ULTRA-SHORT PULSES IN LINEAR AND NONLINEAR MEDIA

Y. CHUNG, C. K. R. T. JONES, T. SCHÄFER, AND C. E. WAYNE

ABSTRACT. We consider the evolution of ultra-short optical pulses in linear and nonlinear media. For the linear case, we first show that the initial-boundary value problem for Maxwell's equations in which a pulse is injected into a quiescent medium at the left endpoint can be approximated by a linear wave equation which can then be further reduced to the linear short-pulse equation. A rigorous proof is given that the solution of the short pulse equation stays close to the solutions of the original wave equation over the time scales expected from the non-rigorous multiple scales derivation of the short pulse equation. For the nonlinear case we present the results of a series of numerical computations which compare the predictions of the traditional nonlinear Schrödinger equation (NLSE) approximation which those of the short pulse equation. These computations show clearly that as the pulse length shortens, the NLSE approximation becomes steadily less accurate while the short pulse equation provides a better and better approximation. In an appendix, we also present a new (non-rigorous) way of deriving the nonlinear Schrödinger equation from Maxwell's equations using the renormalization group method.

1. INTRODUCTION

The standard model for describing propagation of pulses in nonlinear Maxwell's equations is the cubic nonlinear Schrödinger equation (NLSE) [1, 2]. Two main assumptions are made in the derivation of the NLSE from Maxwell's equations: First, it is assumed that the response of the material attains a quasi-steady-state and second that the pulse width is large in comparison to the oscillations of the carrier frequency [3].

In most of the applications of the NLSE in the past, i.e., in the case of pulse propagation in optical fibers, both assumptions were well satisfied. At present, however, technology for creating very short pulses has advanced a lot and experiments with pulses which are as short as a few cycles of the carrier wave have become possible [4]. The description of those pulses lies beyond the slowly varying envelope approximation leading to the NLSE [5]. Various approaches have been proposed to replace the NLSE – see, for example [6, 7, 8] for a sample of these methods. In [9], building on work of [10], two of us proposed an alternative model to approximate the evolution of very short pulses in nonlinear media.

In the present paper we study further the short pulse equation derived in [9]. There are two main sections of this paper. In the first we concentrate on giving a rigorous justification of several of the assumptions and approximations made in [9], for the *linearized* short pulse equation. Among the approximations used in [9] were:

(i) The linearized polarizability of the medium could be approximated by the Fourier transformed expression

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

(see equation (3) of Ref. [9]) Typically, in deriving (1) one assumes that the medium has reached some quasi-stationary state, and as J. Rauch pointed out to us, it is not clear that for these very short pulses the medium will have time to reach such a state before the pulse passes.

(ii) If the expression (1) is an accurate approximation to the polarizability of the medium, then can one really approximate solutions of the resulting equation which correspond to "short" pulses by solutions of the "short pulse" equation derived in [9] and if so, over what time interval does it provide an accurate approximation to the true evolution.

Note that the first of these points is a question just about the linear problem and thus in the next section we consider points (i) and (ii) in the context of a medium whose polarization is assumed to be linear. We study solutions of an initial-boundary value problem in which a linear wave equation is coupled to a medium whose polarization is modeled by a damped, linear oscillator. We inject a (short) pulse into the left end of this material and study how that pulse evolves with time. We prove rigorously that one can, even in this short pulse regime, approximate the polization of the material by the quasi-stationary approximation (1), and we show that if the pulse length is measured by the small parameter ϵ, then the linearized version of the short pulse equation accurately describes the true solution of the equation over time scales of $\mathcal{O}(1/\epsilon)$.

The second part of the paper is primarily a numerical study of the effect of pulse length on the propagation of pulses in nonlinear materials. We compare the (numerical) evolution of solutions of a one-dimensional, nonlinear version of Maxwell's equations with approximations given both by the NLSE and the short pulse equation. We find as expected from the assumptions that underlie the formal derivation of these two equations that for slowly modulated pulses the NLSE does a better job of approximating the evolution, but as the pulses become shorter and shorter the NLSE approximation breaks down and the short-pulse equation provides a better approximation to the true dynamics.

2. Formulation of the linear initial-boundary value problem

In this section we study the propagation of a short pulse injected at one end of a semiinfinite fiber with linear polarizability. If we assume that the polarization of the electric field is transverse to the fiber then the magnitude of the electric field, $u(x,t)$, in appropriately non-dimensional units satisfies the partial differential equation

$$u_{xx} = u_{tt} + p_{tt} \tag{2}$$

where $p(x,t)$ stands for the (magnitude of the) polarization of the material, which we assume is parallel to the electric field.

We model the polarization of the material by a damped harmonic oscillator, so $p(x,t)$ satisfies the equation

$$p_{tt} + \Gamma p_t + \omega_0^2 p = \chi_0 u. \tag{3}$$

Remark 2.1. *In general, the polarization of the medium is modeled as a sum of oscillators, each with its own natural frequency and damping constant. However, as was remarked in* [9], *for infrared pulses in silica fibers, the polarizability can be accurately modeled with a single resonance. Furthermore, in the present circumstances, the linear nature of the problem means that if the polarization was modelled as a sum of several oscillators, we could write the solution as a sum of the solutions to problems of the type (3) with different resonant frequencies and damping constants.*

Remark 2.2. *Experimentally, the damping of these oscillators is quite weak –* Γ *is small. Thus, throughout this paper we will assume that* $\Gamma^2 - 4\omega_0^2 < 0$ *and also that* $\chi_0 > 0$.

Physically, we are interested in the situation where our medium is semi-infinite with one end at the origin and all fields in the medium are initially zero. We enforce this condition by taking initial conditions

$$u(x,0) = u_t(x,0) = p(x,0) = p_t(x,0) = 0 \text{ for } x > 0 \tag{4}$$

We inject a pulse into the left end of the fiber and model this by assuming a boundary condition at $x = 0$ of the form

$$u(0,t) = U_0(t) \tag{5}$$

where $U_0(t)$ presents the optical pulse. Since we are interested in short pulses, we will assume that the injected pulse has the form $U_0(t) = \mathcal{U}_0(t/\epsilon)$. We also assume that $\mathcal{U}_0$ smooth and that it is supported in the interval $[0,1]$.

Summing up, the equations (2), (3) together with the boundary condition (5) and the initial conditions (4) form the initial-boundary value problem (IBVP) describe the short-pulse propagation.

Remark 2.3. *In general one can also expect an instantaneous contribution to the polarization – we ignore this case here since it can be incorporated simply by changing the coefficient in front of the term* u_{tt}.

We prove two approximation results about the solutions of the IBVP (2). Recall that from the formal derivation of the short pulse equation in [9], we expect that approximation of solutions of (2) by the (linearized) short pulse equation should be valid for times $\mathcal{O}(\frac{1}{\epsilon})$. In fact, following the usual convention in nonlinear optics the multiple scale expansion in [9] was made in terms of a long *space* scale, ϵx, rather than a long time scale ϵt, and thus one might expect an approximation result valid over *length* scales of $\mathcal{O}(1/\epsilon)$. However, since pulses in Eqn. (2) travel with a speed $\mathcal{O}(1)$, we can translate our approximation result into a result valid over long time scales. Our first result shows that for times of this order we can approximate solutions of (2) by solutions of the single equation

$$\tilde{u}_{xx} = \tilde{u}_{tt} + \chi_0 \tilde{u} \ , \tag{6}$$

with initial conditions $\tilde{u}(x,0) = \tilde{u}_t(x,0) = 0$ and boundary condition $\tilde{u}(0,t) = U_0(t)$. Note that this corresponds precisely to the approximation in equation (5) in [9] and thus gives a rigorous justification of the heuristic argument of that paper. More precisely we prove the following:

Proposition 2.4. *Let* $T_0 > 0$ *be fixed. Then there exist* $\epsilon_0 > 0$ *and* $C_0 > 0$ *such that if* $0 < \epsilon < \epsilon_0$, *and* $u(x,t)$ *satisfies the IBVP* $\{(2), (3), (5), (4)\}$ *and* $\tilde{u}(x,t)$

satisfies (6), with zero initial data and the boundary condition (5) then

$$\sup_{0\le t\le T_0/\epsilon}(\sup_{x>0}|u(x,t)-\tilde{u}(x,t)|)\le C_0\epsilon^{1/2}\ . \tag{7}$$

In this context we can also show that the linearized version of the short pulse equation derived in [9] correctly describes the propagation of solutions in either (2) or (6). The linearized short pulse equation describing the evolution of a function of two variables $\mathcal{U}=\mathcal{U}(\phi,T)$ is

$$2\partial_\phi\partial_T\mathcal{U}=\chi_0\mathcal{U}\ . \tag{8}$$

We also prove

Proposition 2.5. *Let $T_0>0$ be fixed. Then there exist $\epsilon_0>0$ and $C_0>0$ such that if $0<\epsilon<\epsilon_0$, and $u(x,t)$ satisfies the IBVP $\{(2), (3), (5), (4)\}$ then there exists a solution of the pulse equation (8), $\mathcal{U}(\phi,T)$ such that*

$$\sup_{1\le t\le T_0/\epsilon}(\sup_{x>0}|u(x,t)-\mathcal{U}(\frac{t-x}{\epsilon},\epsilon t)|)\le C_0\epsilon^{1/2}\ . \tag{9}$$

Remark 2.6. *Note that in this proposition we do not begin to compare the solution of the "true" evolution (2) with the solution of the pulse equation until a time $t>1$ – i.e. until the pulse has been injected at the left boundary of the domain. This is because we don't expect the pulse equation to describe the evolution before there is a pulse present in the system!*

Remark 2.7. *In both of these propositions the constants C_0 depend on the profile of the injected pulse, $\mathcal{U}_0$, in a way we make precise in the proof.*

We note that in [11], Alterman and Rauch study in detail the properties of a linear short pulse equation similar to (8), but lacking the term $\chi_0\mathcal{U}$.

