Derivation of the equation for an ultrashort pulse in a fibre

Nijhof, J.H.B.; Ferwerda, H.A.; Hoenders, B.J.

Published in:
Pure and Applied Optics

DOI:
10.1088/0963-9659/4/3/007

Publication date:
1995

Pure Appl. Opt. 4 (1995) 199–218. Printed in the UK

Derivation of the equation for an ultrashort pulse in a fibre

J H B Nijhof†, H A Ferwerda and B J Hoenders

Institute for Theoretical Physics, University of Groningen, Nijenborgh 4, 9747 AG Groningen, the Netherlands

Received 10 November 1994

Abstract. Pulses propagating in a non-linear dispersive (glass) fibre can be described by the non-linear Schrödinger equation if the pulse is longer than a picosecond; for shorter pulses, this equation must be extended. In this paper we systematically derive this extended equation using the method of multiple scales. By using an inherent freedom in the method of multiple scales, a technique is developed such that perturbation terms are greatly simplified. The limits of validity of the derived equation are discussed. It is shown to be valid for pulses longer than 30 fs.

1. Introduction

When one wants to propagate short pulses over long distances along fibres, pulse spreading because of group velocity dispersion becomes significant. However, this can be counteracted by non-linear dispersion. When both are present, the envelope of the pulse is governed by the non-linear Schrödinger equation (NLS). This equation has soliton solutions, that is, it has stable solutions that will propagate undeformed.

The soliton 'industry' started in 1967 when Gardner *et al* [11] discovered the inverse scattering method for solving the Korteweg–de Vries equation. Subsequently Zakharov and Shabat [30] extended the method and solved the non-linear Schrödinger equation (NLS) with it. Nowadays there are plenty of introductory texts on solitons, e.g. [7, 9, 10].

In 1973, Hasegawa and Tappert found that pulses in non-linear dispersive fibres could be described by the NLS, both in the case of anomalous dispersion [16] and in the case of normal dispersion [17], which meant that the fibres should be able to support soliton pulses. This was experimentally verified (in the case of anomalous dispersion) by Mollenauer, Stolen and Gordon in 1980 [25].

The NLS describes a pulse excellently for a pulse longer than a picosecond. For shorter pulses, atomic and molecular processes can no longer be considered to be instantaneous, and higher order terms become important. Most important of these is the stimulated Raman scattering (SRS) term, which leads to a down shift in the pulse frequency, as discovered by Mitschke and Mollenauer [22] and explained by Gordon [12].

Usually, the extended NLS, including these higher order terms, is derived making *ad hoc* assumptions (often, the second order derivative of the amplitude with respect to the propagation coordinate, A_{zz}, is neglected, like in [2, 21]; however, careful analysis shows that the first order contribution to A_{zz} is not negligible, but cancels with another

† e-mail address: nijhof@th.rug.nl.

0963-9659/95/030199+20 $19.50

neglected term). But it is possible to derive the extended NLS in a consistent manner, by using the *method of multiple scales*, in optics literature also known as the slowly varying envelope (or amplitude) approximation or as the rotating wave approximation. Kodama and Hasegawa [18, 19] were the first to give a proper derivation with the higher order terms included.

In this paper, the derivation of the propagation equation for the envelope of the pulse is presented, following (and extending) an approach of Newell and Moloney [27]. In this approach, functions of the frequency are not used at the centre frequency, but they are expanded like $f(\omega) \to f(\omega_1) + i\varepsilon f'(\omega_1)\,\partial_T + \cdots$. This expansion will be applied both to terms linear in the pulse amplitude, and to terms non-linear in the amplitude. It will be shown that this will lead to an automatic cancellation of some of the arising perturbation terms.

This technique should make it possible to apply the method of multiple scales rigorously and still leave parts of the propagation, notably the Raman part (see below) in an integrodifferential form.

A secondary advantage of the method presented in this paper is that one gets expressions for the perturbation to the transverse profile of the field almost for free. So in the case of a pulse in a slab or fibre, one can also see how the shape of the profile will differ from the shape of the linear mode.

Still, it turns out that for pulses of the order of ten femtoseconds, the term corresponding to the stimulated Raman scattering cannot be properly incorporated in a single partial differential equation for the envelope of the pulse. For such short pulses, the SRS term has to be taken into account either as a convolution product, leaving an integrodifferential equation [6, 21], or as a separate differential equation, leaving a system of differential equations, e.g. [1, 14]. These equations can be derived in a similar way.

General books on solitons in fibres are [2, 3, 15]. A general book on shorter pulses (femtosecond) is [4].

1.1. Assumptions

In general, a rigorous analysis of the propagation of a pulse through a slab or fibre is too difficult, so one has to make some assumptions.

First of all, the polarization is supposed to be local. That is, the polarization is allowed to depend on the past of the electric field, but it may only depend on the electric field at the same position.

Secondly, the material is supposed to be isotropic. Also, the fibre or waveguide should be single-mode and polarization-preserving. Without this assumption, mixing of modes could occur and one would have to consider coupled equations.

Today, pulse lengths shorter than 1 ps are of interest. However, a lower limit on the pulse length is necessary, or the asymptotic expansion will break down. In section 6 it is shown that for the derived equation to be valid, the pulse length must be greater than 30 fs.

The non-linear polarization will be treated in the Born–Oppenheimer approximation, i.e. the motion of the nuclei and the motion of the electrons will be treated separately. This simplifies the expression for the non-linear susceptibility.

Gordon–Haus jitter [13] or bandwidth limited amplification [5, 23] have not been included in the derivation.

1.2. Notational conventions

To keep in line with the most common convention, Fourier transform with respect to t is defined by $\mathscr{F}_t[f](\omega) = \int_{-\infty}^{\infty} f(t) \exp(i\omega t)\, dt$ and Fourier transform with respect to z is defined by $\mathscr{F}_z[f](k) = \int_{-\infty}^{\infty} f(z) \exp(-ikz)\, dt$. Therefore the inverse Fourier transforms are

$$\mathscr{F}_t^{-1}[g](t) = (2\pi)^{-1} \int_{-\infty}^{\infty} g(\omega) \exp(-i\omega t)\, d\omega$$

and

$$\mathscr{F}_z^{-1}[g](z) = (2\pi)^{-1} \int_{-\infty}^{\infty} g(k) \exp(ikz)\, dk$$

respectively. Fourier transform with respect to t is also denoted with a hat accent, $\hat{f}(\omega) = \mathscr{F}_t[f](\omega)$, and its inverse also with the check accent, $\check{g}(t)$.

Convolution products over time, written as $f * g$, are defined by $(f * g)(t) = \int_{-\infty}^{\infty} f(t - t_1) g(t_1)\, dt_1$.

A prime will always denote a derivative with respect to the frequency ω, never any other derivative or a real or imaginary part; so $f' = df/d\omega$ etc. A dot over a symbol denotes a derivative with respect to time.

The complex conjugate of a function is per definition given by $f^*(z) \equiv [f(z)]^*$.

SI units are used throughout.

2. Physical background

2.1. The Maxwell equations

The starting point is the Maxwell equations in a medium without free charges or currents,

$$\begin{aligned} \nabla \cdot \boldsymbol{D} &= 0 \qquad & \nabla \times \boldsymbol{E} &= -\partial_t \boldsymbol{B} \\ \nabla \cdot \boldsymbol{B} &= 0 \qquad & \nabla \times \boldsymbol{H} &= \partial_t \boldsymbol{D}. \end{aligned} \tag{1}$$

In paramagnetic and diamagnetic materials, the magnetic induction is proportional to the magnetic field. In good approximation $\boldsymbol{B} = \mu_0 \boldsymbol{H}$. The electric displacement $\boldsymbol{D}$ depends on the electric field as $\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$. The dependence of the polarization $\boldsymbol{P}$ on the electric field $\boldsymbol{E}$ will be discussed in section 2.2. This is where the non-linearity comes in.

Two of the four Maxwell equations can be combined to yield

$$-\nabla \times \nabla \times \boldsymbol{E} = \nabla^2 \boldsymbol{E} - \nabla(\nabla \cdot \boldsymbol{E}) = \mu_0 \partial_t^2 \boldsymbol{D}. \tag{2}$$

$\nabla \cdot \boldsymbol{D} = 0$ does not imply $\nabla \cdot \boldsymbol{E} = 0$, since the susceptibility does not have to be a constant.

2.2. The polarization

In general, the polarization at a certain time not only depends on the electric field at the same time, but also on the past of the electric field. However, the polarization is

supposed only to depend on the elecric field at the same position, and the material is supposed to be isotropic.

The polarization can be expanded in powers of the electric field,

$$\boldsymbol{P} = \boldsymbol{P}^{(1)} + \boldsymbol{P}^{(2)} + \boldsymbol{P}^{(3)} + \boldsymbol{P}^{(4)} + \cdots \tag{3}$$

with $\boldsymbol{P}^{(n)}$ homogeneous of degree n in $\boldsymbol{E}$. If two fields $\boldsymbol{E}_a$ and $\boldsymbol{E}_b$ satisfy $\boldsymbol{E}_b(t) = A\boldsymbol{E}_a(t)$ for all t, then the corresponding polarizations $\boldsymbol{P}_a^{(n)}$ and $\boldsymbol{P}_b^{(n)}$ will satisfy $\boldsymbol{P}_b^{(n)}(t) = A^n\boldsymbol{P}_a^{(n)}(t)$.

In isotropic materials, so in amorphous materials in particular, all even orders in the polarization are missing. Namely, when the electric field reverses sign (for all t), the polarization has to do the same.

