incorporate the effects of the nonlinear and dispersive terms in the equation we make a multiple scales *ansatz* of the form

$$u(x,t) = \epsilon A_0(\phi, x_1, x_2, ...) + \epsilon^2 A_1(\phi, x_1, x_2, ...) + ... \tag{9}$$

with

$$\phi = \frac{t-x}{\epsilon}, \qquad x_n = \epsilon^n x. \tag{10}$$

Note that when $x = 0$, we have $u(x=0,t) = \epsilon A_0(\frac{c_1 t}{\epsilon}) + \epsilon^2 A_1(\frac{c_1 t}{\epsilon})$ so that this does represent a short pulse if ϵ is small.

Inserting (10) into (8) we find that all terms of $\mathcal{O}(\frac{1}{\epsilon})$ cancel because of our choice of the form of the multiple scale *ansatz* and there are no terms of $\mathcal{O}(\epsilon^0)$.

In order to cancel the terms of order ϵ, the envelope equation A_0 must satisfy

$$-2\partial_{x_1}\partial_\phi A_0 = \frac{1}{c_2^2}A_0 + \chi^{(3)}\partial_\phi^2 A_0^3 \ . \tag{11}$$

This is our short pulse equation, and in the next section we will present numerical computations which show that it does indeed do a good job of approximating the behavior of solutions of (1).

We close this section by briefly recapping the conditions under which we expect (11) to accurately describe the true evolution. The key requirement in this derivation was that the linear susceptibility could be approximated by a polynomial (in λ) as in (5). We have shown that this is the case for the experimentally determined susceptibility of silica in the wavelength range from 1600-3000 nm. This approximation should be contrasted with the standard derivation of the Nonlinear Schrödinger equation in nonlinear optics which involves expanding the susceptibility in the frequency – this would correspond to an expansion in *inverse* powers of the wavelength and illustrates how different the regimes governed by the two equations are.

Two questions then arise:

(1) Is such an approximation valid for light in the visible range?
(2) Given that we are considering short pulses, their Fourier spectrum will be widely spread in frequency and we must ask whether or not the frequency range over which our approximation of the susceptibility is accurate is wide enough to encompass the frequency range spanned by the pulse.

In the first case, we can only say that we have not found such an approximation yet. Indeed, on physical grounds we expect that in the visible range the two resonances at $\lambda = 0.116\mu$m and 0.068μm would play an increasingly important role and these will introduce additional time scales into the problem which may necessitate a more complicated approximation procedure.

For the second question we can give a more concrete answer. One can currently construct experimental pulses whose lengths are between 2 and 10 cycles of the central frequency [4]. If we consider a pulse of length six cycles which a central wavelength of 2100 nm one finds that the Fourier transform of this pulse falls off to less than 10% of its maximum amplitude (which means that the power spectrum of the pulse falls to less than 1% of its maximum) outside the frequency range $\omega_{min} = 2.1$ and $\omega_{max} = 3.91$ (in our non-dimensionalized units.) These correspond to physical wavelengths of 2900 and 1600 nm respectively, and thus almost all of the energy of the pulse is concentrated in the wavelength interval where our approximation of $\chi^{(1)}$ is valid. Thus, (11) should provide an accurate approximation for the propagation of such pulses.

If on the other hand, we consider a 3-cycle pulse, again with a central wavelength of 2100 nm, we find that the Fourier transform of this pulse is much broader – extending from about 1300 nm to 5300 nm. If one compares (5) to (4) over this extended frequency range, the approximation is not very good – one gets errors of about 30% in the non-constant part of $\chi^{(1)}$. If, however, one replaces (5) by a new approximation

$$\hat{\chi}^{(1)}(\lambda) \approx 1.1079 - 0.12\lambda^2 \tag{12}$$

one finds that the error decreases to only about 10% over the expanded range. If we replace (5) by (12) in Maxwell's equation we can repeat the multiple scales calculation and we find that the only change in the precise values of the constants in the short pulse equation (11). Thus, (11) should still be able to approximation the evolution of such pulses, albeit with a larger error in the dispersion than for the 6-cycle pulse.

The other source of error in our approximation comes from the higher order terms in ϵ which we ignored in (11). To estimate their importance we need to know how large ϵ is. Roughly speaking ϵ is determined by the "shortness" of the pulse relative to the time scale determined by the resonance. If we again consider pulses with central wavelength of 2100 nm and a length of 3-4 cycles, we find ϵ in the range of 0.2 to 0.25. Thus, ϵ is not very small, and for currently accessible experimental regimes this probably represents the largest source of error in the use of (11) to approximate solutions of Maxwell's equations.