3. Proofs of the approximation results

We begin with the proof of Proposition 2.4. In the proofs of both propositions we will work with the Fourier sine and cosine transforms. Given a function $v(x)$ defined on the positive half-line we define:

$$\hat{v}^s(k)=\int_0^\infty \sin(kx)v(x)dx\ , \tag{10}$$

$$\hat{v}^c(k)=\int_0^\infty \cos(kx)v(x)dx\ . \tag{11}$$

We will also use the Laplace transform which we denote by $\mathcal{L}$. If we take both the Fourier-sine and Laplace transforms of (2) and use the boundary and initial conditions we find

$$-k^2\mathcal{L}[\hat{u}^s]+k\mathcal{L}[U_0]=s^2\mathcal{L}[\hat{u}^s]+s^2\mathcal{L}[\hat{p}^s]\ , \tag{12}$$

$$(s^2+s\Gamma+\omega_0^2)\mathcal{L}[\hat{p}^s]=\chi_0\mathcal{L}[\hat{u}^s]\ . \tag{13}$$

We can combine these two expressions to obtain a single equation for $\mathcal{L}[\hat{u}^s]$, namely

$$\mathcal{L}[\hat{u}^s](k,s)=\left\{\frac{k}{(k^2+s^2)+\left(\frac{\chi_0 s^2}{s^2+s\Gamma+\omega_0^2}\right)}\right\}\mathcal{L}[U_0](s)\ . \tag{14}$$

If we now take the inverse Laplace transform of this expression and use the fact that $U_0(t) = \mathcal{U}(\frac{t}{\epsilon})$ we find that

$$\hat{u}^s(k,t) = \epsilon \int_0^{\frac{t}{\epsilon}} \mathcal{L}^{-1}\left\{ \frac{k}{(k^2+s^2) + \left(\frac{\chi_0 s^2}{s^2+s\Gamma+\omega_0^2}\right)} \right\} (\epsilon[\frac{t}{\epsilon} - \sigma])\mathcal{U}_0(\sigma)d\sigma \ . \tag{15}$$

We now rewrite this expression with the aid of the following standard lemma about Laplace transforms:

Lemma 3.1. *If $F(s) = \mathcal{L}[f](s)$, then*

$$\epsilon f(\epsilon t) = \mathcal{L}^{-1}[F(\frac{\cdot}{\epsilon})](t) \ .$$

Using Lemma 3.1 and defining $p = \epsilon k$, we find that

$$\hat{u}^s(\frac{p}{\epsilon},t) = \int_0^{\frac{t}{\epsilon}} \mathcal{L}^{-1}[F(s,p;\epsilon)](\frac{t}{\epsilon} - \sigma)\mathcal{U}_0(\sigma)d\sigma \ , \tag{16}$$

where

$$F(s,p;\epsilon) = \frac{(\epsilon p)(s^2+\epsilon s\Gamma+\epsilon^2\omega_0^2)}{(s^2+\epsilon s\Gamma+\epsilon^2\omega_0^2)(p^2+s^2)+\epsilon^2\chi_0 s^2} \ . \tag{17}$$

One can compute the inverse Laplace transform of F via it's partial fraction expansion and for that we need to find the roots of the polynomial in the denominator of F, i.e.

$$Q(s;p,\epsilon) = (s^2+\epsilon s\Gamma+\epsilon^2\omega_0^2)(p^2+s^2)+\epsilon^2\chi_0 s^2 \ . \tag{18}$$

To this end we use the following series of Lemmas whose proofs are elementary but somewhat involved. Hence we relegate the proofs to Appendix A.

Lemma 3.2. *For all values of p and ϵ all the eigenvalues of Q have non-positive real part.*

Lemma 3.3. *There exist $\epsilon_0 > 0$ and $C_0, C_1 > 0$ such that for $|\epsilon| < \epsilon_0$ and $p > C_0\epsilon$, Q has a pair of roots of the form*

$$s_\pm^0 = -\frac{\epsilon}{2}(\Gamma \pm \sqrt{\Gamma^2 - 4\omega_0^2}) + \epsilon\sigma_\pm^0(p)$$

with

$$|\sigma_\pm^0(p)| \leq \frac{C_1\epsilon^2}{p^2} \ .$$

Lemma 3.4. *There exists $\epsilon_0 > 0$ and $C_0 > 0$ such that for $|\epsilon| < \epsilon_0$ and $p > C_0\epsilon$, Q has a pair of roots of the form*

$$s_\pm^1 = \pm i\sqrt{p^2+\epsilon^2\chi_0} + \sigma_\pm^1(p)$$

with

$$|\sigma_\pm^1(p)| \leq \frac{C_0\epsilon^3}{p^2} \ .$$

Furthermore, the real part of $\sigma_\pm^1$ is negative.

Lemma 3.5. *Let ϵ_0 and C_0 be as in Lemma 3.4. There exists $c_{min} > 0$ such that for $p < C_0\epsilon$ the roots of Q can be written as $s_\pm^0(p) = \epsilon\tilde{s}_\pm^0(p/\epsilon)$ and $s_\pm^1(p) = \epsilon\tilde{s}_\pm^1(p/\epsilon)$ with $\tilde{s}_\pm^0(q)$ and $\tilde{s}_\pm^1(q)$ all distinct. Furthermore,*

$$\min(|\tilde{s}_+^0(q) - \tilde{s}_-^0(q)|, |\tilde{s}_\pm^0(q) - \tilde{s}_\pm^1(q)|) \geq c_{min} \ ,$$

while

$$|\tilde{s}^1_+(q) - \tilde{s}^1_-(q)| \geq c_{min} q \ .$$

Remark 3.6. *Note that a corollary of the proof of Lemma 3.5 is that for $p > 0$ the roots of Q are all distinct.*

With these estimates on the eigenvalues of Q in hand we now construct the partial fraction expansion of F. Note that since the eigenvalues of Q depend continuously on p (and ϵ) we can label the eigenvalues as $s^0_\pm(p)$ and $s^1_\pm(p)$ for all $p \geq 0$, and then one can write

$$F(s,p;\epsilon) = \frac{A^0_+(p)}{s - s^0_+(p)} + \frac{A^0_-(p)}{s - s^0_-(p)} + \frac{A^1_+(p)}{s - s^1_+(p)} + \frac{A^1_-(p)}{s - s^1_-(p)} \ . \tag{19}$$

If we also write

$$F(s,p;\epsilon) = \frac{(\epsilon p)(s^2 + \epsilon s \Gamma + \epsilon^2 \chi_0)}{(s - s^0_+(p))(s - s^0_-(p))(s - s^1_+(p))(s - s^1_-(p))} \ , \tag{20}$$

then we see that A^0_+ has the form

$$A^0_+ = \frac{(\epsilon p)((s^0_+)^2 + \epsilon(s^0_+)\Gamma + \epsilon^2 \omega_0^2)}{((s^0_+) - s^0_-)((s^0_+) - s^1_+)((s^0_+) - s^1_-)} \ . \tag{21}$$

First note that if $p < C_0 \epsilon$, the numerator of this expression can be bounded by $C\epsilon^4$ by Lemma 3.5 while the same lemma guarantees that the denominator is bounded below by $c\epsilon^3$ for some $c > 0$. Thus for p in this range $|A^0_+| \leq \epsilon$.

Now suppose that $p > C_0 \epsilon$. Since s^0_+ is a root of Q, we have

$$|((s^0_+)^2 + \epsilon(s^0_+)\Gamma + \epsilon^2 \omega_0^2)| = \left| \frac{\epsilon^2 \chi_0 (s^0_+)^2}{(s^0_+)^2 + p^2} \right| .$$

Now we see that this expression is bounded by $C\epsilon^4/p^2$. Using the asymptotic expressions for the roots of Q coming from Lemma 3.3 and 3.4 the denominator in (21) can be bounded from below by ϵp^2 in this range and hence we can bound

$$|A^0_+| \leq \frac{C\epsilon^3}{p^2} \ ,$$

for p in this range. We can combine the estimates on $|A^0_+|$ in these two ranges of p along with identical estimates on A^0_- to obtain

Lemma 3.7. *There exists a constant $C_A > 0$ such that for all $p > 0$,*

$$|A^0_\pm(p)| \leq \frac{C_A \epsilon^3}{p^2 + \epsilon^2} \ .$$

We now estimate the coefficients $A^1_\pm$. Following an argument similar to that above we find

$$A^1_+ = \frac{(\epsilon p)((s^1_+)^2 + \epsilon(s^1_+)\Gamma + \epsilon^2 \omega_0^2)}{((s^1_+) - s^1_-)((s^1_+) - s^0_+)((s^1_+) - s^0_-)} \ . \tag{22}$$

For $p < C_0 \epsilon$, using the asymptotic values of the roots s^1_+ and $s^0_\pm$ given in Lemma 3.5 we see that

$$|A^1_+(p)| \leq C\epsilon \ .$$

For $p \geq C_0\epsilon$ we use Lemma 3.3 and 3.4 to rewrite

$$\frac{(s_+^1)^2 + \epsilon(s_+^1)\Gamma + \epsilon^2\omega_0^2}{((s_+^1) - s_+^0)((s_+^1) - s_-^0)} = \frac{((s_+^1)^2 + \epsilon(s_+^1)\Gamma + \epsilon^2\omega_0^2)}{((s_+^1)^2 + \epsilon\Gamma s_+^1 - \epsilon(\sigma_+^0 + \sigma_-^0)s_+^1 + s_+^0 s_-^0)}$$
$$= 1 + \mathcal{E}(p;\epsilon) \ , \tag{23}$$

where

$$|\mathcal{E}(p;\epsilon)| \leq \frac{C_E\epsilon^2}{p^2+\epsilon^2} \ . \tag{24}$$

Thus,

$$A_+^1 = \frac{\epsilon p(1+\mathcal{E}(p;\epsilon))}{(2i\sqrt{p^2+\epsilon^2\chi_0} + (\sigma_+^1 - \sigma_-^1))} \ , \tag{25}$$

or if we write

$$A_+^1 = \frac{\epsilon p}{2i\sqrt{p^2+\epsilon^2\chi_0}} + \Delta A_+^1 \ , \tag{26}$$

we see that for $p > C_0\epsilon$, (and C_0 sufficiently large)

$$|\Delta A_+^1(p)| \leq \frac{C_A\epsilon^3 p}{(p^2+\epsilon^2)^{3/2}} \ . \tag{27}$$

Similar estimates hold for A_-^1 and we have established:

Lemma 3.8. *There exists a constant $C_A > 0$ such that for $p < C_0\epsilon$,*

$$|A_\pm^1(p)| \leq C_A\epsilon \ ,$$

while for $p > C_0\epsilon$,

$$A_\pm^1(p) = \frac{\pm\epsilon p}{2i\sqrt{p^2+\epsilon^2\chi_0}} + \Delta A_\pm^1 \ ,$$

with

$$|\Delta A_\pm^1(p)| \leq \frac{C_A\epsilon^3 p}{(p^2+\epsilon^2)^{3/2}} \ .$$

With these estimates on the coefficients in the partial fraction decomposition of $F(s,p;\epsilon)$ in hand, we now return to the task of computing the solution $u(x,t)$ of (2). From (16) and (19) we see that

$$\begin{aligned} u(x,t) &= \frac{2}{\pi\epsilon}\int_0^\infty \sin(p(\frac{x}{\epsilon}))\int_0^{\frac{t}{\epsilon}} A_+^0(p)e^{s_+^0(p)(\frac{t}{\epsilon}-\sigma)}\mathcal{U}_0(\sigma)d\sigma dp \\ &+\frac{2}{\pi\epsilon}\int_0^\infty \sin(p(\frac{x}{\epsilon}))\int_0^{\frac{t}{\epsilon}} A_-^0(p)e^{s_-^0(p)(\frac{t}{\epsilon}-\sigma)}\mathcal{U}_0(\sigma)d\sigma dp \\ &+\frac{2}{\pi\epsilon}\int_0^\infty \sin(p(\frac{x}{\epsilon}))\int_0^{\frac{t}{\epsilon}} A_+^1(p)e^{s_+^1(p)(\frac{t}{\epsilon}-\sigma)}\mathcal{U}_0(\sigma)d\sigma dp \\ &+\frac{2}{\pi\epsilon}\int_0^\infty \sin(p(\frac{x}{\epsilon}))\int_0^{\frac{t}{\epsilon}} A_-^1(p)e^{s_-^1(p)(\frac{t}{\epsilon}-\sigma)}\mathcal{U}_0(\sigma)d\sigma dp \ . \end{aligned} \tag{28}$$