Since each higher order of polarization is orders of magnitude smaller than the previous one, $\boldsymbol{P}^{(5)}$ and higher orders can usually be neglected. What remains is

$$\boldsymbol{P} = \boldsymbol{P}^{(1)} + \boldsymbol{P}^{(3)}. \tag{4}$$

$\boldsymbol{P}^{(1)}$ is the linear polarization, and (only) $\boldsymbol{P}^{(3)}$ will be called the non-linear polarization.

2.2.1. Linear polarization. Classically the linear polarization can be modelled by a set of damped harmonic oscillators [20]. In ω-domain, it is given by

$$\hat{\boldsymbol{P}}^{(1)} = \varepsilon_0 \hat{\chi}^{(1)}(\omega)\hat{\boldsymbol{E}}$$

with

$$\hat{\chi}^{(1)}(\omega) = \frac{Ne^2}{m\varepsilon_0 V}\sum_j \frac{f_j}{-\omega^2 - 2i\Gamma_j\omega + \omega_j^2} \tag{5}$$

where $f_j N/V$ is the number of oscillators with frequency ω_j per volume, m is the electron mass, and Γ_j is a damping constant. This expression includes both purely electronic contributions, with time scales of the order of 0.1 fs ($\omega_j \approx 2 \times 10^{16}$ Hz), and contributions involving nuclear oscillations, with time scales of the order of 10 fs ($\omega_j \approx 2 \times 10^{14}$ Hz) [2].

In the time domain, the polarization is given by a convolution product

$$\boldsymbol{P}^{(1)}(t, x) = \varepsilon_0 \tilde{\chi}^{(1)}(t, x) * \boldsymbol{E}(t, x) = \varepsilon_0 \int \tilde{\chi}^{(1)}(t_1, x)\boldsymbol{E}(t - t_1, x)\, dt_1. \tag{6}$$

In the following, the x-dependence of the fields and the susceptibility will be left implicit.

2.2.2. Non-linear polarization. Of the non-linear polarization, part is purely electronic in nature and part is hybrid nuclear–electronic. The time scale associated with the electronic part is of the order of 0.1 to 0.3 fs [2]. For pulses longer than 10 fs, this part of the non-linear polarization can be assumed to be instantaneous. As the medium is isotropic, this part must be of the form

$$\boldsymbol{P}^{(3)}_{\text{electronic}} = \varepsilon_0 \alpha \chi^{(3)}(\boldsymbol{E}(t)\cdot\boldsymbol{E}(t))\boldsymbol{E}(t). \tag{7}$$

α is the fraction of the non-linear polarization that is instantaneous.

Processes like stimulated Raman scattering, where the nuclear vibrations are involved, can be treated in the Born–Oppenheimer approximation, i.e. the motion of the nuclei and the motion of the electrons can be treated separately [8].

In the classical model of Placzek, this approximation is translated into the assumption that the motion of the nuclei enters into the polarization only through the polarizability [8, 29]. This, and the assumed isotropy, implies that the nuclear

contribution to the non-linear polarization has to be of the form

$$\boldsymbol{P}^{(3)}_{\text{nuclear}}(t) = \varepsilon_0(1-\alpha)\chi^{(3)}\boldsymbol{E}(t)\int g_R(t_1)(\boldsymbol{E}(t-t_1)\cdot\boldsymbol{E}(t-t_1))\,dt_1. \tag{8}$$

Equations (7) and (8) can be added to give the total non-linear polarization,

$$\boldsymbol{P}^{(3)}(t) = \varepsilon_0\boldsymbol{E}(t)\int \tilde{\chi}^{(3)}(t_1)(\boldsymbol{E}(t-t_1)\cdot\boldsymbol{E}(t-t_1))\,dt_1 \tag{9}$$

where $\tilde{\chi}^{(3)}$ is the sum of an electronic part and a Raman part as given by Blow and Wood [6],

$$\tilde{\chi}^{(3)}(t) = \chi^{(3)}\cdot(\alpha\delta(t) + (1-\alpha)g_R(t)). \tag{10}$$

If only one vibrational mode is important, with linewidth $1/\tau_2$ and eigenfrequency $1/\tau_1$, the Raman response function g_R is given by

$$g_R(t) = \frac{\tau_1^2+\tau_2^2}{\tau_1\tau_2^2}\,e^{-t/\tau_2}\sin(t/\tau_1) \qquad (t>0). \tag{11}$$

Blow and Wood [6] by fitting experimental results estimate $\alpha = 0.7$, $\tau_1 = 12.2$ fs and $\tau_2 = 32$ fs.

In the ω-domain, in principle the non-linear susceptibility is a function of three frequencies $\hat{\chi}^{(3)}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$; here the notational convention of Butcher and Cotter [8] has been followed, according to which ω_σ is the resulting frequency, $\omega_\sigma = \omega_1 + \omega_2 + \omega_3$.

However, by Fourier transforming (9), and comparing the result with the general expression for $\boldsymbol{P}^{(3)}$,

$$\boldsymbol{P}^{(3)}(t) = \int \hat{\chi}^{(3)}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)\,e^{-i\omega_\sigma t}\hat{\boldsymbol{E}}(\omega_1)\hat{\boldsymbol{E}}(\omega_2)\hat{\boldsymbol{E}}(\omega_3)\,d\omega_1\,d\omega_2\,d\omega_3$$

one finds that $\hat{\chi}^{(3)}$ is a function of the sum of the frequencies only:

$$\hat{\chi}^{(3)}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = \hat{\chi}^{(3)}(\omega_2+\omega_3). \tag{12}$$

It is conventional to write the susceptibility tensor in a symmetric form, thus instead of (12) one would have

$$\hat{\chi}^{(3)}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = \tfrac{1}{3}\hat{\chi}^{(3)}(\omega_1+\omega_2) + \tfrac{1}{3}\hat{\chi}^{(3)}(\omega_1+\omega_3) + \tfrac{1}{3}\hat{\chi}^{(3)}(\omega_2+\omega_3). \tag{13}$$

However, this leads to the same expression for $\boldsymbol{P}^{(3)}$ once the integrations over ω_1, ω_2 and ω_3 have been carried out. Therefore it is more convenient to use the simpler expression (12).

If $\tilde{\chi}^{(3)}$ is given by (10) and (11), then $\hat{\chi}^{(3)}(\omega)$ is given by

$$\hat{\chi}^{(3)}(\omega) = \chi^{(3)}\cdot(\alpha + (1-\alpha)\hat{g}_R(\omega)) = \chi^{(3)}\cdot\left(\alpha + (1-\alpha)\frac{1/\tau_1^2+1/\tau_2^2}{-\omega^2 - 2i\omega/\tau_2 + (1/\tau_1^2+1/\tau_2^2)}\right). \tag{14}$$

Now the total polarization is given by the sum of (6) and (10),

$$\boldsymbol{P} = \varepsilon_0\tilde{\chi}^{(1)} * \boldsymbol{E} + \varepsilon_0(\tilde{\chi}^{(3)} * (\boldsymbol{E}\cdot\boldsymbol{E}))\boldsymbol{E}. \tag{15}$$

Insert this in (2) and the equation to be solved is

$$\mathscr{L}E \equiv \nabla^2 E - \nabla\nabla\cdot E - \frac{1}{c^2}\partial_t^2(E + \tilde{\chi}^{(1)} * E) = \frac{1}{c^2}\partial_t^2(\tilde{\chi}^{(3)} * (E\cdot E)E). \qquad (16)$$

3. The method of multiple scales

We want to solve (16) using perturbation theory, treating the right-hand side as the perturbation. However, the simple iterative method, taking the first order solution plus perturbation terms, fails: secular terms appear which cause the perturbative corrections to grow linearly in time, which violates the (implicit) assumption that the perturbation terms give small corrections.

In the *method of multiple scales*, the first-order solution is allowed to vary slowly, which makes it possible to eliminate the secular terms that cause unbounded perturbations. The variant described below, called the derivative expansion method by Sturrock and Nayfeh, has been developed by, among others, Sturrock [28] and Nayfeh around 1960. According to Nayfeh [26], 'the method of multiple scales is so popular that it is being rediscovered just about every 6 months'. In optics the method is also known as the rotative wave approximation or the slowly varying amplitude (or envelope) approximation. The method will be explained using the pendulum as an example.

3.1. Example: the pendulum

Consider the 'real' pendulum, given by $\ddot{\phi} + \omega_1^2 \sin\phi = 0$. When the amplitude is small enough, ϕ can be approximated by $\phi = \frac{1}{2}a\,e^{-i\omega_1 t} + \frac{1}{2}a^*\,e^{i\omega_1 t}$.

For finite amplitudes, simply trying $\phi = \varepsilon(\frac{1}{2}a\,e^{-i\omega_1 t} + \frac{1}{2}a^*\,e^{i\omega_1 t}) + \varepsilon^2\phi^{(2)} + \varepsilon^3\phi^{(3)} + \cdots$ does not yield a better approximation. Namely, if one Taylor expands $\sin\phi$, and demands that all orders of ε are zero separately, at order ε^2 one finds $\varepsilon^2(\ddot{\phi}^{(2)} + \omega_1^2\phi^{(2)}) = 0$, the same equation as for $\phi^{(1)}$. So one can take $\phi^{(2)}$ to be zero (by replacing $\phi^{(1)}$ by $\phi^{(1)} + \varepsilon\phi^{(2)}$).

But at order ε^3

$$\varepsilon^3(\ddot{\phi}^{(3)} + \omega_1^2\phi^{(3)} - \tfrac{1}{48}\omega_1^2(a^3\,e^{-3i\omega_1 t} + 3a^2a^*\,e^{-i\omega_1 t} + \text{c.c.})) = 0.$$

So $\phi^{(3)} = -\frac{1}{384}a^3\,e^{3i\omega_1 t} + \frac{1}{32}i|a|^2a\omega_1 t\,e^{i\omega_1 t} + \text{c.c.}$ But the second term, the *secular term*, is linear in t. So it will become infinite as $t \to \infty$, whereas one would want the perturbation to remain small, or at least bounded.