3 Comparison to numerics

Following the considerations of the previous section, we limit our considerations to the case where we have one resonance in the susceptibility at a wavelength that is much larger than the wavelengths covered by the pulse's spectrum. In the code, we use

$$\hat{\chi}^{(1)}(\omega) = \frac{\chi_\alpha}{\omega_r^2 - \omega^2}\left(\mathcal{H}(\omega - \omega_c) + \mathcal{H}(-\omega_c - \omega)\right) \tag{13}$$

with $\mathcal{H}(x) = 1$ for $x \leq 0$ and $\mathcal{H}(x) = 0$ elsewhere. This accounts for the fact that the pulse cannot propagate in the whole frequency spectrum and that there is a cut-off before its spectrum hits the resonance frequency ω_r. χ_α is a constant representing the strength of the resonance. As an initial condition for eq. (1) we take a very short pulse with a small amplitude

$$u(x = 0, t) = \epsilon u_0(t/\epsilon)\cos(\omega_0 t). \tag{14}$$

Here, it is necessary to introduce a center frequency ω_0 of the pulse to keep its spectrum away from the resonance ω_r. Ideally, we have $\omega_r \ll \omega_c \ll \omega_0$ but because of $u_0(t/\epsilon)$ the pulse is still broad in Fourier domain. In the case $\chi^{(3)} = 0$ the system is linear and, as it can be seen from straightforward calculations, it has forwards and backwards propagating wave solutions. Here, we look only at the forward propagating wave and therefore we choose the initial condition $u_x(x=0,t)$ to be

$$u_x(x=0,t) = \frac{1}{2\pi}\int i\beta(\omega)\hat{u}_0(\omega)\exp(-i\omega t)d\omega \tag{15}$$

with

$$\hat{u}_0(\omega) = \int u_0(t/\epsilon)\exp(i\omega t)dt, \qquad \beta(\omega) = \omega\sqrt{1+\hat{\chi}^{(1)}(\omega)}. \tag{16}$$

Figure 2 shows the initial condition and the susceptibility for the parameters that were used in the numerical simulations of (1). In order to check the

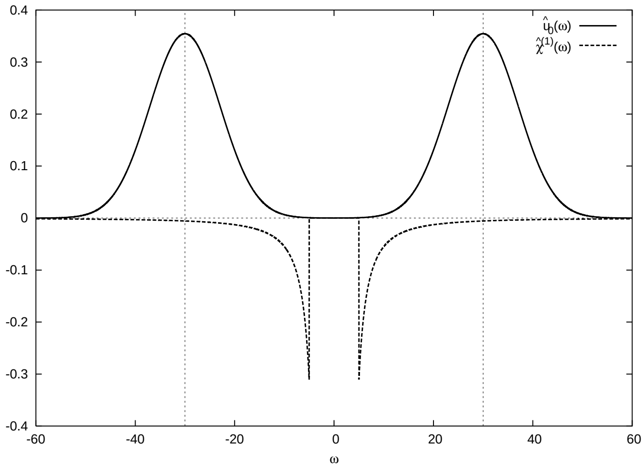

Fig. 2. Graph of the Fourier transform of the initial distribution $\hat{u}_0(\omega)$ and linear susceptibility $\hat{\chi}^{(1)}(\omega)$. Here the parameters are $\epsilon = 0.2$, $\omega_r = 3$, $\omega_c = 5$, $\omega_0 = 30$, $\chi_\alpha = 5$.

validity of the above approach we have solved numerically (1) and (11) by standard methods. As we can see from (9,10), the main linear effect is a shift of the initial distribution on the t-axis with the speed ± 1. Dispersion and nonlinearity change the pulse shape slowly. In the numerical simulations, we consider real initial data of the form

$$u(x=0,t) = \epsilon u_0(t/\epsilon)\cos(\omega_0 t).$$

and choose the initial condition in a way that we will obtain a shift of the initial profile towards positive t, hence here we look at the case

$$\phi = \frac{t-x}{\epsilon}.$$

The transport of the pulse can be seen from Figure 3. Now, we want to compare

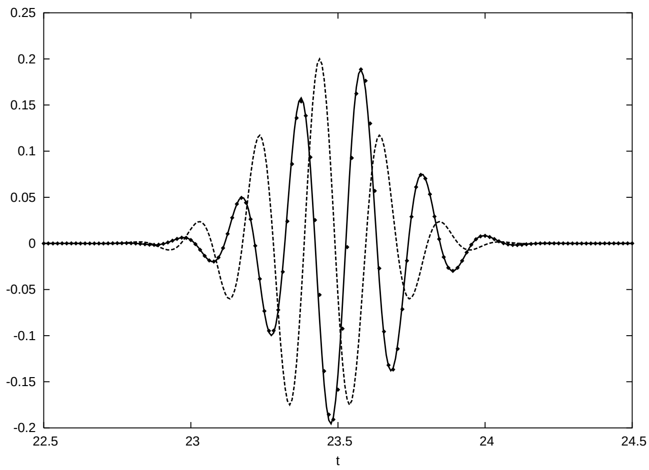

Fig. 3. Transport of the wave front. The solid line presents the pulse evolution of the linear problem A_{lin}, i.e. $\chi^{(3)} = 0$ in (1) after a propagation to $x_e \approx 23.4$. Diamonds present the solution A_m of the corresponding nonlinear problem with $\chi^{(3)} = 0.016$. The dashed line shows the initial pulse.