We can immediately bound the first two terms on the right hand side of (28) by using the facts that the real parts of $s_\pm^0$ are negative, so that the exponential factor is bounded by 1, as is the factor of $\sin(px/\epsilon)$, and bounding $A_\pm^0$ by the bound in Lemma 3.7. Integrating over p and σ we then see that these two lines are bounded by $C\epsilon\|\mathcal{U}_0\|_{L^1}$. We now turn to the last two lines in (28). First of all, rewriting them with the aid of Lemma 3.4 and 3.8 as

$$\begin{aligned}
&\frac{2}{\pi\epsilon}\int_{p=C_1\sqrt{\epsilon}}^{\infty}\int_0^{\frac{t}{\epsilon}} \frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}\sin(p(\frac{x}{\epsilon}))\sin(\sqrt{p^2+\epsilon^2\chi_0}(\frac{t}{\epsilon}-\sigma))\mathcal{U}_0(\sigma)d\sigma dp\\
&+\frac{2}{\pi\epsilon}\int_{p=C_1\sqrt{\epsilon}}^{\infty}\int_0^{\frac{t}{\epsilon}}(\Delta A_+^1(p)e^{s_+^1(p)(\frac{t}{\epsilon}-\sigma)}\\
&\qquad +\Delta A_-^1(p)e^{s_-^1(p)(\frac{t}{\epsilon}-\sigma)})\sin(p(\frac{x}{\epsilon}))\mathcal{U}_0(\sigma)d\sigma dp\\
&+\frac{2}{\pi\epsilon}\int_{p=C_1\sqrt{\epsilon}}^{\infty}\int_0^{\frac{t}{\epsilon}}\sin(p(\frac{x}{\epsilon}))\{\frac{1}{2i}\frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}e^{i\sqrt{p^2+\epsilon^2\chi_0}}(e^{\sigma_+^1(\frac{t}{\epsilon}-\sigma)}-1)\\
&\qquad -\frac{1}{2i}\frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}e^{-i\sqrt{p^2+\epsilon^2\chi_0}}(e^{\sigma_-^1(\frac{t}{\epsilon}-\sigma)}-1)\}\mathcal{U}_0(\sigma)d\sigma dp \qquad (29)\\
&+\frac{2}{\pi\epsilon}\int_0^{p=C_1\sqrt{\epsilon}}\int_0^{\frac{t}{\epsilon}}\sin(p(\frac{x}{\epsilon}))\left\{A_+^1(p)e^{s_+^1(p)(\frac{t}{\epsilon}-\sigma)}+A_-^1(p)e^{s_-^1(p)(\frac{t}{\epsilon}-\sigma)}\right\}\mathcal{U}_0(\sigma)d\sigma dp.
\end{aligned}$$

The last of these integrals can be immediately bounded by $C_A\sqrt{\epsilon}\|\mathcal{U}_0\|_{L^1}$. The next to last integral is estimated by using Lemma 3.4 and the fact that for $0\le t\le T_0/\epsilon$,

$$|e^{\sigma_\pm^1(\frac{t}{\epsilon}-\sigma)}-1|\le\frac{C\epsilon}{p^2}\ .$$

With this estimate the integral can be bounded by

$$C\int_{C_1\sqrt{\epsilon}}^{\infty}\left(\frac{p}{\sqrt{p^2+\epsilon^2\chi_0}}+\frac{\epsilon^3 p}{(p^2+\epsilon^2)^{3/2}}\right)\frac{\epsilon}{p^2}\|\mathcal{U}_0\|_{L^1}dp\le C\sqrt{\epsilon}\|\mathcal{U}_0\|_{L^1}\ .$$

Finally, in the second term we use the bounds on $\Delta A_\pm^1$ from Lemma 3.8 to estimate this integral by

$$C\epsilon^2\int_{C_1\sqrt{\epsilon}}^{\infty}\frac{p}{(p^2+\epsilon^2)^{3/2}}dp\|\mathcal{U}_0\|_{L^1}\le C\epsilon\|\mathcal{U}_0\|_{L^1}\ .$$

Combining these estimates with those on the first two integrals in (28) we see that

$$\begin{aligned}
\sup_{0\le t\le T_0/\epsilon}\Bigg(\sup_{x>0}\Bigg|u(x,t)-\frac{2}{\pi\epsilon}\int_{p=C_1\sqrt{\epsilon}}^{\infty}\int_0^{\frac{t}{\epsilon}}\frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}\sin(p(\frac{x}{\epsilon}))\times \qquad (30)\\
\times\ \sin\left(\sqrt{p^2+\epsilon^2\chi_0}(\frac{t}{\epsilon}-\sigma)\right)\mathcal{U}_0(\sigma)d\sigma dp\Bigg|\Bigg)\le C\sqrt{\epsilon}\|\mathcal{U}_0\|_{L^1}\ .
\end{aligned}$$

We now note that since

$$\begin{aligned}
&\left|\frac{2}{\pi\epsilon}\int_{p=0}^{C_1\sqrt{\epsilon}}\int_0^{\frac{t}{\epsilon}}\frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}\sin(p(\frac{x}{\epsilon}))\sin(\sqrt{p^2+\epsilon^2\chi_0}(\frac{t}{\epsilon}-\sigma))\mathcal{U}_0(\sigma)d\sigma dp\right|\\
&\qquad\le C\sqrt{\epsilon}\|\mathcal{U}_0\|_{L^1} \qquad (31)
\end{aligned}$$

we can subtract it from the left hand side of (30) without changing the bound on the right hand side of this expression. That is, we can bound the difference between $u(x,t)$ and the integral

$$\frac{2}{\pi\epsilon}\int_{p=0}^{\infty}\int_0^{\frac{t}{\epsilon}}\frac{\epsilon p}{\sqrt{p^2+\epsilon^2\chi_0}}\sin(p(\frac{x}{\epsilon}))\sin(\sqrt{p^2+\epsilon^2\chi_0}(\frac{t}{\epsilon}-\sigma))\mathcal{U}_0(\sigma)d\sigma dp.$$

by $C\sqrt{\epsilon}\|\mathcal{U}_0\|_L^1$.

By taking the Fourier sine transform of (6) we see that this integral is exactly $\tilde{u}(x,t)$, the solution of (6) and thus we have completed the proof of Proposition 2.4.

We next prove Proposition 2.5. By the results established so far it suffices to prove that the solution $\tilde{u}(x,t)$ of (6) can be approximated by a solution of the pulse equation (8) over the relevant time intervals. First note that using trigonometric identities for the sine and cosine we can rewrite

$$\tilde{u}(x,t) = u^L(x,t) + u^R(x,t) \tag{32}$$

where

$$u^R(x,t) = \frac{1}{\pi\epsilon}\int_0^\infty \int_0^{\frac{t}{\epsilon}} \frac{\epsilon p}{\omega_\epsilon(p)} \cos(\frac{1}{\epsilon}(px - \omega_\epsilon(p)t) + \sigma\omega_\epsilon(p))\mathcal{U}_0(\sigma) d\sigma dp \ , \tag{33}$$

and

$$u^L(x,t) = -\frac{1}{\pi\epsilon}\int_0^\infty \int_0^{\frac{t}{\epsilon}} \frac{\epsilon p}{\omega_\epsilon(p)} \cos(\frac{1}{\epsilon}(px + \omega_\epsilon(p)t) - \sigma\omega_\epsilon(p))\mathcal{U}_0(\sigma) d\sigma dp \ , \tag{34}$$

where $\omega_\epsilon(p) = \sqrt{p^2 + \epsilon^2\chi_0}$.

Roughly speaking, u^R and u^L represent the left and right moving parts of the pulse. In particular, for $t > 1$, we expect that the left moving part of the solution will no longer be relevant since we are only interested in the solution for $x > 0$. To prove this we use another trigonometric identity to rewrite u^L as

$$\begin{aligned} u^L(x,t) = &-\frac{1}{\pi}\int_0^\infty \frac{p}{\omega_\epsilon(p)} \cos(\frac{1}{\epsilon}(px + \omega_\epsilon(p)t)\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\ &-\frac{1}{\pi}\int_0^\infty \frac{p}{\omega_\epsilon(p)} \sin(\frac{1}{\epsilon}(px + \omega_\epsilon(p)t)\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp \ , \end{aligned} \tag{35}$$

and we recall that $\hat{\mathcal{U}}_0^c$ and $\hat{\mathcal{U}}_0^s$ are the cosine and sine transforms of the boundary data $\mathcal{U}_0$. (We have used here the fact that since the limit on the σ integral exceeds the limits on the support of $\mathcal{U}_0$ we can integrate from 0 to ∞.) We now prove that both of these terms are $\mathcal{O}(\sqrt{\epsilon})$ in the L^∞ norm and thus can be ignored to the order of approximation that we are concerned with.

We'll consider in detail the first of the two terms in (35). The second is handled in an almost identical fashion and we leave the details as an exercise. Rewrite that integral using a trigonometric identity as

$$\begin{aligned} &-\frac{1}{\pi}\int_0^\infty \{\cos(\frac{p}{\epsilon}(x+t))\cos(\frac{t}{\epsilon}(\omega_\epsilon(p) - p)\}\frac{p}{\omega_\epsilon(p)}\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\ &\frac{1}{\pi}\int_0^\infty \{\sin(\frac{p}{\epsilon}(x+t))\sin(\frac{t}{\epsilon}(\omega_\epsilon(p) - p)\}\frac{p}{\omega_\epsilon(p)}\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \ . \end{aligned} \tag{36}$$

Once again these two integrals are estimated in an almost identical fashion so we provide the details for the first and leave the second as an exercise. Integrating by parts, the first integral becomes

$$\begin{aligned}\frac{1}{\pi}\int_0^\infty \frac{\epsilon}{x+t}\sin(\frac{p}{\epsilon}(x+t))\{&\frac{t}{\epsilon}(\omega_\epsilon'(p)-1)\frac{p}{\omega_\epsilon(p)}\sin(\frac{t}{\epsilon}(\omega_\epsilon(p)-p))\hat{\mathcal{U}}_0^c(\omega_\epsilon(p)) \\ &-\cos(\frac{t}{\epsilon}(\omega_\epsilon(p)-p))[\frac{\omega_\epsilon(p)-p\omega_\epsilon'(p)}{(\omega_\epsilon(p))^2}]\hat{\mathcal{U}}_0^c(\omega_\epsilon(p) \\ &-\cos(\frac{t}{\epsilon}(\omega_\epsilon(p)-p))\frac{p}{\omega_\epsilon(p)}\omega_\epsilon'(p)\hat{\mathcal{U}}_0^{c\prime}(\omega_\epsilon(p))\}dp\ .\end{aligned} \tag{37}$$

Note that there exists a constant $C_1 > 0$, independent of ϵ such that the the various quotients appearing in the integrand of (37) can be bounded as follows:

$$p|\frac{\omega_\epsilon'(p)-1}{\omega_\epsilon(p)}| \le \begin{cases} C_1 & \text{for all } p > 0, \\ \frac{C_1\epsilon^2}{p^2} & \text{for all } p > C_0\sqrt{\epsilon}. \end{cases}$$

$$|\frac{\omega_\epsilon(p)-p\omega_\epsilon'(p)}{(\omega_\epsilon(p))^2}| \le \begin{cases} \frac{C_1}{\epsilon} & \text{for all } p > 0, \\ \frac{C_1\epsilon^2}{p^3} & \text{for all } p > C_0\sqrt{\epsilon}. \end{cases}$$

$$|\frac{p\omega_\epsilon'(p)}{\omega_\epsilon(p)}| \le C_1 \text{ for all } p > 0$$