A way out is to allow a to vary slowly in time, da/dt must be expanded in ε as well

$$\frac{da}{dt} = \varepsilon f^{(1)}(a, a^*) + \varepsilon^2 f^{(2)}(a, a^*) + \cdots.$$

Symbolically this is written as

$$\frac{da}{dt} = \varepsilon\frac{da}{dT} = \varepsilon\frac{\partial a}{\partial T_1} + \varepsilon^2\frac{\partial a}{\partial T_2} + \cdots.$$

The variables $T = T_1 = \varepsilon t$, $T_2 = \varepsilon^2 t$ are the *slow times*. One must be aware that the T_ns are not independent variables. In particular, the derivatives with respect to them do

not necessarily commute, $\partial_{T_1}\partial_{T_2}a = \partial_{T_1}f^{(2)} = f^{(1)}\partial_a f^{(2)} + f^{(1)*}\partial_{a^*}f^{(2)}$ may differ from $\partial_{T_2}\partial_{T_1} = \partial_{T_2}f^{(1)} = f^{(2)}\partial_a f^{(1)} + f^{(2)*}\partial_{a^*}f^{(1)}$.

The perturbation terms $\phi^{(n)}$ can contain terms with frequencies 0, ω_1, $2\omega_1$, up to $n\omega_1$. To separate these, $\phi^{(n)}$ is expanded again with respect to $e^{-i\omega_1 t}$. Thus

$$\phi^{(n)} = \sum_{m\geqslant 0} \operatorname{Re}\phi_m^{(n)} e^{-mi\omega_1 t} = \frac{1}{2}\sum_{m\geqslant 0}(\phi_m^{(n)} e^{-mi\omega_1 t} + \phi_m^{(n)*} e^{mi\omega_1 t}).$$

At order ε^2 one finds

$$\varepsilon^2 e^{i\phi}\left(-\frac{\partial a}{\partial T_1}\right) + \text{c.c.} = 0$$

so here the *compatibility condition*, i.e. the demand that the secular term is zero, is simply $\partial a/\partial T_1 = 0$.

Next, at order ε^3 the equation reads

$$(-4\omega_1^2\phi_3^{(3)} - \tfrac{1}{48}\omega_1^2 a^3) e^{-3i\omega_1 t} + \left(-\tfrac{1}{16}\omega_1^2|a|^2 a - i\omega_1\frac{\partial a}{\partial T_2}\right) e^{-i\omega_1 t} + \text{c.c.} = O(\varepsilon). \tag{17}$$

Note that though $\phi_1^{(3)}$ and $\phi_3^{(3)}$ may depend slowly on t, their derivatives have been left out, because they are of a higher order in ε. One should solve one order of ε at a time. The secular term in (17) will be zero if

$$i\omega_1\frac{\partial a}{\partial T_2} + \frac{1}{16}\omega_1^2|a|^2 a = 0. \tag{18}$$

Since this equation implies $\partial|a|^2/\partial T_2 = 0$, its solution is $a(t) = a_0\exp(\frac{1}{16}i\varepsilon^2|a_0|^2\omega_1 t)$, whence $\phi_3^{(3)} = -\frac{1}{192}a^3$ and $\phi_1^{(3)} = 0$. So

$$\phi = \varepsilon a_0 e^{-i\omega_1(1-(1/16)\varepsilon^2|a_0|^2)t} - \tfrac{1}{384}\varepsilon^3 a_0^3 e^{-3i\omega_1(1-(1/16)\varepsilon^2|a_0|^2)t} + \text{c.c.} + O(\varepsilon^5). \tag{19}$$

In figure 1 the linear approximation is given together with the 'real' solution as given by (19) for an amplitude of 0.5 rad. One can clearly see that though the frequency difference between the real solution and the linear approximation is only small, by just

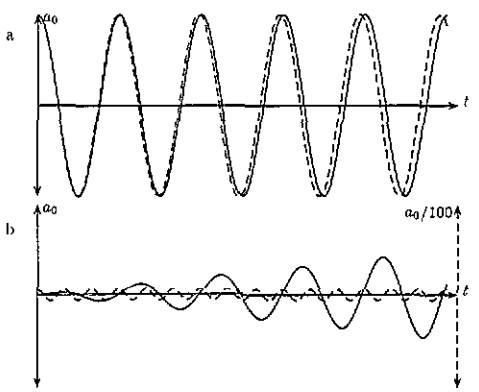

Figure 1. (a) The linear approximation $\varepsilon a_0\cos\omega_1 t$ (dashed) and the exact solution $\phi(t)$ (solid), for an amplitude of 0.5 rad. (b) The difference $\phi(t) - \varepsilon a_0\cos\omega_1(t)$ (solid), and the perturbation $\varepsilon^3\phi^{(3)}$ magnified 100 times (dashed).

adding a perturbation term one cannot compensate for it; a small difference in phase leads to a large difference between the function values $\phi(t) - \varepsilon a_0 \cos \omega_1 t$.

3.2. Asymptotic expansions of convolutions involving slowly varying functions

If a function A depends slowly on t, then a convolution product $\check{f} * (A\, e^{-i\omega_1 t})$ can be written as $e^{-i\omega_1 t} f(\omega_1 + i\varepsilon\partial_T) A(T)$. This can be shown by taking the Fourier transform of the convolution:

$$\begin{aligned}\mathscr{F}_t [\check{f} * (A\, e^{-i\omega_1 t})] &= f(\omega)\hat{A}(\omega - \omega_1) \\ &= (f(\omega_1) + f'(\omega_1)\Omega + \tfrac{1}{2} f''(\omega_1)\Omega^2 + \cdots)\hat{A}(\Omega) \qquad (20)\end{aligned}$$

with $\Omega = \omega - \omega_1$. The inverse Fourier transform of $\Omega\hat{A}(\Omega)$ is $i\partial_t A$, and $\partial_t A = \varepsilon\partial_T A$, so the inverse Fourier transform of the above expression is

$$\check{f} * (A\, e^{-i\omega_1 t}) = e^{-i\omega_1 t}(f(\omega_1) + i\varepsilon f'(\omega_1)\partial_T - \tfrac{1}{2}\varepsilon^2 f''(\omega_1)\partial_T^2 + \cdots)A. \qquad (21)$$

Symbolically, this can be written as

$$\check{f} * (A\, e^{-i\omega_1 t}) = e^{-i\omega_1 t} f(\omega_1 + i\varepsilon\partial_T) A. \qquad (22)$$

This is exact as long as the series (20) converges. However, notice that while the given function $f(\omega)$, often determined by the material, is expanded in a series, A is the quantity of interest. This means that $\hat{A}$ has to meet some requirements. In particular, where $f(\omega)$ is singular, $\hat{A}(\omega)$ must be zero. Hence $\hat{A}$ must have a limited bandwidth, and therefore A cannot have a fast time dependence.

Besides, at some stage in the calculations, the series will be cut off, so then from a certain n onwards, $\varepsilon^n \partial_T^n A$ must be negligible.

Similarly, if A depends slowly on z, the inverse Fourier transform of some function of k times $\mathscr{F}_z(A)$ can be written as $\mathscr{F}_z^{-1}[f(k)F_z[A](k - k_1)] = e^{ik_1 z} f(k_1 - i\varepsilon\partial_Z) A$.

3.3. Asymptotic expansion for a pulse in a waveguide

Suppose $\hat{E}$ is roughly centred at frequency ω_1. The wavenumber of the guided mode of the linear waveguide corresponding to this frequency is denoted by k_1, see section 4.

We separate the rapidly varying part of the electric field from the slowly varying part by writing $E = \frac{1}{2}\varepsilon E_1^{(1)} \exp i(k_1 z - \omega_1 t) + \text{c.c.} + O(\varepsilon^2)$. Abbreviate

$$\phi \equiv k_1 z - \omega_1 t. \qquad (23)$$

In higher order approximation, higher powers of the phase factor $e^{i\phi}$ will occur as well, so E is expanded in powers of ε and of $e^{i\phi}$:

$$E = \sum_{n>0} \varepsilon^n E^{(n)} = \text{Re} \sum_{\substack{n>0 \\ m\geqslant 0}} \varepsilon^n E_m^{(n)}\, e^{mi\phi} = \frac{1}{2} \sum_{\substack{n>0 \\ m\geqslant 0}} \varepsilon^n (E_m^{(n)}\, e^{mi\phi} + E_m^{(n)*}\, e^{-mi\phi}). \qquad (24)$$

The upper index (n) denotes the order of ε, the lower index m denotes the frequency, f_m is the component of f at frequency $m\omega_1$.

The $E_m^{(n)}$'s are supposed to depend slowly on t and z, they are functions of the slow variables $T = \varepsilon t$ and $Z = \varepsilon z$.

Notice that the epsilons used for scaling z, t and E are identical. One could have started with three different epsilons, ε_z, ε_t and ε_E. But then to relate the time behaviour to the z-behaviour, the epsilons must be related to one another. It turns out that choosing all epsilons the same is the simplest solution. This is not the only solution; for instance, choosing $\varepsilon_t = \varepsilon_E = \varepsilon_z^2$ works as well. In the end this would lead to the same results, i.e. in the end the same perturbation terms would show up, but they would appear at different orders of the perturbation parameter.