the simulations of (1) with the predictions of the equation (11). But as we can see from Figure 3, the influence of the nonlinearity on the pulse shape is small in comparison to the oscillations of the whole signal. Therefore, we carry out the comparison in the following way: We take A_{lin}, the solution of the linear part of (1), i.e. we solve this equation with $\chi^{(3)} = 0$ and take the difference of A_m and A_{lin}, where A_m is the solution of (1) with $\chi^{(3)} > 0$ and compare this difference to $A_s - A_{lin}$ where A_s is the solution of (11) with the same $\chi^{(3)} > 0$ that was used in order to obtain A_m. From Figure 4 we see that the equation (11) provides a very good approximation of the original equation (1). Under the condition discussed above, (11) plays the same role as the nonlinear Schrödinger equation does for broad pulses: Both equations, each in its range of validity, are a considerable simplification of Maxwell's equations - but still keep the important properties of that system.

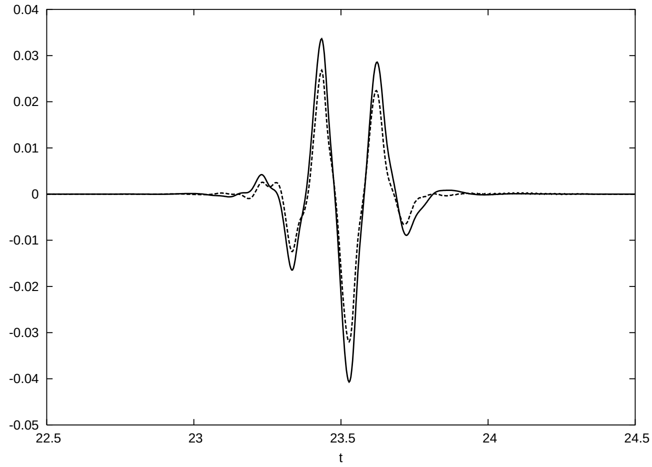

Fig. 4. Comparison of the solution of the approximate equation A_s to the solution of full Maxwell's equation A_m. The figure shows $A_s - A_{lin}$ (line) and $A_m - A_{lin}$ (dashes) where A_{lin} is the solution of the corresponding linear problem. Here, we used the parameters presented in Figure 2 and Figure 3.

4 Existence of solutions

In this section we prove that the Cauchy problem for (11) is well-posed. Consider

$$-\partial_\phi \partial_x A_0 = \alpha \; A_0 + \frac{\chi^{(3)}}{2} \partial_\phi^2 A_0^3, \;\; A_0|_{x=0} = \mathcal{A}_0 \tag{17}$$

We then have:

Theorem 4.1 *If $\mathcal{A}_0 \in H^s$ with $s \geq 2$ then there exists $X_0 > 0$ such that (17) has a unique solution $A_0 \in C^0([0, X_0]; H^s)$.*

The proof of this theorem is somewhat complicated by the combination of the rather unusual nature of the linear part of the equation and the quasi-linear nonlinearity. We begin the proof by simplifying (17). Let $S(x)$ be the semigroup associated with the linear equation

$$-\partial_\phi \partial_x A = \alpha A \;, \;\; A|_{x=0} = \mathcal{A}_0 \;. \tag{18}$$

As noted in [1],

$$\widehat{S(x)A}(k) = e^{-i\frac{\alpha}{k}x} \hat{\mathcal{A}}_0(k) \;. \tag{19}$$

Here the "hat" denotes Fourier transform with respect to ϕ. Note that S defines a strongly continuous (and norm-preserving) semigroup on H^s for all $s \geq 0$. Defining $A_0 = S(x)w$, we see that (17) is equivalent to

$$
\begin{aligned}
\partial_x w &= (S(x))^{-1}\{[S(x)w]^2 S(x)\partial_\phi w\} \\
w|_{x=0} &= w_0 = \mathcal{A}_0
\end{aligned} \tag{20}
$$

Here, we have assumed without loss of generality that $3\chi^{(3)} = 2$. Thus, solving the initial value problem for (20) yields a solution for (17). For (20) we have actually a slightly stronger result than Theorem 4.1, namely

Theorem 4.2 *If $w_0 \in H^s$, $s \geq 2$, then there exists $X_0 > 0$ such that (20) has a unique solution $w \in C^0([0, X_0]; H^s) \cap C^1([0, X_0]; H^{s-1})$. Furthermore, w depends continuously on w_0.*

Note that Theorem 4.1 follows immediately from Theorem 4.2, so the remainer of this section is devoted to proving this latter result. Our proof of this result is modelled on Kato's method [18] for solving quasi-linear PDE's in that we construct the solution of (20) as the limit of a sequence of functions $\{w^n\}$ which solve *linear* equations, though our manner of treating the linear equations differs from Kato's. More precisely, define $w^{(0)}(\phi, x) = w_0(\phi)$, and take $w^{(n)}$, for $n \geq 1$ to be the solution of the initial value problem

$$
\begin{aligned}
\partial_x w^{(n)} &= (S(x))^{-1}\{[S(x)w^{(n-1)}]^2 S(x)\partial_\phi w^{(n)}\} \\
w^{(n)}|_{x=0} &= w_0 = \mathcal{A}_0
\end{aligned} \tag{21}
$$