Thus, bounding the factors of sine and cosine in (37) by 1 we see that this integral can be bounded by

$$\begin{aligned}&\frac{C}{x+t}\int_0^{C_0\sqrt{\epsilon}}(|t|+1)|\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))|dp \\ &+\frac{C\epsilon^2}{x+t}\int_{C_0\sqrt{\epsilon}}^\infty(|t|+1)(\frac{1}{p^2}+\frac{1}{p^3})|\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))|dp+\frac{C\epsilon}{x+t}\int_0^\infty|\hat{\mathcal{U}}_0^{c\prime}(\omega_\epsilon(p))|dp.\end{aligned} \tag{38}$$

In the first two of these integrals we bound $|\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))|$ by $C\|\mathcal{U}_0\|_{L^1}$ Thus, these first two integrals can be bounded by $C\sqrt{\epsilon}\|\mathcal{U}_0\|_{L^1}$. To bound the final integral write it as the sum

$$\int_0^1|\hat{\mathcal{U}}_0^{c\prime}(\omega_\epsilon(p))|dp+\int_1^\infty|\hat{\mathcal{U}}_0^{c\prime}(\omega_\epsilon(p))|dp$$

The first integral can again be bounded by $C\|\mathcal{U}_0\|_{L^1}$, while the second is bounded by $C\int_1^\infty|\hat{\mathcal{U}}_0^{c\prime}(\xi)|d\xi$, by making the change of variables $\xi = \omega_\epsilon(p)$. This integral can be bounded by $C(\int_1^\infty(1+\xi^2)|\hat{\mathcal{U}}_0^{c\prime}(\xi)|^2d\xi)^{1/2}$ by the Cauchy-Schwartz. Applying Parseval's equality and the fact taht $\mathcal{U}_0$ has finite support this integral is bounded by $\|\mathcal{U}_0\|_{H^1}$. Note that since $\mathcal{U}_0$ has finite support, one can also bound the L^1 norm of $\mathcal{U}_0$ by a constant times the H^1 norm, and thus, (38) is bounded by $\frac{C(t+1)}{x+t}\sqrt{\epsilon}\|\mathcal{U}_0\|_{H^1}$ A similar estimate applies to the remaining terms in u^L and so we conclude that

$$\sup_{t\ge1}\sup_{x>0}|\tilde{u}-u^R(x,t)| \le C\sqrt{\epsilon}\|\mathcal{U}_0\|_{H^1}\ . \tag{39}$$

We now examine u^R more closely and show that it can be approximated by a solution of the pulse equation. Begin by writing it as

$$\begin{aligned}u^R(x,t) &= \frac{1}{\pi}\int_0^\infty\frac{p}{\omega_\epsilon(p)}\cos(\frac{1}{\epsilon}(px-\omega_\epsilon(p)t)\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\ &+\frac{1}{\pi}\int_0^\infty\frac{p}{\omega_\epsilon(p)}\sin(\frac{1}{\epsilon}(px-\omega_\epsilon(p)t)\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp\ .\end{aligned} \tag{40}$$

Now define $\phi = \left(\frac{x-t}{\epsilon}\right)$ and $T = \epsilon t$. Then

$$\begin{aligned}
&\mathcal{U}^R(\phi, T) \equiv u^R(x(\phi,T), t(\phi,T)) = \\
&\quad \frac{1}{\pi}\int_0^\infty \frac{p}{\omega_\epsilon(p)} \cos(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2})\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\
&\quad +\frac{1}{\pi}\int_0^\infty \frac{p}{\omega_\epsilon(p)} \sin(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2})\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp \ ,
\end{aligned} \tag{41}$$

We now define

$$\begin{aligned}
\mathcal{U}(\phi, T) &= \frac{1}{\pi}\int_{\epsilon^{1/2}}^\infty \frac{p}{\omega_\epsilon(p)} \cos(p\phi - \frac{\chi_0 T}{2p})\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\
&\quad +\frac{1}{\pi}\int_{\epsilon^{1/2}}^\infty \frac{p}{\omega_\epsilon(p)} \sin(p\phi - \frac{\chi_0 T}{2p})\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp \ .
\end{aligned} \tag{42}$$

Note that by an easy and explicit computation $\mathcal{U}(\phi, T)$ satisfies the linearized pulse equation (8), hence, Proposition 2.5 will follow if we can show that $\Delta\mathcal{U} = \mathcal{U}^R - \mathcal{U}$ is small. Subtracting (42) from (41) we obtain

$$\begin{aligned}
\Delta\mathcal{U}(\phi, T) &= \frac{1}{\pi}\int_{\epsilon^{1/2}}^\infty \frac{p}{\omega_\epsilon(p)} \left\{\cos(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2}) - \cos(p\phi - \frac{\chi_0 T}{2p})\right\}\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\
&+\frac{1}{\pi}\int_{\epsilon^{1/2}}^\infty \frac{p}{\omega_\epsilon(p)} \left\{\sin(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2}) - \sin(p\phi - \frac{\chi_0 T}{2p})\right\}\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp \\
&+\frac{1}{\pi}\int_0^{\epsilon^{1/2}} \frac{p}{\omega_\epsilon(p)} \cos(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2})\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \\
&+\frac{1}{\pi}\int_0^{\epsilon^{1/2}} \frac{p}{\omega_\epsilon(p)} \sin(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2})\hat{\mathcal{U}}_0^s(\omega_\epsilon(p))dp \ .
\end{aligned} \tag{43}$$

Note that the integrals over $[0, \sqrt{\epsilon}]$ can be easily bounded by noting that $|p/\omega_\epsilon(p)| \leq 1$, hence both of these integrals are bounded by $C\sqrt{\epsilon}(|\hat{\mathcal{U}}_0^s|_{L^\infty} + |\hat{\mathcal{U}}_0^c|_{L^\infty})$. The two remaining integrals are bounded in a similar fashion – we give the details of the bound on the integral containing the difference of cosines and leave the other as an exercise. By the mean value theorem there exists some ξ such that

$$\begin{aligned}
&|\cos(p\phi + (p-\omega_\epsilon(p))\frac{T}{\epsilon^2}) - \cos(p\phi - \frac{\chi_0 T}{2p})| = |\sin(\xi)\left\{(p-\omega_\epsilon(p))\frac{T}{\epsilon^2} + \frac{\chi_0 T}{2p}\right\}| \\
&\leq CT\frac{\chi_0^2\epsilon^2}{p^3} \ ,
\end{aligned}$$

where the last inequality used Taylor's theorem to bound $\left\{(p-\omega_\epsilon(p))\frac{T}{\epsilon^2} + \frac{\chi_0 T}{2p}\right\}$ and the fact that for $p \geq \epsilon^{1/2}$, $\frac{\epsilon^3}{p^3} << 1$. Inserting this estimate into the first of the integral terms in (43) we see that it is bounded by

$$CT\int_{\epsilon^{1/2}}^\infty \frac{\chi_0^2\epsilon^2}{p^2\omega_\epsilon(p)}\hat{\mathcal{U}}_0^c(\omega_\epsilon(p))dp \leq CT\sqrt{\epsilon}|\hat{\mathcal{U}}_0^c|_{L^\infty} \ .$$

A similar estimate holds for the remaining term in the definition of $\Delta\mathcal{U}$ and Proposition 2.5 follows.

4. Nonlinear pulse dynamics

Summarizing the results of the previous section we now know (rigorously) that if we ignore nonlinear effects we can approximate the motion of a short pulse injected into one end of an optical fiber by the linearized short-pulse equation (8). Since nonlinear effects are important for a lot of optical phenomena it is now the next step to investigate, how incorporating nonlinear terms into the polarization affects (8). Here, we consider as a first step the most simple form of a nonlinear contribution p_{nl} given by

$$p_{\text{nl}} = \chi_3 u^3. \tag{44}$$

In order to answer this question we start from a wave equation similar to (6) with an additional nonlinear term

$$u_{xx} = u_{tt} + \chi_0 u + \chi_3 (u^3)_{tt}. \tag{45}$$

Remark 4.1. *The equation (45) can be derived from Maxwell's wave equation*

$$u_{xx} - u_{tt} = (p_{\text{lin}})_{tt} + (p_{\text{nl}})_{tt}$$

writing the linear part of the polarization as

$$p_{\text{lin}}(x,t) = \int \chi(t-\tau) u(x,\tau) d\tau$$

and making the approximation

$$\hat{\chi}(\omega) = -\frac{\chi_0}{\omega^2 - i\Gamma\omega - \omega_0^2} \approx -\frac{\chi_0}{\omega^2}.$$

This means physically that the frequency range of the pulse under consideration is far from the resonance frequency of the material. It is also possible to consider other forms of the susceptibility in frequency domain, leading to different types of equations for the short pulse. If we ignore the nonlinear term then Proposition 2.4 shows that this approximation leads to a small error. The choice of nonlinear part of the susceptibility corresponds to

$$p_{\text{nl}}(x,t) = \int \chi^{(3)}(t-\tau_1, t-\tau_2, t-\tau_3) u(x,\tau_1) u(x,\tau_2) u(x,\tau_3) d\tau_1 d\tau_2 d\tau_3$$

with the assumption that the nonlinear contribution is instantaneous, hence

$$\chi^{(3)}(t-\tau_1, t-\tau_2, t-\tau_3) = \chi_3 \delta(t-\tau_1)\delta(t-\tau_2)\delta(t-\tau_3).$$

On the basis of formal asymptotic calculations we believe that this is the most important contribution to the nonlinearity and thus in the present paper we limit ourself to the consideration of this case. However, we also stress that, for ultra-short pulses, it is very interesting to extend the analysis to more complicated forms of the nonlinear susceptibility [12] *including delay in the response of the material.*

Based on the results of the previous section, it is reasonable to assume that for short pulses we can approximate the solution of (45) by an ansatz of the form

$$u(x,t) = \epsilon \mathcal{U}_0\left(\frac{t-x}{\epsilon}, \epsilon x\right) + \epsilon^2 \mathcal{U}_1\left(\frac{t-x}{\epsilon}, \epsilon x\right) + \ldots \tag{46}$$

into (45). The equation on the leading nontrivial order for $\mathcal{U}_0$ is then given by

$$-2\partial_\phi \partial_X \mathcal{U}_0 = \chi_0 \mathcal{U}_0 + \chi_3 \partial_{\phi\phi} (\mathcal{U}_0)^3, \qquad X = \epsilon x. \tag{47}$$

This nonlinear short-pulse equation describes the influence of the nonlinear contribution of the polarization to the pulse.

As mentioned in the introduction, the standard model describing the nonlinear pulse evolution is the cubic nonlinear Schrödinger equation (NLSE). To emphasize the different regimes to which the NLSE and short pulse equations apply, we briefly review how one derives (47) from (45). The main idea is, that we assume a broad, rather than a short, pulse in the sense that we introduce time scales that are *slower* than the oscillations of the carrier wave that is oscillating at a fixed frequency $\tilde{\omega}$ with a wavenumber $\tilde{\beta}$. Therefore, the NLSE is an equation describing the slowly varying amplitude of the optical signal.