As in section 3.1, it is necessary to asymptotically expand the derivatives of $\boldsymbol{E}^{(n)}$. However, only expanding one of ∂_Z and ∂_T further will do, for this will give a new term at every order of ε, and at each order of ε there will be only one compatibility condition to be satisfied. Expanding both $\partial_Z \boldsymbol{E}^{(n)}$ and $\partial_T \boldsymbol{E}^{(n)}$ would lead to an underdetermined set of equations. It is more convenient to expand only ∂_Z any further, since the time dependence of the polarization is more involved than its spatial dependence. Namely, we do assume locality, while the polarization may depend on the past of the electric field.

So define $Z_n = \varepsilon^n z$ and for a function $F = F(T, Z)$ let

$$\partial_t F = \varepsilon F_T \qquad \text{and} \qquad \partial_z F = \varepsilon F_Z = \varepsilon F_{Z_1} + \varepsilon^2 F_{Z_2} + \cdots. \tag{25}$$

4. The linear waveguide

In a linear slab or fibre, the electric field satisfies

$$\check{\mathscr{L}} \boldsymbol{E} \equiv -\nabla \times \nabla \times \boldsymbol{E} - \frac{1}{c^2} \partial_t^2 (\boldsymbol{E} + \tilde{\chi}^{(1)} * \boldsymbol{E}) = \nabla^2 \boldsymbol{E} - \nabla(\nabla \cdot \boldsymbol{E}) - \frac{1}{c^2} \partial_t^2 (\boldsymbol{E} + \tilde{\chi}^{(1)} * \boldsymbol{E}) = 0. \tag{26}$$

The response function $\tilde{\chi}^{(1)}$ is assumed to be a function of the transverse coordinates $\boldsymbol{x}_\perp$ and of time, not of z, i.e. the refractive index profile is assumed not to change along the propagation direction.

The Fourier transform of (26) with respect to t is

$$\nabla^2 \hat{\boldsymbol{E}} - \nabla(\nabla \cdot \hat{\boldsymbol{E}}) + \frac{\omega^2}{c^2} (\hat{\boldsymbol{E}} + \tilde{\chi}^{(1)}(\omega, \boldsymbol{x}_\perp) \hat{\boldsymbol{E}}) = 0. \tag{27}$$

The index of refraction $n(\omega, \boldsymbol{x}_\perp)$ is defined by

$$n^2(\omega, \boldsymbol{x}_\perp) = 1 + \tilde{\chi}^{(1)}(\omega, \boldsymbol{x}_\perp). \tag{28}$$

Fourier transforming with respect to z as well, we find

$$\mathscr{L}(\omega, k) \mathscr{F}_z \hat{\boldsymbol{E}} = \begin{pmatrix} \partial_y^2 - k^2 + n^2 \dfrac{\omega^2}{c^2} & -\partial_x \partial_y & -ik\partial_x \\ -\partial_x \partial_y & \partial_x^2 - k^2 + n^2 \dfrac{\omega^2}{c^2} & -ik\partial_y \\ -ik\partial_x & -ik\partial_y & \partial_x^2 + \partial_y^2 + n^2 \dfrac{\omega^2}{c^2} \end{pmatrix} \mathscr{F}_z \hat{\boldsymbol{E}} \tag{29}$$

(obtained by simply replacing ∂_z by ik).

The fibre or slab is assumed to be single-mode and polarization-preserving, so at each frequency there is only one mode. At frequency ω this mode is denoted by $\boldsymbol{U}(\omega, \boldsymbol{x}_\perp)$, where $\boldsymbol{U}(\omega, \boldsymbol{x}_\perp)$ and the corresponding wavenumber $k(\omega)$ are both determined by

$$\mathscr{L}(\omega, k(\omega))\boldsymbol{U}(\omega, \boldsymbol{x}_\perp) = 0. \tag{30}$$

In other words, $\boldsymbol{U}(\omega, \boldsymbol{x}_\perp)\, e^{ik(\omega)z - i\omega t}$ is a solution of (26).

Assume for now, that $n(\omega)$ is real, i.e. neglect the damping. This is allowed far from resonance. Later on the damping will be taken into account as a perturbation. If the damping is too large to be treated perturbatively, the pulse will not propagate anyway. Then it will simply die.

With this assumption, $\mathscr{L}(\omega, k(\omega))$ is a Hermitian operator for real ω and k. Namely, define the inner product

$$\langle \boldsymbol{F}, \boldsymbol{G} \rangle \equiv \int_{X_\perp} (\boldsymbol{F}^* \cdot \boldsymbol{G})\, \mathrm{d}\boldsymbol{x}_\perp \tag{31}$$

on the space functions for which $\langle \boldsymbol{F}, \boldsymbol{F} \rangle < \infty$. Then the operator $\mathscr{L}(\omega, k)$ satisfies $\langle \boldsymbol{F}, \mathscr{L}(\omega, k)\boldsymbol{G} \rangle = \langle \mathscr{L}(\omega, k)\boldsymbol{F}, \boldsymbol{G} \rangle$. In particular, every $\boldsymbol{F}$ satisfies

$$\langle \boldsymbol{U}(\omega), \mathscr{L}(\omega, k(\omega))\boldsymbol{F} \rangle = \langle \mathscr{L}(\omega, k(\omega))\boldsymbol{U}(\omega), \boldsymbol{F} \rangle = \langle \boldsymbol{0}, \boldsymbol{F} \rangle = 0. \tag{32}$$

This property will be used extensively in the following sections.

Also, $k(\omega)$ is real, and $k(-\omega) = -k(\omega)$, and $\boldsymbol{U}(-\omega, \boldsymbol{x}_\perp) = \boldsymbol{U}^*(\omega, \boldsymbol{x}_\perp)$. This can be shown as follows. When $\boldsymbol{U}$ and k satisfy $\mathscr{L}(\omega, k(\omega))\boldsymbol{U}(\omega) = 0$,

$$\check{\mathscr{L}}\, \boldsymbol{U}(\omega, \boldsymbol{x}_\perp)\, e^{i(k(\omega)z - \omega t)} = 0.$$

Now since the operator $\check{\mathscr{L}}$ is real (in the space–time domain), this means that both the real part and the imaginary part of $\boldsymbol{E}(\boldsymbol{x}_\perp, z, t) = \boldsymbol{U}(\omega, \boldsymbol{x}_\perp) \exp i(k(\omega)z - \omega t)$ must satisfy $\mathscr{L}\ \mathrm{Re}\ \boldsymbol{E}(\boldsymbol{x}_\perp, z, t) = 0$ and $\mathscr{L}\ \mathrm{Im}\ \boldsymbol{E}(\boldsymbol{x}_\perp, z, t) = 0$ separately. So also the complex conjugate

$$\boldsymbol{E}^*(\boldsymbol{x}_\perp, z, t) = \boldsymbol{U}^*(\omega) \exp i(-k(\omega)z + \omega t)$$

satisfies $\check{\mathscr{L}}\boldsymbol{E}^*(\boldsymbol{x}_\perp, z, t) = 0$. Fourier transforming this in time and z-domain then shows that

$$\mathscr{L}(-\omega, -k(\omega))\boldsymbol{U}^*(\omega, \boldsymbol{x}_\perp) = 0.$$

This concludes the proof.

A related solution of (30) is $\{-\omega, k(\omega), \boldsymbol{U}(\omega)\}$. This describes a wave propagating in the opposite direction.

The wavenumber k is determined by the material dispersion and by the geometry of the waveguide. The inverse of $k' = \mathrm{d}k/\mathrm{d}\omega$ is the group velocity, and k'' is called the group velocity dispersion parameter. In fibre optics literature, the dispersion parameter D is used instead of k''. It is defined by $D = \mathrm{d}k'/\mathrm{d}\lambda \approx -(2\pi c/\lambda^2)k''$.

For bulk silica, D is zero at $\lambda_D = 1.312\ \mu\mathrm{m}$. In dispersion shifted fibres, this value is shifted, commonly to $\lambda \approx 1.55\ \mu\mathrm{m}$, in order to be able to use the erbium-doped fibre amplifiers that operate at that wavelength [2, 24]. For wavelengths greater than λ_D, D is positive and k'' is negative. This is referred to as anomalous dispersion.

4.1. The first guess

Consider a pulse with a 'carrier frequency' ω_1 (*the exact value of* ω_1 *does not matter*). In first order E should satisfy $\mathscr{L}E=0$, so $E_1^{(1)}$ has to be something like $U(\omega_1, x_\perp)A(Z, T)$. But the only demand for $E_1^{(1)}$ is that $\mathscr{L}(\varepsilon E_1^{(1)} e^{i\phi} + \text{c.c.}) = O(\varepsilon^2)$, so to $E_1^{(1)}$ we can add anything we like, as long as it is of order ε or smaller. We can use this freedom to choose $E_1^{(1)}$ such that some perturbation terms that will occur are accounted for automatically.

It turns out that a convenient choice is [27]

$$E_1^{(1)} = U(\omega_1 + i\varepsilon\partial_T)A(Z, T) \tag{33}$$

$$E_{-1}^{(1)} = E_1^{(1)*} = U^*(\omega_1 + i\varepsilon\partial_T)A^*(Z, T). \tag{34}$$

The advantage of this choice can be seen by Fourier transforming (33) as described in section 3.2. One then finds that $\varepsilon E^{(1)} = \frac{1}{2}\varepsilon E_1^{(1)} e^{i\phi} + \text{c.c.}$ satisfies the linear equation exactly, so additional terms in E are necessary only because of the non-linear right-hand side of (16).