For the initial value problem (21) we prove the following well-posedness result

Theorem 4.3 *Suppose that $w_0 \in H^s$ with $s \geq 2$, and that $C_0 > \|w_0\|_{H^s}$. There exists $X_0 > 0$ and $C_1 > 0$ such that if $w^{(n-1)} \in C^0([0, X_0]; H^s) \cap C^1([0, X_0]; H^{s-1})$ satisfies $w^{(n-1)}|_{x=0} = w_0$ and*

$$
\|w^{(n-1)}\|_{C^0([0,X_0];H^s)} \leq C_0 \quad \text{and} \quad \|w^{(n-1)}\|_{C^1([0,X_0];H^{s-1})} \leq C_1 \ .
$$

then (21) has a unique solution $w^{(n)} \in C^0([0, X_0]; H^s) \cap C^1([0, X_0]; H^{s-1})$ which satisfies

$$
\|w^{(n)}\|_{C^0([0,X_0];H^s)} \leq C_0 \quad \text{and} \quad \|w^{(n)}\|_{C^1([0,X_0];H^{s-1})} \leq C_1 \ .
$$

Furthermore $\|w^{(n)}(\cdot, x)\|_{H^s}$ depends continuously on w_0.

Proof (of Proposition 4.3) We begin by rewriting (21) as

$$
\begin{aligned}
\partial_x w^{(n)} &= V^{(n-1)}\partial_\phi w^{(n)} + L^{(n-1)} w^{(n)} \\
w^{(n)}|_{x=0} &= w_0 \ ,
\end{aligned} \tag{22}
$$

where $V^{(n-1)} \equiv [S(x)w^{(n-1)}]^2 \in C^0([0, X_0]; H^s)$ and
$L^{(n-1)}w \equiv (S(x))^{-1}\{V^{(n-1)}S(x)\partial_\phi w\} - V^{(n-1)}\partial_\phi w\}$.

Due to its commutator structure, $L^{(n-1)}$ has a smoothing effect, detailed in the following:

Lemma 4.4 *If $V \in H^s$ with $s \geq 2$, there exists a $C > 0$ such that for $0 \leq x \leq X_0$*

$$\|Lw\|_{H^s} \leq C\|V\|^2_{H^s}\|w\|_{H^s}.$$

The proof of this lemma is a somewhat lengthy computation and so we relegate it to an appendix.

Now consider the transport equation

$$\partial_x W = V\partial_\phi W \ , \ W|_{x=\xi} = W_0 \ . \tag{23}$$

Using the method of characteristics, one can construct the solutions of this equation explicitly in terms of the solutions of the nonautonomous ODE:

$$\dot{\phi} = V(x,\phi) \tag{24}$$

Writing the solution in terms of the solution operator $\Sigma(x,\xi)$, we see immediately that $W \equiv \Sigma(x;\xi)W_0 \in C^0([0,X_0];H^s) \cap C^1([0,X_0];H^{s-1})$. Furthermore elementary estimates of the solution yield:

Lemma 4.5 *There exist C^0_Σ, C^1_Σ and X_{00}, positive constants, depending only on $\|V\|_{C^0([0,X_0];H^s)}$ such that for all $0 \leq \xi \leq x \leq X_{00}$ one has*

$$\begin{aligned} \|\Sigma(x;\xi)W_0\|_{H^s} &\leq C^0_\Sigma\|W_0\|_{H^s} \\ \|\partial_x\Sigma(x;\xi)W_0\|_{H^{s-1}} &\leq C^1_\Sigma\|W_0\|_{H^s} \ . \end{aligned}$$

Remark 4.6 *These estimates imply immediately that the solution of (23) depends continuously on the initial conditions. Furthermore, the form of the solution given by the method of characteristics, and the fact that the solutions of an ODE depend continuously on the vectorfield imply that Σ also depends continuously on V. That is, if $\tilde{\Sigma}$ is the solution operator for*

$$\partial_x\tilde{W} = \tilde{V}\partial_\phi\tilde{W} \ , \ \tilde{W}|_{X=\xi} = W_0 \ . \tag{25}$$

then there exists a constant $\tilde{C}$ such that for $0 \leq \xi \leq s \leq X_{00}$ one has

$$\begin{aligned} &\|(\Sigma(x,\xi) - \tilde{\Sigma}(x,\xi))W_0\|_{H^s} \\ &\leq \tilde{C}|x-\xi|e^{(C|x-\xi|)}\|V - \tilde{V}\|_{C^0([0,X_0];H^s)}\|W_0\|_{H^s} \ . \end{aligned} \tag{26}$$

Now consider (21). Using the solution operator $\Sigma^{(n-1)}$ for (23) (with $V = V^{(n-1)}$) we can rewrite it as:

$$w^{(n)}(x) = \Sigma^{(n-1)}(x,0)w_0 + \int_0^x \Sigma^{(n-1)}(x,\xi)L^{(n-1)}w^{(n)}(\xi)d\xi \ . \tag{27}$$