The separation of those time scales can be done by a usual expansion in multiple scales. In the present work, however, we utilize the so-called renormalization group (RG) method to derive the NLSE. This perturbative technique was first developed by Chen, Goldenfeld, and Oono as a tool for asymptotic analysis (see [13] and [14]). In [14], the validity of the RG method has been justified by applying to various examples of ordinary differential equations involving multiple scales, boundary layers and WKB analysis. See also, [15] and [16] for some examples of the rigorous use the of the renormalization group in the study of partial differential equations. The mathematical study of this method has been also presented by Ziane in [17]. The author explicitly described the RG method in the general setting of autonomous nonlinear systems of differential equations.

We will follow the approach given in [17] and explain how to obtain the perturbative solution of Maxwell's equation. Although the RG method will lead to the same results as in multiple scale technique, it is worth mentioning that there are advantages of using this method. First, the RG method does not require one to introduce all the different scales in the beginning of the ansatz since these will appear naturally in the *renormalization group equation.* This implies that one can assume a naive perturbation series in any given problems involving multiple scales. A second argument in favor of this method is that the algebraic calculations are simpler than when other perturbation techniques are used, especially when one considers higher order approximations.

To see how the RG method works in the present case, we start from a slightly more general form than (45)

$$u_{xx} = u_{tt} + \partial_{tt} \int \chi(t-\tau) u(x,\tau) d\tau + \chi_3 (u^3)_{tt}. \tag{48}$$

We assume that the solution of (48) is of small amplitude and concentrated around the carrier frequency. Because of the oscillations of the carrier wave, in the Fourier domain the signal will be concentrated around the frequencies $\tilde{\omega}$ and $-\tilde{\omega}$. Therefore, we can write our solution as a wave packet in the form of

$$u(x,t) = A(x,t) e^{i(\tilde{\beta}x - \tilde{\omega}t)} + A^*(x,t) e^{-i(\tilde{\beta}x - \tilde{\omega}t)}. \tag{49}$$

Taking the Fourier transform of (48), we find

$$\left(\frac{\partial^2}{\partial x^2} + \beta^2(\omega) \right) \hat{u}(x,\omega) = -\omega^2 \chi_3 \widehat{u^3}(x,\omega), \tag{50}$$

where the wavenumber β is given by

$$\beta(\omega) = \omega \sqrt{1 + \hat{\chi}(\omega)}. \tag{51}$$

The main idea is now to make a Taylor expansion of the dispersion β around the carrier frequency. This assumes that the signal is localized in Fourier domain, corresponding to a slowly varying amplitude approximation in time domain. Because of this *local* character of this expansion, the specific form of $\chi(\omega)$ is not essential. The Taylor expansion of β at $\tilde{\omega}$ yields

$$\beta^2(\omega) = \beta^2(\tilde{\omega}) + \left.\frac{\partial \beta^2}{\partial \omega}\right|_{\omega=\tilde{\omega}} (\omega - \tilde{\omega}) + \frac{1}{2}\left.\frac{\partial^2 \beta^2}{\partial \omega^2}\right|_{\omega=\tilde{\omega}} (\omega - \tilde{\omega})^2 + \cdots . \tag{52}$$

Let us denote $\tilde{\beta} = \beta(\tilde{\omega})$. Applying the inverse Fourier transform to (50), the straightforward calculation yields

$$\begin{aligned}
\frac{\partial^2}{\partial x^2} u(x,t) \quad &+ \quad \mathrm{e}^{i(\tilde{\beta}x - \tilde{\omega}t)} \sum_k \frac{1}{k!} \left.\frac{\partial^k \beta^2}{\partial \omega^k}\right|_{\omega=\tilde{\omega}} \left(i\frac{\partial}{\partial t}\right)^k A(x,t) \\
&+ \quad \mathrm{e}^{-(i\tilde{\beta}x - \tilde{\omega}t)} \sum_k \frac{1}{k!} \left.\frac{\partial^k \beta^2}{\partial \omega^k}\right|_{\omega=-\tilde{\omega}} \left(i\frac{\partial}{\partial t}\right)^k A^*(x,t) \\
&= \quad \chi_3 \frac{\partial^2}{\partial t^2} u^3(x,t).
\end{aligned} \tag{53}$$

Now we are ready to introduce a slow time t_1 by setting

$$t_0 = t, \qquad t_1 = \epsilon t, \tag{54}$$

and writing (49) as

$$u(x,t_0,t_1) = A(x,t_1) e^{i(\tilde{\beta}x - \tilde{\omega}t_0)} + A^*(x,t_1) e^{-i(\tilde{\beta}x - \tilde{\omega}t_0)}. \tag{55}$$

Here, notice that we do not separate the scales in the evolution variable x at this step. Now we use a small amplitude expansion of the function u

$$u = \epsilon u_0 + \epsilon^2 u_1 + \epsilon^3 u_2 + \cdots . \tag{56}$$

Solving now (53) order by order, we can determine the equation for $A(x,t_1)$ which describes the slowly varying amplitude of the electric field u. A detailed calculation can be found in the Appendix. Since we want to analyze the effect of nonlinearity, we need to explore the order $\mathcal{O}(\epsilon^3)$ where the effects of nonlinearity first occur. The result of the application of the RG method yields

$$u(x,t_0,t_1) = \epsilon \Gamma(x,t_1) e^{i(\tilde{\beta}x - \tilde{\omega}t_0)} + \epsilon \Gamma^*(x,t_1) e^{-i(\tilde{\beta}x - \tilde{\omega}t_0)} + \mathcal{O}(\epsilon^4) \tag{57}$$

where the function $\Gamma(x,t_1)$ satisfies the NLSE of the form (72b)

$$\begin{aligned}
\frac{\partial \Gamma(x,t_1)}{\partial x} \quad &= \quad -\epsilon \tilde{\beta}' \frac{\partial \Gamma(x,t_1)}{\partial t_1} \\
&\quad + \epsilon^2 i \left(-\frac{\tilde{\beta}''}{2} \frac{\partial^2}{\partial t_1^2} \Gamma(x,t_1) + \frac{3\chi_3}{2} \frac{\tilde{\omega}^2}{\tilde{\beta}} \Gamma(x,t_1) |\Gamma(x,t_1)|^2 \right).
\end{aligned} \tag{58}$$

In order to carry out a direct comparison between (47) and (58) the only remaining step is to specify in (58) the dispersion $\beta(\omega)$. We calculate this from the assumed form of the susceptability, $\hat{\chi}(\omega) = -\chi_0/\omega^2$. In this case, (51) yields

$$\tilde{\beta}' = \frac{\tilde{\omega}}{\sqrt{\omega^2 - \chi_0}}, \qquad \tilde{\beta}'' = \frac{-\chi_0}{(\omega^2 - \chi_0)^{3/2}}. \tag{59}$$

We expect on the basis of this derivation and extensive numerical and experimental evidence that for broad pulses the nonlinear Schrödinger equation is an excellent approximation but for ultra-short pulses, the nonlinear short-pulse equation should be a more appropriate approximation than NLSE. Intuitively this is clear by the scaling of the ansatz that was used in both derivations: In order to derive (47) we started from a pulse of the form $\mathcal{U}(t/\epsilon)$ whereas the NLSE describes the envelope on a slow time scale $A(x, \epsilon t)$. On the other hand, it is not clear how far we can push each of those perturbations. It would be interesting to compare (47) and (58) analytically to the solution of Maxwell's equations given by (45), but this is an extremely complicated problem. Therefore, we approach this problem numerically.

5. Numerical comparison of the approximations to Maxwell's equations

We perform the following (numerical) experiment: Consider Maxwell's equation (45) with an initial data that corresponds to pulse carried by a carrier wave:

$$u(x=0,t) = a\,\mathrm{e}^{-b^2t^2/2}\ \cos\omega_0 t. \tag{60}$$

Here, we choose x to be our evolution variable and choose $u_x(x=0,t)$ such that our initial conditions correspond to a forward traveling wave of the linear problem. The factor a in (60) corresponds to the amplitude of the pulse and the parameter b determines the pulse width. For the susceptibility $\chi(\omega)$ we use in the numerical simulations the following form

$$\hat{\chi}(\omega) = -\frac{\chi_0}{\omega^2}\left(\mathcal{H}(\omega-\omega_c) + \mathcal{H}(-\omega_c-\omega)\right) \tag{61}$$

with $\mathcal{H}(x) = 1$ for $x \le 0$ and $\mathcal{H}(x) = 0$ elsewhere. This accounts for the fact that the pulse cannot propagate for very low frequencies. We assume that the amplitude of the pulse decays to almost zero at ω_c. This condition is satisfied by an appropriate choice of the carrier frequency $\omega_0 > \omega_c$. First, we discuss the "classical" case for small b, where the NLSE applies. To apply the NLSE model, we first extract out of (60) the corresponding initial condition for the slowly varying amplitude, compute the evolution of this initial data according to (58) and then use (57) in order to reconstruct the electric field $u(x_{\mathrm{end}}, t)$. Figure 1 shows the typical result of the case where the nonlinear Schrödinger approximation holds. Due to the choice of $b = 0.2$ the width of the pulse is sufficiently large in comparison to the period of the carrier frequency. Since we are interested in the *nonlinear* evolution of the signal, we present the result of the NLSE approximation in the following form: First, we propagate the initial pulse in the linear Maxwell's equations by setting $\chi_3 = 0$. This solution u_{lin} serves as reference data: Now we propagate the same pulse in the corresponding nonlinear setting and, after obtaining u_{maxwell}, we compute the difference from the linear solution

$$\Delta u_{\mathrm{maxwell}} = u_{\mathrm{maxwell}} - u_{\mathrm{lin}}. \tag{62}$$

Then we find by solving the NLSE the corresponding u_{nlse} and compute

$$\Delta u_{\mathrm{nlse}} = u_{\mathrm{nlse}} - u_{\mathrm{lin}}. \tag{63}$$

The question is now how well Δu_{nlse} approximates $\Delta u_{\mathrm{maxwell}}$. As we can see from figure 1, in this case, the approximation is excellent. Let us now increase b corresponding to making the pulse shorter. Let's first choose $b = 2$. Notice that this parameter already formally violates the basic assumption about separation of time scales made in the derivation of the NLSE. Therefore, from figure 2, it is surprising how well the NLSE still works. In this parameter regime, the short-pulse equation