N.B. Since per definition

$$U^*(z) \equiv [U(z)]^* \qquad U^*(\omega_1 + i\varepsilon\partial_T) = U^*(\omega_1) - i\varepsilon U_1'^*\partial_T - \tfrac{1}{2}\varepsilon^2 U_1''^*\partial_T^2 + \cdots.$$

We wish to solve (16) up to order ε^4 inclusive. This means we want to solve

$$\begin{aligned}
&\check{\mathscr{L}}U(\omega_1 + i\varepsilon\partial_T)A\, e^{i\phi} + \varepsilon^2\check{\mathscr{L}}E_1^{(2)} e^{i\phi} + \varepsilon^3\check{\mathscr{L}}E_1^{(3)} e^{i\phi} + \varepsilon^3\check{\mathscr{L}}E_3^{(3)} e^{3i\phi} + \varepsilon^4\check{\mathscr{L}}E_1^{(4)} e^{i\phi} \\
&\quad + \varepsilon^4\check{\mathscr{L}}E_3^{(4)} e^{3i\phi} + \text{c.c.} \\
&= \frac{1}{c^2}\partial_t^2((\varepsilon E_1^{(1)} e^{i\phi} + \varepsilon E_1^{(2)} e^{i\phi} + \text{c.c.}) \\
&\quad \times \tilde{\chi}^{(3)} * (\varepsilon E_1^{(1)} e^{i\phi} + \cdots)\cdot(\varepsilon E_1^{(1)} e^{i\phi} + \cdots)) + O(\varepsilon^5).
\end{aligned} \tag{35}$$

The expansion of $\check{\mathscr{L}}\varepsilon E_1^{(1)} = \text{Re}\, \varepsilon e^{i\phi}(\mathscr{L}U)(\omega_1 + i\varepsilon\partial_T, k_1 - i\varepsilon\partial_Z)A$ will be used up to order ε^4 inclusive. The k-argument in $\mathscr{L}\varepsilon E_1^{(1)}$ is $k_1 - i\varepsilon\partial_Z$, independent of $\omega_1 + i\varepsilon\partial_T$, and the expansion of $\mathscr{L}E_1^{(1)}$ will consist of all terms of the form $\partial_\omega^j\partial_k^l(\mathscr{L}U)(i\varepsilon\partial_T)^j(-i\varepsilon\partial_Z)^l A$.

Now consider

$$\mathscr{N}(\omega) \equiv \mathscr{L}(\omega, k(\omega))U(\omega, x_\perp) \tag{36}$$

which is a function of ω only. It is identically equal to zero, so in particular $\mathscr{N}(\omega_1 + i\varepsilon\partial_T)A = 0$. Therefore one can freely subtract $\frac{1}{2}e^{i\phi}\mathscr{N}(\omega_1 + i\varepsilon\partial_T)A + \text{c.c.}$ from $\mathscr{L}E$. This will eliminate all terms with factors $\partial_\omega^j(\mathscr{L}U)$, what is left of the expansions of $\varepsilon\mathscr{L}E_1^{(1)}$ is:

$$\begin{aligned}
&\varepsilon(\mathscr{L}U)(\omega_1 + i\varepsilon\partial_T, k_1 - i\varepsilon\partial_Z)A = \varepsilon(\mathscr{L}U)(\omega_1 + i\varepsilon\partial_T, k_1 - i\varepsilon\partial_Z)A - \varepsilon\mathscr{N}(\omega_1 + i\varepsilon\partial_T)A \\
&\quad = \varepsilon(\mathscr{L}U)A + i\varepsilon^2(\mathscr{L}U)_\omega\partial_T A - i\varepsilon^2(\mathscr{L}_k U)\partial_Z A - \tfrac{1}{2}\varepsilon^3(\mathscr{L}U)_{\omega\omega}\partial_T^2 A \\
&\qquad + \varepsilon^3(\mathscr{L}U)_{\omega k}\partial_T\partial_Z A - \tfrac{1}{2}\varepsilon^3(\mathscr{L}U)_{kk}\partial_Z^2 A - \tfrac{1}{6}i\varepsilon^4(\mathscr{L}U)_{\omega\omega\omega}\partial_T^3 A \\
&\qquad + \tfrac{1}{2}\varepsilon^4(\mathscr{L}U)_{\omega\omega k}\partial_T^2\partial_Z A - \tfrac{1}{2}\varepsilon^4(\mathscr{L}U)_{\omega kk}\partial_T\partial_Z^2 A + \tfrac{1}{6}i\varepsilon^4(\mathscr{L}U)_{kkk}\partial_Z^3 A \\
&\qquad - \varepsilon(\mathscr{L}U)A - i\varepsilon^2[(\mathscr{L}U)_\omega + k_1'(\mathscr{L}U)_k]\partial_T A + \tfrac{1}{2}\varepsilon^3[(\mathscr{L}U)_{\omega\omega} \\
&\qquad + 2k_1'(\mathscr{L}U)_{\omega k} + k_1'^2(\mathscr{L}U)_{kk} + k_1''(\mathscr{L}U)_k]\partial_T^2 A + \tfrac{1}{6}i\varepsilon^4[(\mathscr{L}U)_{\omega\omega\omega} \\
&\qquad + 3k_1'(\mathscr{L}U)_{\omega\omega k} + 3k_1'^2(\mathscr{L}U)_{\omega kk} + k_1'^3(\mathscr{L}U)_{kkk} + 3k_1''(\mathscr{L}U)_{\omega k} \\
&\qquad + 3k_1'k_1''(\mathscr{L}U)_{kk} + k_1'''(\mathscr{L}U)_k]\partial_T^3 A + O(\varepsilon^5) \\
&\quad = -i\varepsilon^2(\mathscr{L}_k U)(\partial_Z + k_1'\partial_T)A \\
&\qquad + \tfrac{1}{2}\varepsilon^3(\mathscr{L}_k U)k_1''\partial_T^2 A + \tfrac{1}{6}i\varepsilon^4 k_1'''(\mathscr{L}_k U)\partial_T^3 A
\end{aligned}$$

$$+ \tfrac{1}{2}\mathrm{i}\varepsilon^4 k_1'' \frac{\mathrm{d}}{\mathrm{d}\omega}(\mathscr{L}_k U)\partial_T^3 A$$
$$+ \varepsilon^3[(\mathscr{L}U)_{\omega k}\partial_T - \tfrac{1}{2}(\mathscr{L}U)_{kk}(\partial_Z - k_1'\partial_T)](\partial_Z + k_1'\partial_T)A$$
$$+ \mathrm{i}\varepsilon^4[\tfrac{1}{2}(\mathscr{L}U)_{\omega\omega k}\partial_T^2 - \tfrac{1}{2}(\mathscr{L}U)_{\omega kk}\partial_T(\partial_Z - k_1'\partial_T)$$
$$+ \tfrac{1}{6}(\mathscr{L}U)_{kkk}(\partial_Z^2 - k_1'\partial_Z\partial_T + \partial_T^2)](\partial_Z + k_1'\partial_T)A + \mathrm{O}(\varepsilon 5). \tag{37}$$

In the expanded form, $\mathscr{L}$, $\mathscr{L}_\omega$, U, U_ω, etc are evaluated at $\omega = \omega_1$ and $k = k_1$. To save on brackets later on, $(\mathscr{L}U)_k$ has been replaced by the equivalent expression $\mathscr{L}_k U$ (U does not depend explicitly on k).

5. Derivation of the amplitude equation

5.1. Order ε and ε^2

The non-linear right-hand side of equation (16) is of third order in ε, so in first and second order, E has to satisfy the linear equation $\check{\mathscr{L}}E = 0$.

In the right-hand side of the expansion of $\varepsilon\mathscr{L}E_1^{(1)}$, (37), the first term is of order ε^2. In this term, $\partial_Z A$ must be expanded to $\partial_Z A = \partial_{Z_1}A + \varepsilon\partial_{Z_2}A + \cdots$. Then at order ε^2, for $E_1^{(2)}$ one finds

$$-\mathrm{i}\mathscr{L}_k U(k_1'\partial_T + \partial_{Z_1})A + \mathscr{L}E_1^{(2)} = \mathrm{O}(\varepsilon). \tag{38}$$

Take the inner product with $U = U(\omega_1, x_\perp)$. This yields

$$-\mathrm{i}\langle U, \mathscr{L}_k U\rangle(k_1'\partial_T + \partial_{Z_1})A + \langle U, \mathscr{L}E_1^{(2)}\rangle = \mathrm{O}(\varepsilon). \tag{39}$$

The last term vanishes, since $\langle U, \mathscr{L}E_1^{(2)}\rangle = \langle \mathscr{L}U, E_1^{(2)}\rangle = \langle \mathbf{0}, E_1^{(2)}\rangle = 0$. One could interpret this as 'the source terms must be perpendicular to the null space of $\mathscr{L}$'.

Therefore $\partial_{Z_1}A$ is given by

$$k_1'\partial_T A + \partial_{Z_1}A = 0, \tag{40}$$

i.e. the amplitude propagates at the group velocity $v_G = 1/k_1'$. Since $E_1^{(2)}$ then satisfies the same equation as $E_1^{(1)}$, it can be taken to be zero (by replacing $E_1^{(1)}$ by $E_1^{(1)} + \varepsilon E_1^{(2)}$).

5.2. Do ∂_T and ∂_{Z_n} commute?

When working with the operators $\partial_T, \partial_Z, \partial_{Z_1}, \partial_{Z_2}, \ldots$, caution is needed. They function like derivative operators, but the Z_ns are not independent, so the derivatives with respect to them do not have to commute.

Since there is only one slow time scale, derivatives with respect to time are easy: for any f, $\partial_t^n f = \varepsilon^n \partial_T^n f$ and $\partial_T^n f = \varepsilon^{-n}\partial_t^n f$. Higher order derivatives with respect to z are more involved.