Given the estimates on $\Sigma^{(n-1)}$ and $L^{(n-1)}$ proven in Lemmas 4.5 and 4.4, a solution $w^{(n)}$ with the properties claimed in Proposition 4.3 follows immediately from a standard contraction mapping argument. $\square$

We now turn to the proof of Proposition 4.2. By the Arzela-Ascoli theorem we can extract from $\{w^{(n)}\}$ a subsequence $\{w^{(n_j)}\}$ converging in $C^0([0,X_0];H^{s-1})$. Let $w^* = \lim_{n_j\to\infty} w^{(n_j)}$. By the continuity properties of $L^{(n)}$ and $\Sigma^{(n)}$, we see that w satisfies

$$w^*(x) = \Sigma^*(x,0)w_0 + \int_0^x \Sigma^*(x,\xi)L^*w^*(\xi)d\xi \ , \tag{28}$$

where Σ^* is the solution operator of (23) with $V = [S(x)w^*]^2$ and L^* is the commutator whose properties were studied in Lemma 4.4, also with $V = [S(x)w^*]^2$. Note that from the method of characteristics it is easy to see that if $V^* \in C^0([0,X_0];H^{s-1})$ the solution of (23), $\Sigma^*(x,0)w_0 \in C^1([0,X_0];H^{s-1})$. This, combined with the fact that w^* satisfies (28) immediately implies that $w^* \in C^1([0,X_0];H^{s-1})$.

Thus, the only points that remain to be proven in Proposition 4.2 are that $w^* \in C^0([0,X_0];H^s)$ and that it is the unique solution of (20). Both of these results follow from *a priori* energy-type estimates. We first note the following estimate:

Lemma 4.7 *Suppose that $w \in C^1([0,X_0];H^s)$, $s \geq 2$, is a solution of (20). Then for $0 < x < X_0$,*

$$|\frac{1}{2}\partial_x(\partial_\phi^s w, \partial_\phi^s w)_{L^2}| \leq C\|w\|_{H^s}^4 \ .$$

Note that with this estimate it is standard to show that the solution, w^* of (20) constructed above is in $C^0([0,X_0];H^s)$, for some $X_0 > 0$ if the initial condition w_0 is in H^s. To prove the lemma we simply note that

$$\frac{1}{2}\partial_x(\partial_\phi^2 w, \partial_\phi^2 w)_{L^2} = (\partial_\phi^s w, \partial_\phi^s (S(x))^{-1}\{[S(x)w]^2\partial_\phi(S(x)w)\})_{L^2} \ .$$

Using the product rule to evaluate the derivatives in the right hand side of this inner product we see that

$$\frac{1}{2}\partial_x(\partial_\phi^s w, \partial_\phi^s w)_{L^2} = (S(x)\partial_\phi^s w, \{[S(x)w]^2\partial_\phi^{s+1}(S(x)w))_{L^2} + \mathcal{R}$$

where $\mathcal{R}$ consists of terms involving only derivatives of w of order s or less. The latter can all clearly be bounded by $C\|w\|_{H^s}^4$. For the remaining term we rewrite

$$\begin{aligned}(S(x)\partial_\phi^s w, \{[S(x)w]^2\partial_\phi^{s+1}(S(x)w))_{L^2} &= \frac{1}{2}([S(x)w]^2, \partial_\phi[\partial_\phi^s(S(x)w)]^2)_{L_2} \\ &= -((S(x)w)\partial_\phi(S(x)w), [\partial_\phi^s(S(x)w)]^2)_{L^2} \\ &\leq C\|w\|_{H^s}^4 \end{aligned} \tag{29}$$

This completes the proof of Lemma 4.7.

Uniqueness of the solution follows from a similar estimate, namely.

Lemma 4.8 *Suppose that w and $\tilde{w}$ are two solutions of (20) in $C^1([0, X_0]; H^{s-1}) \cap C^0([0, X_0]; H^s)$. Then there exists a constant C_u, depending on $\|w\|_{C^1([0,X_0];H^{s-1})} + \|w\|_{C^0([0,X_0];H^s)}$ and $\|\tilde{w}\|_{C^1([0,X_0];H^{s-1})} + \|\tilde{w}\|_{C^0([0,X_0];H^s)}$ such that*

$$\frac{1}{2}\partial_x\|\partial_\phi^{s-1}(w - \tilde{w})\|_{L^2} \leq C_u\|w - \tilde{w}\|_{H^{s-1}} \ .$$

The proof of this lemma is very similar to that of Lemma 4.7 and thus we leave it as an exercise to the reader.