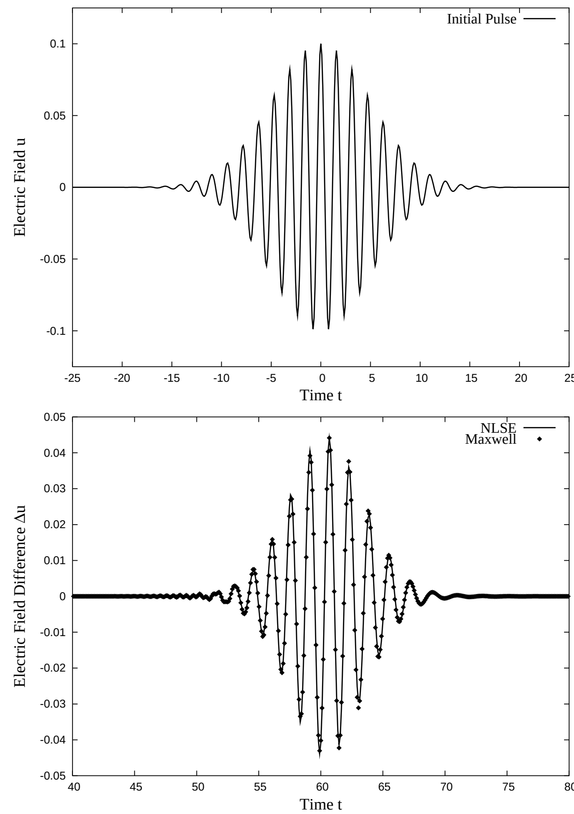

FIGURE 1. Comparison of the solution of Maxwell's equation and the cubic nonlinear Schrödinger equation. The figure above shows the initial pulse at $x = 0$. The figure below compares the difference of nonlinear Maxwell's equation and the solution of the corresponding linear problem to the prediction of the cubic nonlinear Schrödinger equation. The total propagation distance is $x_{\text{end}} = 50$. The parameters of the simulation are $a = 0.1$, $b = 0.2$, $\omega_c = 2.5$, $\omega_0 = 4$, $\chi_0 = 5$ and $\chi_3 = 0.5$. Eq. (59) yields $\tilde{\beta}' \approx 1.2$.

is not better than the NLSE since for $b = 2$ the essential assumption about a t/ϵ dependence of the initial condition is not satisfied. It is possible to extend the validity of NLSE to shorter pulses by incorporating higher order terms. In the Appendix, we also give a derivation for the next order that appears in the RG expansion. In this article, however, we want to focus on the comparison between the leading order approximation of Maxwell's equations and the ultra-short pulse equation. Going to shorter pulses, e.g. for $b = 3.0$, already, the short-pulse equation starts to do a better job than the NLSE as we can see from Figure 2. Setting $b = 5$, we finally arrive to the domain of ultra-short pulses. Here, the NLSE still predicts the rough shape of the pulse, but does not give correct information about the pulse shape as we can see from figure 3. On the other hand we can see now that the short-pulse equation already in this chosen case of $b = 5.0$ corresponding to $\epsilon = 0.2$ provides an excellent approximation of Maxwell's equation. Note that we propagated till

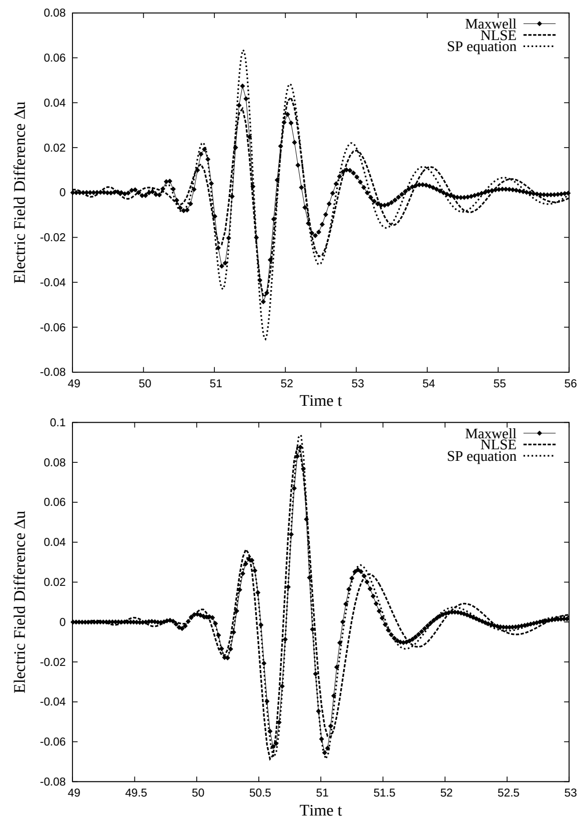

FIGURE 2. Comparison of the solution of Maxwell's equation and the cubic nonlinear Schrödinger equation and the short-pulse equation for a short pulse. Again, the figures compare the difference of nonlinear Maxwell's equation and the solution of the corresponding linear problem to the prediction of the cubic nonlinear Schrödinger equation and the short-pulse equation. The parameters for this simulation are the same as in figure 1 with the exception of $b = 2$, $\omega_0 = 6.5$ for the figure above and $b = 3$, $\omega_0 = 13$ for the figure below.

$x_{\text{end}} = 50 \sim \mathcal{O}\left(\frac{1}{\epsilon^2}\right)$. This numerical experiment cannot substitute for a more thorough analytical investigation, but it is an encouraging sign that, for ultra-short pulses, (47) can be used in order to approximate the solutions of (45). In further numerical studies we found that even at propagation distances up to $x_{\text{end}} = 200$, the short-pulse equation is a good approximation to Maxwell's equations.

6. Conclusion

In this paper we presented two main results: First we showed that if we ignore nonlinear effects one can *rigorously* approximate the evolution of a very short pulse injected into one end of an optical fiber by a solution of the (linear) short-pulse equation (8). Second, in the case of nonlinear pulse propagation, we have shown numerically that the (nonlinear) short-pulse equation gives an excellent approximation to the solution of Maxwell's equation. The analytical proof that those two

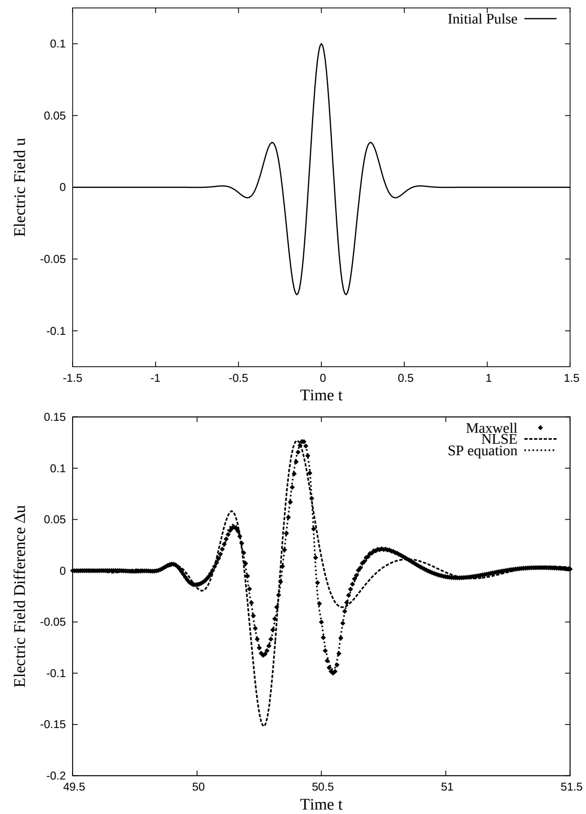

FIGURE 3. Comparison of the solution of Maxwell's equation and the short-pulse equation and the NLSE for an ultra-short pulse. Again, the figure above shows the initial pulse at $x = 0$ and the figure below compares the difference of nonlinear Maxwell's equation and the solution of the corresponding linear problem to the prediction of short-pulse equation and NLSE. The parameters for this simulation are the same as in figure 1 with the exception of $b = 5$ and $\omega_0 = 20$.

solutions stay close, at least for an evolution up to $\mathcal{O}(1/\epsilon)$ is a challenging task and subject to future research.

Appendix A. Bounds on the roots of Q

We begin by proving the results in Lemmas 3.3 and 3.4 on the large p asymptotics of the roots of Q. We give all the details of the calculations needed to prove Lemma 3.3 and leave the very similar details of the proof of Lemma 3.4 to the reader. Begin by defining $\sigma^0_\pm$ by

$$s^0_\pm = -\frac{\epsilon}{2}\left(\Gamma \pm \sqrt{\Gamma^2 - 4\omega_0^2}\right) + \epsilon\sigma^0_\pm \ . \tag{64}$$

Then $s_\pm^0$ is a root of $Q(s;p,\epsilon)$ if and only if $\sigma_\pm^0$ is a solution of

$$g(\sigma_\pm^0;p,\epsilon) = (\sigma_\pm^0)^2 - \sigma_\pm^0\left(\pm\sqrt{\Gamma^2-4\omega_0^2}\right) + \frac{\epsilon^2\chi_0\left(-\frac{1}{2}\left(\Gamma\pm\sqrt{\Gamma^2-\omega_0^2}+\sigma_\pm^0\right)\right)^2}{p^2}$$
$$+\frac{\epsilon^2\left((\sigma_\pm^0)^2-\sqrt{\Gamma^2-4\omega_0^2}\right)\left(-\frac{1}{2}\left(\Gamma\pm\sqrt{\Gamma^2-4\omega_0^2}\right)+\epsilon\sigma_\pm^0\right)^2}{p^2} = 0\ .$$

If we let $\xi = p^{-2}$, and define $\tilde{q}(\sigma_\pm^0;\xi,\epsilon) = q(\sigma_\pm^0;p,\epsilon)$ then we see that $\tilde{q}(0;0,0)=0$ and $\partial_\sigma\tilde{q}(0;0,0) = -\sqrt{\Gamma^2-4\omega_0^2}\neq 0$, so the implicit function theorem implies that there exists a smooth function $\sigma_\pm^0(\xi,\epsilon)$ such that $\tilde{q}(\sigma_\pm^0(\xi,\epsilon);\xi,\epsilon)=0$. Note that there exist constants A, B, C, and D which depend on Γ, ω_0, and χ_0, but not on ϵ or p such that if we rearrange the equation $\tilde{q}(\sigma_\pm^0(\xi,\epsilon);\xi,\epsilon)=0$, (and replace ξ by p^{-2} it can be written as:

$$|\sigma_\pm^0(\xi,\epsilon)| \leq (1-A\epsilon^2/p^2)^{-1}\left(\frac{B\epsilon^2}{p^2}+C|\sigma_\pm^0(\xi,\epsilon)|^2+\frac{D\epsilon^4}{p^2}|\sigma_\pm^0(\xi,\epsilon)|^4\right)\ . \tag{65}$$

Since $\sigma_\pm^0(\xi=0,\epsilon=0)=0$, (65) immediately implies that there exists $C_0, C_1>0$ such that for $p>C_0\epsilon$,

$$|\sigma_\pm^0(\xi,\epsilon)| \leq \frac{C_1\epsilon^2}{p^2}\ , \tag{66}$$

which completes the proof of Lemma 3.3. □

We now prove Lemma 3.2 We first note that if $p\neq 0$, Q has no purely imaginary eigenvalues. This follows by assuming that there exists such an eigenvalue – say $s=ix$, for $x\in\mathbb{R}$. Inserting this into Q and equating real and imaginary parts we see that x must satisfy

$$\begin{aligned}(\epsilon^2\omega_0^2-x^2)(p^2-x^2)-\epsilon^2x^2\chi_0 &= 0\\ \epsilon x\Gamma(p^2-x^2) &= 0\ .\end{aligned} \tag{67}$$

From the second of these equations we see that either $x=0$ or $x=\pm p$. However, neither of these values of x solves the first equation and hence there are no pure imaginary roots. Next note that from Lemmas 3.3 and 3.4 we know that for p sufficiently large, all four roots lie in the left half plane. But since the roots vary continuously with p, the only way we could obtain a root with positive real part was is one of the roots passed through the imaginary axis. We have just seen that there are no pure imaginary roots for any non-zero value of p and hence we never have a root with positive real part. □