Per definition, $\partial_z f = \varepsilon\partial_{Z_1}f + \varepsilon^2\partial_{Z_2}f + \cdots$. The way in which $\partial_z f$ is split up among the time scales is not given beforehand. At each order of ε, $\partial_{Z_n}A$ is *chosen* in a certain way, where any choice that satisfies the compatibility condition at that order is allowed. Fixing these choices, the $\partial_{Z_n}A$s become functions of A, A^* and their derivatives with respect to time: $\partial_{Z_1}A \equiv A_1(A, A^*, A_T, A_T^*, \ldots)$, $\partial_{Z_2}A \equiv A_2(A, A^*, A_T, A_T^*, \ldots)$ etc. Then the $\partial_{Z_n}A$s are given by the complex conjugate of $\partial_{Z_n}A$, so $\partial_{Z_1}A^* = A_1^*(A, \ldots)$, etc.

For higher derivatives, what $\partial_T \partial_{Z_n}$ should mean, is clear: it is

$$\partial_T \partial_{Z_n} A = \frac{\partial A_n}{\partial A} A_T + \frac{\partial A_n}{\partial A^*} A_T^* + \frac{\partial A_n}{\partial A_T} A_{TT} + \cdots.$$

What $\partial_{Z_n} \partial_T A$ should mean is less clear. So far, we only know how to apply ∂_{Z_n} to (functions of) A and A^*. But we can just give $\partial_{Z_n} \partial_T A$ the most intuitive definition, i.e. we can define it to be equal to $\partial_T \partial_{Z_n} A$. Then per definition, ∂_{Z_n} and ∂_T commute.

Now we can also calculate higher order derivatives with respect to z, namely

$$\begin{aligned}\partial_{Z_n} \partial_{Z_m} A &= \frac{\partial A}{\partial Z_n} \frac{\partial A_m}{\partial A} + \frac{\partial A^*}{\partial Z_n} \frac{\partial A_m}{\partial A^*} + \frac{\partial A_T}{\partial Z_n} \frac{\partial A_m}{\partial A_T} + \cdots \\ &= A_n \frac{\partial A_m}{\partial A} + A_n^* \frac{\partial A_m}{\partial A^*} + \frac{\partial A_n}{\partial T} \frac{\partial A_m}{\partial A_T} + \cdots.\end{aligned}$$

It is obvious from this expression that $\partial_{Z_n} \partial_{Z_m} A$ does not have to be equal to $\partial_{Z_m} \partial_{Z_n} A$. But as we just saw in (40), A satisfies $A_{Z_1} = -k_1' A_T$; A^* satisfies the same equation. And also for e.g. A_T,

$$\partial_{Z_1}(A_T) = \partial_T A A_{Z_1} = \partial_T(-k_1' A_T) = -k_1' \partial_T(A_T).$$

So $\partial_{Z_1} f = -k_1' \partial_T f$ for any function f of A, A^* and their derivatives with respect to time. Now since ∂_T and ∂_{Z_n} commute for all n, so will ∂_{Z_1} and ∂_{Z_n}.

As a consequence, any function of A and A^* and their derivatives with respect to T and Z_n will satisfy $(\partial_{Z_1} + k_1' \partial_T) f(A, A^*, \ldots) = 0$.

However, probably ∂_{Z_2} and ∂_{Z_3} will not commute anymore.

5.3. *Order* ε^3

After substituting $A_{Z_1} + k_1' A_T = 0$ in (37), the left-hand side of (16) becomes

$$\begin{aligned}\check{\mathscr{L}} \boldsymbol{E} &= \tfrac{1}{2}\varepsilon\, \mathrm{e}^{\mathrm{i}\phi} (\mathscr{L}\boldsymbol{U})(\omega_1 + \mathrm{i}\partial_T, k_1 - \mathrm{i}\partial_Z) A(Z, T) + \text{c.c.} \\ &\quad + \varepsilon^3 \check{\mathscr{L}} \boldsymbol{E}^{(3)} + \varepsilon^4 \check{\mathscr{L}} \boldsymbol{E}^{(4)} + \mathrm{O}(\varepsilon^5) \\ &= \tfrac{1}{2}\varepsilon\, \mathrm{e}^{\mathrm{i}\phi} [-\mathrm{i}\varepsilon^2 \mathscr{L}_k \boldsymbol{U}(\partial_{Z_2} + \varepsilon \partial_{Z_3}) A + \tfrac{1}{2}\varepsilon^2 k_1'' \mathscr{L}_k \boldsymbol{U} \partial_T^2 A + \tfrac{1}{6}\mathrm{i} k_1''' \varepsilon^3 \mathscr{L}_k \boldsymbol{U} \partial_T^3 A \\ &\quad + \tfrac{1}{2}\mathrm{i}\varepsilon^2 k_1'' \frac{\mathrm{d}}{\mathrm{d}\omega} (\mathscr{L}_k \boldsymbol{U}) \partial_T^3 A + \varepsilon^2 [(\mathscr{L}\boldsymbol{U})_{\omega k} \partial_T - \tfrac{1}{2}(\mathscr{L}\boldsymbol{U})_{kk} (\partial_{Z_1} - k_1' \partial_T)] \varepsilon \partial_{Z_1} A] \\ &\quad + \text{c.c.} + \varepsilon^3 \check{\mathscr{L}} \boldsymbol{E}^{(3)} + \varepsilon^4 \check{\mathscr{L}} \boldsymbol{E}^{(4)} + \mathrm{O}(\varepsilon^5) \\ &= \tfrac{1}{2}\varepsilon^3\, \mathrm{e}^{\mathrm{i}\phi} \mathscr{L}_k \boldsymbol{U}(-\mathrm{i} A_{Z_2} + \tfrac{1}{2} k_1'' A_{TT}) + \text{c.c.} + \varepsilon^3 \check{\mathscr{L}} \boldsymbol{E}^{(3)} \\ &\quad + \tfrac{1}{2}\varepsilon^4\, \mathrm{e}^{\mathrm{i}\phi} [\mathscr{L}_k \boldsymbol{U}(-\mathrm{i} A_{Z_3} + \tfrac{1}{6}\mathrm{i} k_1''' A_{TTT}) + \frac{\mathrm{d}}{\mathrm{d}\omega} (\mathscr{L}_k \boldsymbol{U}) \partial_T (\partial_{Z_2} + \tfrac{1}{2}\mathrm{i} k_1'' \partial_T^2) A] \\ &\quad + \text{c.c.} + \varepsilon^4 \check{\mathscr{L}} \boldsymbol{E}^{(4)} + \mathrm{O}(\varepsilon^5). \qquad (41)\end{aligned}$$

The following has been used:

$$[(\mathscr{L}\boldsymbol{U})_{\omega k} \partial_T - \tfrac{1}{2}(\mathscr{L}\boldsymbol{U})_{kk} (\partial_{Z_1} - k_1' \partial_T)] \varepsilon \partial_{Z_2} A$$

$$= [((\mathscr{L}U)_{\omega k} + k_1'(\mathscr{L}U)_{kk})\partial_T - \tfrac{1}{2}(\mathscr{L}U)_{kk}(\partial_{Z_1} + k_1'\partial_T)]\varepsilon\partial_{Z_2}A$$

$$= \varepsilon \frac{\mathrm{d}}{\mathrm{d}\omega}(\mathscr{L}_k U)\partial_T\partial_{Z_2}A$$

since $(\partial_{Z_1} + k_1'\partial_T)\partial_{Z_2}A = \partial_{Z_2}(\partial_{Z_1} + k_1'\partial_T)A = 0$.

The right-hand side of (16) now comes into play. Expanded up to second order, E is given by

$$E = \tfrac{1}{2}(\varepsilon UA\,\mathrm{e}^{\mathrm{i}\phi} + \mathrm{i}\varepsilon^2 U'A_T\,\mathrm{e}^{\mathrm{i}\phi} + \varepsilon U^*A^*\,\mathrm{e}^{-\mathrm{i}\phi} - \mathrm{i}\varepsilon^2 U'^*A_T^*\,\mathrm{e}^{-\mathrm{i}\phi}) + \mathrm{O}(\varepsilon^3). \tag{43}$$

Define

$$V_1^{(3)}(\omega) \equiv -\frac{\omega^2}{4c^2}[2(U\cdot U^*)U\hat{\chi}^{(3)}(0) + (U\cdot U)U^*\hat{\chi}^{(3)}(2\omega)] \tag{44}$$

$$V_3^{(3)}(3\omega) \equiv -\frac{9\omega^2}{4c^2}(U\cdot U)U\hat{\chi}^{(3)}(2\omega) \tag{45}$$

$$\begin{aligned} V_1^{(4)}(\omega) \equiv \frac{\omega^2}{2c^2}[&\hat{\chi}^{(3)\prime}(2\omega)(U\cdot U)U^* + \hat{\chi}^{(3)}(2\omega)(U\cdot U')U^* + \hat{\chi}^{(3)}(2\omega)(U\cdot U)U'^* \\ &- \hat{\chi}^{(3)\prime}(0)(U\cdot U^*)U + \hat{\chi}^{(3)}(0)(U'\cdot U^*)U + 2\hat{\chi}^{(3)}(0)(U\cdot U'^*)U \\ &+ \hat{\chi}^{(3)}(0)(U\cdot U^*)U'] \end{aligned} \tag{46}$$

with $U = U(\omega)$ and $U^* = U^*(\omega) = U(-\omega)$. Then the right-hand side of (16) is given by

$$\begin{aligned} \frac{1}{c^2}\partial_t^2(E\tilde{\chi}^{(3)} * (E\cdot E)) = {} & \tfrac{1}{2}\varepsilon^3\,\mathrm{e}^{\mathrm{i}\phi}V_1^{(3)}|A|^2A + \tfrac{1}{2}\varepsilon^3\,\mathrm{e}^{3\mathrm{i}\phi}V_3^{(3)}A^3 + \tfrac{1}{2}\mathrm{i}\varepsilon^4\,\mathrm{e}^{\mathrm{i}\phi}V_1^{(3)\prime}\partial_T(|A|^2A) \\ &+ \tfrac{1}{2}\mathrm{i}\varepsilon^4\,\mathrm{e}^{3\mathrm{i}\phi}V_3^{(3)\prime}\partial_T(A^3) + \tfrac{1}{2}\mathrm{i}\varepsilon^4\,\mathrm{e}^{\mathrm{i}\phi}V_1^{(4)}A\partial_T(|A|^2) + \text{c.c.} + \mathrm{O}(\varepsilon^5). \end{aligned} \tag{47}$$

N.B.

$$V_3^{(3)\prime} = \frac{\mathrm{d}}{\mathrm{d}\omega}V_3^{(3)}(\omega)|_{\omega=3\omega_1} \qquad \text{not } \frac{\mathrm{d}}{\mathrm{d}\omega}V_3^{(3)}(3\omega)|_{\omega=\omega_1}.$$

$V_1^{(3)} = V_1^{(3)}(\omega_1)$, $V_3^{(3)} = V_3^{(3)}(3\omega^1)$, etc.