5 Nonexistence of pulse solutions

In this section we prove that the equation (17) does not have real valued, smooth, pulse solutions that are stationary in a moving frame. More precisely we prove:

Proposition 5.1 *There are no solutions of (17) of the form $A(x, \phi) = \psi(\phi + \gamma x)$ with $\psi \in C^2 \cap H^2$.*

Suppose that $A(x, \phi) = \psi(\phi + \gamma x)$. Here we write A instead of A_0 and x instead of x_1. Then ψ satisfies

$$-\gamma\psi'' = \alpha\ \psi + (\psi^3)''. \tag{30}$$

Here, we have set without loss of generality $\chi^{(3)} = 1$. First note that we can assume that $\alpha \neq 0$ since if $\alpha = 0$ (30) implies that $\psi^2 = \gamma$, that is $\psi \equiv$ const. and the conditions at $\pm\infty$ imply that $\psi \equiv 0$. If $\alpha \neq 0$ we can also assume that $\gamma \neq 0$ since otherwise, ψ satisfies

$$-3\psi\psi'' = \alpha + 6(\psi')^2 \tag{31}$$

at any point where $\psi \neq 0$. But since $\psi(\xi) \to 0$ and $\psi'(\xi) \to 0$ as $\xi \to \infty$, this implies $\alpha = 0$ which is a contradiction.

Let $w = \psi$ and $v = \psi'$. Then if $\gamma + 3\psi^2 \neq 0$, the above equation is equivalent to

$$w' = v, \qquad v' = \left(\frac{-\alpha}{\gamma + 3w^2}\right) w - \left(\frac{6}{\gamma + 3w^2}\right) v^2 w. \tag{32}$$

Suppose first that α and γ are of the same sign. In this case, the linearization $(\tilde{v}, \tilde{w})$ of (32) at the origin is given by

$$\tilde{w}' = \tilde{v}, \qquad \tilde{v}' = -\frac{\alpha}{\gamma}\tilde{w} \tag{33}$$

and therefore (32) cannot have solutions that approach zero as $\xi \to \pm\infty$.[1] Next assume that α and γ have different signs. First we consider the case with $\alpha > 0$ and $\gamma < 0$. If $\psi \to 0$ as $\xi \to \infty$ and $\psi(\xi) > 0$ for large ξ, then ψ must have a positive local maximum. If at that maximum $\gamma + 3\psi^2 \leq 0$, then

$$-(\gamma + 3\psi^2)\psi'' = \alpha\psi. \tag{34}$$

But this is a contradiction since the left hand side of this equation is less or equal than zero while the right hand side is positive.

Now suppose that at some point we have $\gamma + 3\psi^2 > 0$. Then, as $\psi \in C^2$ is smooth and $\gamma < 0$ and $\psi(\xi) \to 0$ for $\xi \to \infty$, there exists a ξ_0 where $\gamma + 3\psi^2(\xi_0) = 0$ implying

$$\alpha\psi(\xi_0) + 6\psi(\xi_0)(\psi'(\xi_0)^2) = 0 \tag{35}$$

which is impossible if $\alpha > 0$. Note that in the case where $\psi(\xi) < 0$ for large ξ, ψ must have a negative local minimum and the same argument will rule out

[1] Of course, incorporating the nonlinear terms in (32) may change the phase portrait from a center as in the linearized equation to either an inward or outward spiral, but then the solution will fail to converge to zero either as $\xi \to -\infty$ or as $\xi \to \infty$.

the existence of a solitary wave. If we assume $\alpha < 0$ and $\gamma > 0$ we can apply a similar argument. $\square$

Given the important role that solitons have played in the applications of the NLS to fiber optics we want to point out several questions that the non-existence of solitary wave solutions for the pulse equation raises. It has recently been proven that Maxwell's equations do have true traveling pulse solutions of the sort represented by the NLS soliton [19]. However, these solutions represent very "slow" modulations of an underlying carrier wave and hence they are very far from the physical regime we are considering here. We see at least two possibilities:

(1) Maxwell's equations do not have traveling wave solutions in the short pulse regime, as suggested by the nonexistence result for the pulse equation. If this is the case one should investigate how rapidly an initial pulse looses its pulse-like shape since if this occurs too fast, it will be difficult to use these very short signals in communications.
(2) As is the case with the NLS approximation, we only expect (11) to approximate the true behavior of Maxwell's equations for a (long but) finite time. Since the nonexistence result of the present section gives no indication of how long it takes the pulse to break down, it may be that the breakdown occurs only after the equation has ceased to accurately approximate Maxwell's equation and that Maxwell's equation does have traveling pulse solutions even in the short pulse regime.

Since the existence of traveling wave solutions requires a very delicate balance between dispersion and nonlinearity and since we see no reason that this balance will hold in the short pulse regime we favor the first of these two scenarios, but at the moment we have no way of proving which is correct and that remains a matter for future research.

6 Conclusion

We have derived a new nonlinear wave equation directly from Maxwell's equation. This equation describes the evolution of a short pulse in nonlinear media if the pulse center is far from the nearest resonance frequency of the material's susceptibility. In some sense it represents the opposite extreme from the NLS approximation since that results from expanding the susceptibility in the frequency while equation results from expanding the susceptibility in the wavelength. Our analytical result was verified by numerical simulations showing that solutions of Maxwell's equation are well approximated by solutions of our short pulse equation. We concluded by proving the local existence and uniqueness of solutions of our approximate equation and showed that the approximating equation has no pulse-like solutions that propagate with fixed shape in a moving frame.