Finally, we prove Lemma 3.5. Note that if we define $s=\epsilon\tilde{s}$ and $q=p/\epsilon$, then $Q(s;p,\epsilon)=\tilde{Q}(\tilde{s},q,\epsilon)$, with $\tilde{Q}(\tilde{s},q,\epsilon)=(\tilde{s}^2+\tilde{s}^2\Gamma+\omega_0^2)(\tilde{s}^2+q^2)+\tilde{s}^2\chi_0$, and the fact that the roots of Q can be written as $s_\pm^{0,1}(p)=\epsilon\tilde{s}_\pm^{0,1}(p/\epsilon)$ follows. Next note that for q small, an easy perturbative argument shows that $\tilde{Q}$ has a pair of complex conjugate roots $\tilde{s}_\pm^0$ of size $\mathcal{O}(1)$ (which correspond to the roots $s_\pm^0$ of Q) and a pair of complex conjugate roots $\tilde{s}_\pm^1=\pm iq+\mathcal{O}(q^2)$. Thus, Lemma 3.5 will follow if we can show that for all values of q the roots of $\tilde{Q}$ are distinct. Since the coefficients of $\tilde{Q}$ are real, the only way it can have multiple roots is if there is a multiple root on the (negative) real axis, or if there is a double complex root $\tilde{s}$ and a second double root equal to the complex conjugate of $\tilde{s}$. We can immediately rule out the first possibility by noting that

$$\tilde{Q}(\tilde{s},q,\epsilon)=\tilde{s}^2(\tilde{s}^2+\tilde{s}\Gamma+\omega_0^2+\chi_0)+q^2(\tilde{s}^2+\tilde{s}\Gamma+\omega_0^2)\ .$$

But since we have assumed that $\Gamma^2<4\omega_0^2$, $(\tilde{s}^2+\tilde{s}\Gamma+\omega_0^2)>0$ and $(\tilde{s}^2+\tilde{s}\Gamma+\omega_0^2+\chi_0)>0$ (for real values of $\tilde{s}$) and hence $\tilde{Q}>0$ for all real values of $\tilde{s}$. To rule out the possibility that the roots of $\tilde{Q}$ are of the form $\tilde{s}$ and $\overline{\tilde{s}}$ assume that that one can factor $\tilde{Q}=(s-\tilde{s})^2(s-\overline{\tilde{s}})^2$. Expanding this expression and equating coefficients of like powers of s with the expression for $\tilde{Q}$ above one finds a similar contradiction. □

Appendix B. Details on the derivation of NLSE by renormalization group method

In this appendix we show how to solve (53) order by order. Without any loss of generality we assume $\chi_3=1$. We first collect terms $\mathcal{O}(\epsilon)$ and find

$$\left(\frac{\partial^2}{\partial x^2}+\tilde{\beta}^2\right)u_0(x,t_0,t_1)=0.$$

Since we assume the solution has the form of (55), we obtain

$$u_0(x,t_0,t_1)=A_0(t_1)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}+A_0^*(t_1)e^{-i(\tilde{\beta}x-\tilde{\omega}t_0)},$$

where $A_0(t_1)$ can be determined from the initial condition for (48). Now, at the second order, $\mathcal{O}(\epsilon^2)$, we find

$$\begin{aligned}\left(\frac{\partial^2}{\partial x^2}+\tilde{\beta}^2\right)u_1(x,t_0,t_1)\quad &+\quad \left.\frac{d\beta^2}{d\omega}\right|_{\omega=\tilde{\omega}}\left(i\frac{d}{dt_1}A_0(t_1)\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}\\ &+\quad \left.\frac{d\beta^2}{d\omega}\right|_{\omega=-\tilde{\omega}}\left(i\frac{d}{dt_1}A_0^*(t_1)\right)e^{-i(\tilde{\beta}x-\tilde{\omega}t_0)}=0.\end{aligned}\quad (68)$$

Solving the above differential equation, we find the solution $u_1(x,t_0,t_1)$,

$$u_1(x,t_0,t_1)=A_1(t_1)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}+A_1^*(t_1)e^{-i(\tilde{\beta}x-\tilde{\omega}t_0)}+Parti(u_1),$$

where $A_1(t_1)$ depends on the initial condition and we denote $Parti(u_1)$ by the particular solution of the given equation (68). One of the simple ways to find this particular solution for the given equation is to assume that

$$Parti(E_1)=e^{i\tilde{\beta}x}a(t_0,t_1)x+e^{-i\tilde{\beta}x}b(t_0,t_1)x,$$

where $a(t_0,t_1),b(t_0,t_1)$ are to be determined later. We plug this into the equation (68) then it follows that

$$Parti(u_1)=-\tilde{\beta}'(\frac{dA_0(t_1)}{dt_1}xe^{i(\tilde{\beta}x-\tilde{\omega}t_0)}+\frac{dA_0^*(t_1)}{dt_1}xe^{-i(\tilde{\beta}x-\tilde{\omega}t_0)}).$$

Finally, we obtain the second order approximated solution

$$\begin{aligned}u^{(2)}(x,t_0,t_1)\quad &=\quad \epsilon(u_0+\epsilon u_1)\\ &=\quad \epsilon\left(A_0(t_1)+\epsilon A_1(t_1)-\epsilon\tilde{\beta}'\frac{dA_0(t_1)}{dt_1}x\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}\\ &+\quad \text{complex conjugate.}\end{aligned}$$

Letting $\tilde{A}_0(t_1) = A_0(t_1) + \epsilon A_1(t_1)$, we find

$$\begin{aligned} u^{(2)}(x, t_0, t_1) &= \epsilon \left(\tilde{A}_0(t_1) - \epsilon \tilde{\beta}' \frac{d\tilde{A}_0(t_1)}{dt_1} x \right) e^{i(\tilde{\beta} x - \tilde{\omega} t_0)} + \mathcal{O}(\epsilon^3) \\ &+ \text{complex conjugate.} \end{aligned}$$

Since $u^{(2)}(x, t_0, t_1)$ needs to be an approximation valid up to order $\mathcal{O}(\epsilon^2)$, the term $\mathcal{O}(\epsilon^3)$ can be neglected. Hence, we have

$$\begin{aligned} u^{(2)}(x, t_0, t_1) &= \epsilon \left(\tilde{A}_0(t_1) - \epsilon \tilde{\beta}' \frac{d\tilde{A}_0(t_1)}{dt_1} x \right) e^{i(\tilde{\beta} x - \tilde{\omega} t_0)} \\ &+ \text{complex conjugate.} \end{aligned}$$

Here, we notice that a secular term appears on the above expression, which corresponds to the term proportional to x. In other words, this approximation is no longer valid when $x \sim \mathcal{O}\left(\frac{1}{\epsilon}\right)$ or higher. In order to get rid of this secular term, we look at the term $\tilde{A}_0(t_1) - \epsilon \tilde{\beta}' \frac{d\tilde{A}_0(t_1)}{dt_1} x$ as the Taylor expansion of order 1 of some function $\Lambda(x, t_1)$ about $x = 0$. Thus, we need to find $\Lambda(x, t_1)$ which satisfies that

$$\Lambda(x, t_1)|_{x=0} = \tilde{A}_0(t_1), \tag{69a}$$

$$\frac{\partial \Lambda(x, t_1)}{\partial x} = -\epsilon \tilde{\beta}' \frac{\partial \Lambda(x, t_1)}{\partial t_1}. \tag{69b}$$

This is the *renormalization group equation.* Now the above form of the equation motivates us to introduce a different scale ϵx. Let us define $x_1 = \epsilon x$ then (69) gives

$$\Lambda(x_1, t_1)|_{x_1=0} = \tilde{A}_0(t_1), \tag{70a}$$

$$\frac{\partial \Lambda(x_1, t_1)}{\partial x_1} = -\tilde{\beta}' \frac{\partial \Lambda(x_1, t_1)}{\partial t_1}. \tag{70b}$$

By solving (70b) provided that the initial condition (70a) is satisfied, we can express $\tilde{A}_0(t_1) - \tilde{\beta}' \frac{d\tilde{A}_0(t_1)}{dt_1} x_1$ as the Taylor expansion of order 1 of $\Lambda(x_1, t_1)$ about $x_1 = 0$. Hence, we finally obtain the second order approximate solution

$$u^{(2)}(x, x_1, t_0, t_1) = \epsilon \Lambda(x_1, t_1) e^{i(\tilde{\beta} x - \tilde{\omega} t_0)} + \text{complex conjugate.}$$

Following the similar steps, let us now find the third order approximation. First, we need to collect $\mathcal{O}(\epsilon^3)$ terms. The usual way of obtaining these is simply plugging the previous ansatz $u = \epsilon u_0 + \epsilon^2 u_1 + \epsilon^3 u_2 + \cdots$ into (53) and collect appropriate terms. This will, however, lead to highly complicated algebraic calculation. We now approach this problem by assuming a different ansatz. Since we have already obtained the second order approximation of the solution, we assume that

$$u = u^{(2)} + \epsilon^3 u_2 + \cdots .$$

For the nonlinear part of the equation (53) we collect $\mathcal{O}(\epsilon^3)$ terms which will give rise to secularities. To do this, we note that

$$\begin{aligned}
\frac{\partial^2}{\partial t^2}u(x,t)^3 &= -\epsilon^3 \mathrm{e}^{3i(\tilde{\beta}x-\tilde{\omega}t)}\left(9\tilde{\omega}^2+6\tilde{\omega}i\frac{\partial}{\partial t}\right)A(x,t)^3 \\
&- \epsilon^3 \mathrm{e}^{i(\tilde{\beta}x-\tilde{\omega}t)}\left(3\tilde{\omega}^2+2\tilde{\omega}i\frac{\partial}{\partial t}\right)|A(x,t)|^2A(x,t) \\
&- \epsilon^3 \mathrm{e}^{-i(\tilde{\beta}x-\tilde{\omega}t)}\left(3\tilde{\omega}^2+2\tilde{\omega}i\frac{\partial}{\partial t}\right)|A(x,t)|^2A^*(x,t) \\
&- \epsilon^3 \mathrm{e}^{-3i(\tilde{\beta}x-\tilde{\omega}t)}\left(9\tilde{\omega}^2+6\tilde{\omega}i\frac{\partial}{\partial t}\right)A^*(x,t)^3. \qquad (71)
\end{aligned}$$

and see that the terms that will lead to resonances are the terms proportional to $\mathrm{e}^{i(\tilde{\beta}x-\tilde{\omega}t)}$ and $\mathrm{e}^{-i(\tilde{\beta}x-\tilde{\omega}t)}$. Thus, at the third order $O(\epsilon^3)$, we have

$$\begin{aligned}
\left(\frac{\partial^2}{\partial x^2}\right. &+ \left.\tilde{\beta}^2\right)u_2(x,t_0,t_1) \\
&+ \left(-\tilde{\beta}\tilde{\beta}''\frac{\partial^2}{\partial t_1^2}\Lambda(x_1,t_1)+3\tilde{\omega}^2\Lambda(x_1,t_1)|\Lambda(x_1,t_1)|^2\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)} \\
&+ \left(-\tilde{\beta}\tilde{\beta}''\frac{\partial^2}{\partial t_1^2}\Lambda^*(x_1,t_1)+3\tilde{\omega}^2\Lambda^*(x_1,t_1)|\Lambda(x_1,t_1)|^2\right)e^{-i(\tilde{\beta}x-\tilde{\omega}t_0)}=0.
\end{aligned}$$