So at order ε^3,

$$\begin{aligned} &\tfrac{1}{2}\varepsilon^3\,\mathrm{e}^{\mathrm{i}\phi}\mathscr{L}_kU(-\mathrm{i}\partial_{Z_2} + \tfrac{1}{2}k_1''\partial_T^2)A + \tfrac{1}{2}\varepsilon^3\mathscr{L}E_1^{(3)}\,\mathrm{e}^{\mathrm{i}\phi} + \tfrac{1}{2}\varepsilon^3\mathscr{L}E_3^{(3)}\,\mathrm{e}^{3\mathrm{i}\phi} \\ &\quad = \tfrac{1}{2}\varepsilon^3\,\mathrm{e}^{\mathrm{i}\phi}V_1^{(3)}|A|^2A + \tfrac{1}{2}\varepsilon^3\,\mathrm{e}^{3\mathrm{i}\phi}V_3^{(3)}A^3 + \mathrm{O}(\varepsilon^4). \end{aligned} \tag{48}$$

For the compatibility condition look at the terms with phase factor $\mathrm{e}^{\mathrm{i}\phi}$ and take the inner product with U. Then again, like in equation (39), $\mathscr{L}E_1^{(3)}$ will drop out since $\langle U, \mathscr{L}E_1^{(3)}\rangle = \langle \mathscr{L}U, E_1^{(3)}\rangle = \langle \mathbf{0}, E_1^{(3)}\rangle = 0$. So

$$\tfrac{1}{2}\langle U, \mathscr{L}_kU\rangle(-\mathrm{i}\partial_{Z_2} + \tfrac{1}{2}k_1''\partial_T^2)A - \tfrac{1}{2}\langle U, V_1^{(3)}\rangle|A|^2A = \mathrm{O}(\varepsilon). \tag{49}$$

Define

$$\gamma(\omega) = \frac{\langle \boldsymbol{U}(\omega), \boldsymbol{V}_1^{(3)}(\omega)\rangle}{\langle \boldsymbol{U}(\omega), (\mathscr{L}_k \boldsymbol{U})(\omega, k(\omega))\rangle}. \tag{50}$$

Then one can recognize the non-linear Schrödinger equation in (49):

$$\mathrm{i}\partial_{Z_2} A - \tfrac{1}{2}k_1''\partial_T^2 A + \gamma_1 |A|^2 A = 0.$$

Equation (49) has to be satisfied up to $O(\varepsilon)$, so nothing prevents us from adding terms of order ε. And like letting $\boldsymbol{E}_1^{(1)} = \boldsymbol{U}(\omega_1 + \mathrm{i}\varepsilon\partial_T)A$ instead of $\boldsymbol{U}(\omega_1)A$ automatically took care of higher order linear terms, it turns out that higher order non-linear terms will be simplied if we replace $\gamma_1 = \gamma(\omega_1)$ by $\gamma(\omega_1 + \mathrm{i}\varepsilon\partial_T)$, and let $\partial_{Z_2} A$ be given by

$$\mathrm{i}\partial_{Z_2} A - \tfrac{1}{2}k_1''\partial_T^2 A - \gamma(\omega_1 + \mathrm{i}\varepsilon\partial_T)(|A|^2 A) = 0. \tag{51}$$

For $\boldsymbol{E}_1^{(3)}$ we make the ansatz

$$\boldsymbol{E}_1^{(3)} = \boldsymbol{U}_1^{(3)}(\omega_1 + \mathrm{i}\varepsilon\partial_T, k_1 - \mathrm{i}\varepsilon\partial_Z)(|A|^2 A). \tag{52}$$

Substituting (52) and (51) in (48), we find that $\boldsymbol{U}_1^{(3)}$ has to satisfy

$$\mathscr{L}(\omega, \mathrm{k})\boldsymbol{U}_1^{(3)}(\omega, k, \boldsymbol{x}_\perp) = \boldsymbol{V}_1^{(3)}(\omega, \boldsymbol{x}_\perp) - \gamma(\omega)\mathscr{L}_k(\omega, k(\omega))\boldsymbol{U}(\omega, \boldsymbol{x}_\perp). \tag{53}$$

The solution of this equation is not unique for $k = k(\omega)$. Namely, since

$$\mathscr{L}(\omega, k(\omega))\boldsymbol{U}(\omega) = 0,$$

any multiple of $\boldsymbol{U}$ can be added to $\boldsymbol{U}_1^{(3)}(\omega, k(\omega))$. However, $\lim_{k \to k(\omega)} \boldsymbol{U}_1^{(3)}(\omega, k)$ is unique, and k-derivatives of $\boldsymbol{U}_1^{(3)}$ that occur in the expansion of (52) will only make sense if this value is chosen for $\boldsymbol{U}_1^{(3)}(\omega, k(\omega))$.

Note, however, that as far as an equation for A is concerned, until order ε^5 the quantity $\boldsymbol{U}_1^{(3)}$ itself is not needed. For the linear left-hand side of (16), we only need to know $\mathscr{L}\boldsymbol{E}_1^{(3)}$; and that quantity is fully determined by (52) and (53). And in the non-linear right-hand side of (16), the first term in which $\boldsymbol{U}_1^{(3)}$ occurs is of order ε^5.

The equation for $\boldsymbol{E}_3^{(3)}$, following from the terms with phase factor $\mathrm{e}^{3\mathrm{i}\phi}$ in (48), is simpler,

$$\boldsymbol{E}_3^{(3)} = \boldsymbol{U}_3^{(3)}(3\omega_1 + \mathrm{i}\varepsilon\partial_T, 3k_1 - \mathrm{i}\varepsilon\partial_Z)A^3 \tag{54}$$

where

$$\mathscr{L}(3\omega, 3k)\boldsymbol{U}_3^{(3)}(3\omega, 3k, \boldsymbol{x}_\perp) = \boldsymbol{V}_3^{(3)}(3\omega, \boldsymbol{x}_\perp). \tag{55}$$

This equation has a unique solution if $k(3\omega_1) \neq 3k(\omega_1)$. If $k(3\omega_1) = 3k(\omega_1)$, the effective indices of refraction are equal. Then this equation is secular too, since the homogeneous equation then has a non-trivial solution, and a second compatibility condition has to be imposed. In physical terms, this is the case of phase matching, in which net long range effects are possible.

We will not pursue this complication any further.

5.4. *Order* ε^4

Thanks to the choice of $\boldsymbol{E}_3^{(3)}$, after elimination of ∂_{Z_1} and ∂_{Z_2} by means of (40) and (51), all fourth-order terms with phase factor $\mathrm{e}^{3\mathrm{i}\phi}$ cancel, so $\boldsymbol{E}_3^{(4)} = 0$. This shows the convenience of the choice of (54).

After expanding (52) to give

$$\mathscr{L}E_1^{(3)} = (\mathscr{L}U_1^{(3)})(\omega_1 + i\varepsilon\partial_T)(|A|^2A)$$
$$= (V_1^{(3)} - \gamma_1\mathscr{L}_kU)|A|^2A + i\varepsilon(V_1^{(3)\prime} - \gamma_1'\mathscr{L}_kU - \gamma(\mathscr{L}_kU)_1')\partial_T(|A|^2A) + O(\varepsilon^2) \quad (56)$$

and (51) to give

$$(-i\partial_{Z_2} + \tfrac{1}{2}k_1''\partial_T^2)A = \gamma_1|A|^2A + i\varepsilon\gamma_1'\partial_T(|A|^2A) + O(\varepsilon^2) \quad (57)$$

the left-hand side of (16), that was expanded already in (41), becomes

$$\check{\mathscr{L}}E = \tfrac{1}{2}\varepsilon^3\,e^{i\phi}V_1^{(3)}|A|^2A + \tfrac{1}{2}i\varepsilon^4\,e^{i\phi}V_1^{(3)\prime}\partial_T(|A|^2A) + \tfrac{1}{2}i\varepsilon^3\,e^{3i\phi}V_3^{(3)}A^3 + \tfrac{1}{2}i\varepsilon^4\,e^{3i\phi}V_3^{(3)\prime}\partial_T(A^3)$$
$$+ \tfrac{1}{2}\varepsilon^4\,e^{i\phi}\mathscr{L}_kU(-i\partial_{Z_3} + \tfrac{1}{6}ik_1'''\partial_T^3)A + \tfrac{1}{2}\varepsilon^4\,e^{i\phi}\mathscr{L}E_1^{(4)} + \text{c.c.} + O(\varepsilon^5) \quad (58)$$

and this should be equal to (47), so $E_1^{(4)}$ should satisfy

$$\mathscr{L}_kU(-i\partial_{Z_3} + \tfrac{1}{6}ik_1'''\partial_T^3)A + \mathscr{L}E_1^{(4)} = iV_1^{(4)}A\partial_T|A|^2 + O(\varepsilon). \quad (59)$$