Appendix: The commutator estimates:

In this appendix, we prove Lemma 4.4.

If we write out the expression for Lu in Fourier space, we find

$$\hat{Lu}(k) = \int \hat{V}(k-p)\left(\mathrm{e}^{i\alpha x(\frac{1}{k}-\frac{1}{p})}-1\right)(ip)\hat{u}(p)dp \tag{36}$$

We define

$$K(k,p,x) = \mathrm{e}^{i\alpha x(\frac{1}{k}-\frac{1}{p})}-1 \tag{37}$$

and assume $v \in C^0([0,X],H^2)$ and bound the H^2 norm of Lu (the case H^s for $s>2$ works in a similar way)

$$\begin{aligned}\|(Lu))\|_{H^2}^2 \leq \int\int\int (1+k^2)^2|\hat{V}(k-p_1)|\;|\hat{V}(k-p_2)|\;|K(k,p_1,x)|\;\cdot \\ |K(k,p_2,x)|\;|p_1|\;|p_2|\;|\hat{u}(p_1)|\;|\hat{u}(p_2)|\;dp_1dp_2dk\end{aligned} \tag{38}$$

We estimate the right hand side of the above equation by breaking the integrals up into regions where k and p_j are "large" or "small". Define

$$\Xi^<(k) = \begin{cases} 1 \text{ if } |k|<C_0 \\ 0 \text{ otherwise}\end{cases} \tag{39}$$

Additionally, we set $\Xi^> = 1-\Xi^<$. The constant C_0 is chosen so that if $|k|,|p|>C_0$.

$$|K(k,p,x)| \leq 2\alpha|x|\left|\frac{1}{k}-\frac{1}{p}\right|. \tag{40}$$

(Recall that $|x|\leq X_0$, so this estimate holds uniformly for x in this range.)

Then (38) can be bounded by a sum of eight terms of the form

$$\int\int\int \Xi(k)^a\Xi(p_1)^b\Xi(p_2)^c(1+k^4)|\hat{V}(k-p_1)|\;|\hat{V}(k-p_2)|\;\cdot$$

$$|K(k,p_1,x)||K(k,p_2,x)|\ |p_1|\ |p_2|\ |\hat{u}(p_1)|\ |\hat{u}(p_2)|\ dp_1dp_2dk \tag{41}$$

where a, b, c take the values in $\{<, >\}$. We now estimate the cases that arise.

Case 1

Let $a =<$ and b, c be anything. Then $k^4 \le C_0^4$ and (41) is bounded by

$$C\int\int\left(\int|\hat{V}(k-p_1)|\ |\hat{V}(k-p_2)|dk\right)|p_1|\ |p_2|\ |\hat{u}(p_1)|\ |\hat{u}(p_2)|\ dp_1dp_2 \tag{42}$$

where we also used the fact that $|K(k,p,x)| \le 2$. Bounding the integral over k with the Cauchy-Schwarz inequality, we obtain an estimate of (42) by

$$C\|V\|_{L^2}^2\left(\int|p|\ |\hat{u}(p)|dp\right)^2 \le \tag{43}$$

$$C\|V\|_{L^2}^2\left(\int\frac{1}{1+|p|}((1+|p|)|p|\ |\hat{u}(p)|)dp\right)^2.$$

Applying the Cauchy-Schwarz inequality again, this is bounded by

$$C\|V\|_{L^2}^2\|u\|_{H^2}^2.$$

Case 2a

Let $a =>$ and $b = c =<$. In this case we bound

$$k^4 \le C((k-p_1)^2+p_1^2)((k-p_2)^2+p_2^2) \tag{44}$$

and then use the fact that all factors of $|p_1|$ and $|p_2|$ are bounded by C_0 to estimate (41) by

$$C\int\int\left(\int(1+|k-p_1|^2)|\hat{V}(k-p_1)|(1+|k-p_2|^2)|\hat{V}(k-p_2)|dk\right)$$
$$\cdot|\hat{u}(p_1)|\ |\hat{u}(p_2)|\ dp_1dp_2. \tag{45}$$

Applying the Cauchy-Schwarz inequality to the k integral, this is bounded by

$$C\|V\|_{H^2}^2\left(\int|\hat{u}(p)|dp\right)^2 \le C\|V\|_{H^2}^2\|u\|_{H^1}^2 \tag{46}$$