We solve the above differential equation and find $u_2(x,t_0,t_1)=A_2(t_1)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}+Parti(u_2)+$ complex conjugate, where

$$Parti(u_2)=-ix\left(\frac{\tilde{\beta}''}{2}\frac{\partial^2}{\partial t_1^2}\Lambda(x_1,t_1)-\frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}\Lambda(x_1,t_1)|\Lambda(x_1,t_1)|^2\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}.$$

Therefore, the third order approximate solution is

$$\begin{aligned}
&u^{(3)}(x,x_1,t_0,t_1)=\epsilon\Lambda(x_1,t_1)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)} \\
&\quad+\epsilon^3\left(A_2(t_1)-ix\left(\frac{\tilde{\beta}''}{2}\frac{\partial^2}{\partial t_1^2}\Lambda(x_1,t_1)-\frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}\Lambda(x_1,t_1)|\Lambda(x_1,t_1)|^2\right)\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)} \\
&\quad+\text{complex conjugate.}
\end{aligned}$$

In order to obtain all the possible secular terms, we rewrite the above equation using (70a), (70b) and the Taylor expansion of order 1 for $\Lambda(x_1,t_1)$ about $x=0$. Then we find

$$\begin{aligned}
&u^{(3)}\sim\epsilon\left(\Lambda(x=0)-\epsilon z\tilde{\beta}'\frac{\partial\Lambda(x_1,t_1)}{\partial t_1}\right. \\
&\quad\left.+\epsilon^2\left(A_2(t_1)+ix\left(-\frac{\tilde{\beta}''}{2}\frac{\partial^2}{\partial t_1^2}\Lambda(x_1,t_1)+\frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}\Lambda(x_1,t_1)|\Lambda(x_1,t_1)|^2\right)\right)\right)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)} \\
&\quad+\text{complex conjugate.}
\end{aligned}$$

Again, to get rid of the secular terms, we now need to find $\Gamma(x_1,t_1)$ satisfying that

$$\begin{aligned}
\Gamma(x=0) &= \Lambda(x=0)+\epsilon^2 A_2(t_1), && \text{(72a)}\\
\frac{\partial\Gamma(x_1,t_1)}{\partial x} &= -\epsilon\tilde{\beta}'\frac{\partial\Gamma(x_1,t_1)}{\partial t_1} && \text{(72b)}\\
&\quad +\epsilon^2 i\left(-\frac{\tilde{\beta}''}{2}\frac{\partial^2}{\partial t_1^2}\Gamma(x_1,t_1)+\frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}\Gamma(x_1,t_1)|\Gamma(x_1,t_1)|^2\right).
\end{aligned}$$

This leads us to introduce a new scale $x_2=\epsilon^2 x$ in addition to $x_1=\epsilon x$. Then it follows that

$$\frac{\partial}{\partial x}=\epsilon\frac{\partial}{\partial x_1}+\epsilon^2\frac{\partial}{\partial x_2}.$$

Therefore, from (72b), we finally obtain

$$\frac{\partial\Gamma(x_1,x_2,t_1)}{\partial x_2}=i\left(-\frac{\tilde{\beta}''}{2}\frac{\partial^2}{\partial t_1^2}\Gamma(x_1,x_2,t_1)+\frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}\Gamma(x_1,x_2,t_1)|\Gamma(x_1,x_2,t_1)|^2\right).$$

This is the *cubic nonlinear Schrödinger equation*. Following the similar steps, we can extend the results to the higher order, $\mathcal{O}(\epsilon^4)$ approximation. First, we assume the ansatz, $u=u^{(3)}+\epsilon^4 u_3+\cdots$ and recall (71). Since $\frac{\partial}{\partial t}=\epsilon\frac{\partial}{\partial t_1}$, at the order $\mathcal{O}(\epsilon^4)$, we find

$$\begin{aligned}
\left(\frac{\partial^2}{\partial x^2}+\tilde{\beta}^2\right)u_3 &+ e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}\left(\frac{1}{3}\tilde{\beta}\tilde{\beta}'''(-i)\frac{\partial^3}{\partial t_1^3}\Gamma(x_1,x_2,t_1)\right.\\
&+ \left. i\left(6\tilde{\omega}-3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(\Gamma(x_1,x_2,t_1)|\Gamma(x_1,x_2,t_1)|^2\right)\right)\right)\\
&+ e^{-i(\tilde{\beta}x-\tilde{\omega}t_0)}\left(\frac{1}{3}\tilde{\beta}\tilde{\beta}'''i\frac{\partial^3}{\partial t_1^3}\Gamma^*(x_1,x_2,t_1)\right.\\
&- \left. i\left(6\tilde{\omega}-3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(\Gamma(x_1,x_2,t_1)|\Gamma(x_1,x_2,t_1)|^2\right)\right)\right)=0.
\end{aligned}$$

Solving the above differential equation, we find

$$\begin{aligned}
u_3(x,x_1,x_2,t_0,t_1) &= A_3(t_1)e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}-\frac{1}{2i\tilde{\beta}}xe^{i(\tilde{\beta}x-\tilde{\omega}t_0)}\left(\frac{1}{3}\tilde{\beta}\tilde{\beta}'''(-i)\frac{\partial^3}{\partial t_1^3}\Gamma(x_1,x_2,t_1)\right.\\
&+ \left. i\left(6\tilde{\omega}-3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(\Gamma(x_1,x_2,t_1)|\Gamma(x_1,x_2,t_1)|^2\right)\right)\right)\\
&+ \text{complex conjugate},
\end{aligned}$$

where $A_3(t_1)$ is a function of t_1 which depends on the initial condition. Hence, we obtain the fourth order approximate solution,

$$\begin{aligned} u^{(4)}(x, x_1, x_2, t_0, t_1) &= \epsilon\Gamma(x_1, x_2, t_1)e^{i(\tilde{\beta}x - \tilde{\omega}t_0)} \\ &+ \epsilon^4 e^{i(\tilde{\beta}x-\tilde{\omega}t_0)}\left(A_3(t_1) + x\left(\frac{1}{6}\tilde{\beta}\tilde{\beta}'''\frac{\partial^3}{\partial t_1^3}\Gamma(x_1, x_2, t_1)\right.\right. \\ &- \left.\left.\frac{1}{2\tilde{\beta}}\left(6\tilde{\omega} - 3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(\Gamma(x_1, x_2, t_1)|\Gamma(x_1, x_2, t_1)|^2\right)\right)\right)\right) \\ &+ \text{complex conjugate.} \end{aligned}$$

After we rewrite $\Gamma(x_1, x_2, t_1)$ as the Taylor expansion of order 1 about $x = 0$ using (72), it is now clear that we need to find $V(x_1, x_2, t_1)$ satisfying

$$V(x=0) = \Gamma(x=0) + \epsilon^3 A_3(t_1), \tag{73a}$$

$$\begin{aligned} \frac{\partial V}{\partial x} &= \epsilon(-\tilde{\beta}')\frac{\partial V}{\partial t_1} + \epsilon^2 i\left(-\frac{1}{2}\tilde{\beta}''\frac{\partial^2}{\partial t_1^2}V + \frac{3}{2}\frac{\tilde{\omega}^2}{\tilde{\beta}}V^2V^*\right) \\ &+ \epsilon^3\left(\frac{1}{6}\tilde{\beta}'''\frac{\partial^3}{\partial t_1^3}V - \frac{1}{2\tilde{\beta}}\left(6\tilde{\omega} - 3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(V|V|^2\right)\right)\right). \end{aligned} \tag{73b}$$

Introducing $x_3 = \epsilon^3 x$, from (73b) we find

$$\begin{aligned} \frac{\partial V}{\partial x_3}(x_1, x_2, x_3, t_1) &= \frac{1}{6}\tilde{\beta}'''\frac{\partial^3}{\partial t_1^3}V(x_1, x_2, x_3, t_1) \\ &- \frac{1}{2\tilde{\beta}}\left(6\tilde{\omega} - 3\frac{\tilde{\beta}'\tilde{\omega}^2}{\tilde{\beta}}\right)\left(\frac{\partial}{\partial t_1}\left(V(x_1, x_2, x_3, t_1)|V(x_1, x_2, x_3, t_1)|^2\right)\right). \end{aligned}$$

Since $\frac{\partial}{\partial t_1}V|V|^2$ includes $|V|^2\frac{\partial}{\partial t_1}V$ and $V\frac{\partial}{\partial t_1}|V|^2$, we see that the Raman scattering and the self-steepening terms appear at the higher order approximation.

Acknowledgments

This material is based upon work supported in part by the National Science Foundation under Grants No. 0073923 and DMS-0103915. TS and CEW also wish to thank Jeffrey Rauch for stimulating correspondence about the propagation of short pulses.

References

[1] A. Hasegawa and Yuki Kodama. *Solitions in optical communications*. Clarendon Press, Oxford, 1995.
[2] G. P. Agrawal. *Nonlinear Fiber Optics*. Academic Press, San Diego, 1995.
[3] A. C. Newell and J. V. Moloney. *Nonlinear Optics*. Addison-Wesley, Redwood City, CA, 1992.
[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] K. J. Blow and D. Wood. Theoretical description of transient stimulated Raman scattering in optical fibers. *IEEE J. Quant. Elect.*, 25:2665–2673, 1989.
[7] P. V. Mamyshev and S. V. Chernikov. Ultrashort-pulse propagation in optical fibers. *Opt. Lett.*, 15:1076–1078, 1990.

[8] T. Brabec and F. Krausz. Nonlinear optical pulse propagation in the single-cycle regime. *Phys. Rev. Lett.*, 78:3282–3285, 1997.
[9] T. Schäfer and C. E. Wayne. Propagation of ultra-short optical pulses in nonlinear media. *Preprint*, 2003.
[10] D. Alterman and J. Rauch. Diffractive short pulse asymptotics for nonlinear wave equations. *Phys. Lett. A*, 264:390–395, 2000.
[11] Deborah Alterman and Jeffrey Rauch. The linear diffractive pulse equation. *Methods Appl. Anal.*, 7(2):263–274, 2000. Cathleen Morawetz: a great mathematician.
[12] R. W. Boyd. *Nonlinear Optics*. Academic Press, Boston, 1992.
[13] L. Y. Chen, N. Goldenfeld, and Y. Oono. Renormalization group theory for global asymptotic analysis. *Phys. Rev. Lett*, 73:1311–1315, 1994.
[14] L. Y. Chen, N. Goldenfeld, and Y. Oono. The renormalization group and singular perturbations: Multiple-scales, boundary layers and reductive perturbation theory. *Phys. Rev. E*, 54:376–394, 1996.
[15] J. Bricmont, A. Kupiainen, and G. Lin. Renormalization group and asymptotics of solutions of nonlinear parabolic equations. *Comm. Pure Appl. Math.*, 47(6):893–922, 1994.
[16] J. L. Bona, K. S. Promislow, and C. E. Wayne. Higher-order asymptotics of decaying solutions of some nonlinear, dispersive, dissipative wave equations. *Nonlinearity*, 8(6):1179–1206, 1995.
[17] M. Ziane. On a certain renormalization group method. *J. Math. Phys.*, 41(5):3290–3299, 2000.