Case 2b

Now let $a =>$ and $b = c =>$. In this case we use (40) to bound the factors of K and estimate (41) by

$$
\begin{aligned}
\mathcal{J} = & C\alpha^2 x^2 \int\int\int k^4 |\hat{V}(k-p_1)|\, |\hat{V}(k-p_2)| \left|\frac{k-p_1}{kp_1}\right| \left|\frac{k-p_2}{kp_2}\right| \\
& \cdot |p_1|\, |p_2|\, |\hat{u}(p_1)|\, |\hat{u}(p_2)| dp_1 dp_2 dk \\
\leq & C\alpha^2 x^2 \int\int\int k^2 |\hat{V}(k-p_1)|\, |\hat{V}(k-p_2)| \\
& \cdot |k-p_1|\, |k-p_2|\, |\hat{u}(p_1)|\, |\hat{u}(p_2)| dp_1 dp_2 dk \\
\leq & C\alpha^2 x^2 \int\int\int (1+|k-p_2|^2) |\hat{V}(k-p_1)| (1+|k-p_2|^2) |\hat{V}(k-p_2)| \\
& \cdot (1+|p_1|)(1+|p_2|) |\hat{u}(p_1)|\, |\hat{u}(p_2)| dp_1 dp_2 dk \\
\leq & C\alpha^2 x^2 \|V\|_{H^2}^2 \left(\int (1+|p_1|)|\hat{u}(p_1)| dp_1\right)^2 \\
\leq & C\alpha^2 x^2 \|V\|_{H^2}^2 \|u\|_{H^2}^2
\end{aligned}
\tag{47}
$$

(Note: We have only estimated the term in the integrand of (41) proportional to k^4. The other terms are easy to bound.)

Case 2c

The last case is for $a =>$ and b and c different. It can be treated by combining the techniques used in the cases 2a and 2b.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. 0073923 and 0103915. The authors thank K. Promislow for first calling to their attention the work [1]. CEW thanks J. Rauch for very useful discussions of the approximation of short pulses. This work was begun while the second author was a sabbatical visitor in the Division of Applied Mathematics at Brown University whose hospitality is gratefully acknowledged.

References

[1] D. Alterman and J. Rauch. Diffractive short pulse asymptotics for nonlinear wave equations. *Phys. Lett. A*, 264:390–395, 2000.

[2] D. Alterman and J. Rauch. Diffractive nonlinear geometric optics for short pulses. Preprint, 2002.

[3] G. P. Agrawal. *Nonlinear Fiber Optics.* Academic Press, Boston, 1989.

[4] N. Karasawa, S. Nakamura, N. Nakagawa, M. Shibata, R. Morita, H. Shigekawa, and M. Yamashita. Comparision between theory and experiment of nonlinear propagation for a-few-cycle and ultrabroadband optical pulses in a fused-silica fiber. *IEEE J. Quant. Elect.*, 37:398–404, 2001.

[5] J. E. Rothenberg. Space-time focusing: breakdown of the slowly varying envelope approximation in the self-focusing of femtosecond pulses. *Opt. Lett.*, 17:1340–1342, 1992.

[6] J. K. Ranka and A. L. Gaeta. Breakdown of the slowly varying envelope approximation in the self-focusing of ultrashort pulses. *Opt. Lett.*, 23:534–536, 1998.

[7] A. C. Newell and J. V. Moloney. *Nonlinear Optics.* Addison-Wesley, Redwood City, CA, 1992.

[8] K. J. Blow and D. Wood. Theoretical description of transient stimulated Raman scattering in optical fibers. *IEEE J. Quant. Elect.*, 25:2665–2673, 1989.

[9] P. V. Mamyshev and S. V. Chernikov. Ultrashort-pulse propagation in optical fibers. *Opt. Lett.*, 15:1076–1078, 1990.

[10] A. L. Berkhoer and V. E. Zakharov. Self excitation of waves with different polarizations in nonlinear media. *Soviet Phys. JETP*, 31:486–490, 1970.

[11] S. V. Chernikov and P. V. Mamyshev. Femtosecond soliton propagation in fibers with slowly decreasing dispersion. *J. Opt. Soc. Am. B*, 8:1633–1641, 1991.

[12] T. Brabec and F. Krausz. Nonlinear optical pulse propagation in the single-cycle regime. *Phys. Rev. Lett.*, 78:3282–3285, 1997.

[13] P. M. Goorjian, A. Taflove, R. M. Joseph, and S. C. Hagness. Computational modeling of femtosecond optical solitons from Maxwell's equations. *IEEE J. Quant. Elect.*, 28:2416–2422, 1992.

[14] C. V. Hile and W. L. Kath. Numerical solutions of Maxwell's equations for nonlinear-optical pulse propagation. *J. Opt. Soc. Am. B*, 13:1135–1145, 1996.

[15] R. W. Boyd. *Nonlinear Optics.* Academic Press, Boston, 1992.

[16] I. H. Maliton. Interspecimen comparison of the refractive index of fused silica. *J. Opt. Soc. Amer.*, 55(10):1205–1210, October 1965.

[17] G. Schneider and C. E. Wayne. The long-wave limit for water wave problem: I the case of zero surface tension. *Comm. Pure Appl. Math.*, 53(12):1475–1535, 2000.

[18] Tosio Kato. Abstract evolution equations, linear and quasilinear, revisited. In *Functional analysis and related topics, 1991 (Kyoto)*, pages 103–125. Springer, Berlin, 1993.

[19] G. Schneider and H. Uecker. Existence and stability of modulating pulse solutions in maxwell's equations describing nonlinear optics. Preprint, 2001.