The inner product of (59) with U yields a compatibility condition again, like in equations (39) and (49), namely

$$\langle U, \mathscr{L}_kU\rangle(-i\partial_{Z_3} + \tfrac{1}{6}ik_1'''\partial_T^3)A - i\langle U, V_1^{(4)}\rangle A\partial_T(|A|^2) = 0. \quad (60)$$

Define the parameter η by

$$\eta(\omega) = \frac{\langle U(\omega), V_1^{(4)}(\omega)\rangle}{\langle U(\omega), (\mathscr{L}_kU)(\omega, k(\omega))\rangle} \quad (61)$$

then $\partial_{Z_3}A$ is given by

$$i\partial_{Z_3}A - \tfrac{1}{6}ik_1'''\partial_T^3A + i\eta(\omega_1 + i\varepsilon\partial_T)A\partial_T(|A|^2) = 0. \quad (62)$$

The perturbation term $E_1^{(4)}$ is given by

$$E_1^{(4)} = U_1^{(4)}(\omega_1 + i\varepsilon\partial_T, k_1 - i\varepsilon\partial_Z)(A\partial_T|A|^2) \quad (63)$$

and

$$\mathscr{L}(\omega, k)U_1^{(4)}(\omega, k) = iV_1^{(4)}(\omega) - i\eta(\omega)\mathscr{L}_kU(\omega, k(\omega)). \quad (64)$$

If one only considers terms up to order ε^4, one can replace $\eta(\omega_1 + i\varepsilon\partial_T)$ and $U_1^{(4)}(\omega_1 + i\varepsilon\partial_T, k_1 - i\varepsilon\partial_Z)$ by $\eta_1 = \eta(\omega_1)$ and $U_1^{(4)}(\omega_1, k_1)$.

5.5. *Collecting results*

So far, the 'small parameter' ε has been left unspecified. But it has been assumed that the amplitude varies slowly over the optical period of the light $2\pi/\omega_1$. If T_0 is the pulse duration, this means that ω_1T_0 must be much larger than unity. Therefore, let $\varepsilon = (\omega_1T_0)^{-1}$, and define a dimensionless time τ in a frame of reference moving at group velocity

$$\tau = \frac{t - k_1'z}{T_0} = \frac{T - k_1'Z}{\varepsilon T_0}. \quad (65)$$

A natural scale for the propagation distance is the dispersion length L_D, defined by $L_D = T_0^2/|k_1''|$. It is the distance over which group velocity dispersion leads to pulse spreading. For instance, for a Gaussian pulse, the pulse duration is proportional to

$\sqrt{1 + (z/L_D)^2}$. So define the dimensionless propagation distance ζ by

$$\zeta = \frac{|k_1''|}{T_0^2} z = \frac{z}{L_D} = \frac{|k_1''|}{\varepsilon T_0^2} Z. \tag{66}$$

In order to arrive at a dimensionless non-linear Schrödinger equation, it turns out that the amplitude εA must be scaled by $\sqrt{\gamma_1 L_D}$, so define a dimensionless amplitude a by

$$a = \sqrt{\frac{T_0^2 \gamma_1}{|k_1''|}}\, \varepsilon A = \sqrt{\gamma_1 L_D}\, \varepsilon A. \tag{67}$$

Now $\partial_\zeta a$ is given by $\partial_\zeta a = \mathrm{i}\varepsilon L_D(\partial_{Z_1} + k_1' \partial_T)a + \mathrm{i}\varepsilon^2 L_D \partial_{Z_2} a + \mathrm{i}\varepsilon^3 L_D \partial_{Z_3} a$. From (40), the first term is zero. From (57), if k_1'' is negative, so in the case of anomalous dispersion,

$$\mathrm{i}\varepsilon^2 L_D \partial_{Z_2} a = \mathrm{i}\sqrt{\gamma_1 L_D}\, L_D \varepsilon^3 \partial_{Z_2} A = -\tfrac{1}{2} a_{\tau\tau} - |a|^2 a - \frac{\gamma_1'}{T_0 \gamma_1} (|a|^2 a)_\tau$$

and from (62),

$$\mathrm{i}\varepsilon^3 L_D \partial_{Z_3} a = \mathrm{i}\, \frac{k'''}{6 T_0 |k''|}\, a_{\tau\tau\tau} - \mathrm{i}\, \frac{\eta_1}{T_0 \gamma_1} (|a|^2)_\tau a.$$

Define the dimensionless parameters δ, τ_{shock} and τ_R by

$$\delta = -\frac{k_1'''}{6 T_0 |k_1''|} \tag{68}$$

$$\tau_{\text{shock}} = \frac{\gamma_1'}{T_0 \gamma_1} \tag{69}$$

$$\tau_R = -\mathrm{i}\, \frac{\eta_1}{T_0 \gamma_1}. \tag{70}$$

The parameter δ indicates the relative size of the third order dispersion, τ_{shock} is related to the self steepening and τ_R is the Raman response time. The definition in (70) makes τ_R positive, possibly with a small imaginary part.

Then the envelope a must satisfy

$$\mathrm{i}a_\zeta + \tfrac{1}{2} a_{\tau\tau} + |a|^2 a + \mathrm{i}\delta a_{\tau\tau\tau} + \mathrm{i}\tau_{\text{shock}}(|a|^2 a)_\tau - \tau_R (|a|^2)_\tau a = 0. \tag{71}$$

5.6. *Damping*

So far, the imaginary part of the linear susceptibility, corresponding to damping, has been neglected. To keep $\mathscr{L}$ Hermitian, the damping must be taken into account as a perturbation, i.e. it must be brought to the right-hand side of equation (16). Thus, equation (48) will change to

$$\begin{aligned} &\tfrac{1}{2}\varepsilon^3 \mathrm{e}^{\mathrm{i}\phi} \mathscr{L}_k \boldsymbol{U}(-\mathrm{i}\partial_{Z_2} + \tfrac{1}{2} k_1'' \partial_T^2) A + \tfrac{1}{2}\varepsilon^3 \mathscr{L} \boldsymbol{E}_1^{(3)} \mathrm{e}^{\mathrm{i}\phi} + \tfrac{1}{2}\varepsilon^3 \mathscr{L} \boldsymbol{E}_3^{(3)} \mathrm{e}^{3\mathrm{i}\phi} \\ &\quad = \tfrac{1}{2}\varepsilon^3 \mathrm{e}^{\mathrm{i}\phi} \boldsymbol{V}_1^{(3)} |A|^2 A + \tfrac{1}{2}\varepsilon^3 \mathrm{e}^{3\mathrm{i}\phi} \boldsymbol{V}_3^{(3)} A^3 - \tfrac{1}{2}\varepsilon \frac{\omega_1^2}{c^2} \mathrm{i}\hat{\chi}_{\text{im}}^{(1)} \boldsymbol{U} A \mathrm{e}^{\mathrm{i}\phi} + \mathrm{O}(\varepsilon^4) \end{aligned} \tag{72}$$

where $\hat{\chi}^{(1)}_{\mathrm{im}}$ is the imaginary part of the linear susceptibility at the central frequency ω_1, $\hat{\chi}^{(1)}_{\mathrm{im}} = \mathrm{Im}\,\hat{\chi}(\omega_1, x_\perp)$.

The damping term in (72) seems to be of order ε, not of order ε^3. However, $\hat{\chi}^{(1)}_{\mathrm{im}}$ was supposed to be small, and we must suppose it to be of order ε^2 or smaller. If it is not, dissipation will be too high and the pulse will not propagate.

Define

$$\Gamma = -\frac{\omega_1^2 L_{\mathrm{D}}}{c^2}\frac{\langle U, \hat{\chi}^{(1)}_{\mathrm{im}} U\rangle}{\langle U, \mathscr{L}_k U\rangle}, \tag{73}$$

in which a factor ε^{-2} is hidden in the factor L_{D}. Then (72) means that equation (51) for $\partial_{z_2} A$ has to be replaced by

$$\mathrm{i}\partial_{z_2} A - \tfrac{1}{2}k_1''\partial_T^2 A + \gamma(\omega_1 - \mathrm{i}\varepsilon\partial_T)|A|^2 A + \mathrm{i}(\varepsilon^2 L_{\mathrm{D}})^{-1}\Gamma A = 0. \tag{74}$$

Therefore equation (71) changes to

$$\mathrm{i}a_\zeta + \tfrac{1}{2}a_{\tau\tau} + |a|^2 a + \mathrm{i}\delta a_{\tau\tau\tau} + \mathrm{i}\tau_{\mathrm{shock}}(|a|^2 a)_\tau - \tau_{\mathrm{R}}(|a|^2)_\tau a = 0. \tag{75}$$

6. Limitations

As mentioned in the introduction, some assumptions have been made. Firstly, the polarization has been supposed to be local. Secondly, the material has been supposed to be isotropic. Also, the fibre or waveguide should be single-mode and polarization-preserving, in order not to have to consider coupled equations. Coupled equations would also arise in the case of phase matching, when $k(3\omega_1) = 3k(\omega_1)$. Finally, the non-linear polarization has been treated in the Born–Oppenheimer approximation. All these assumptions seem to be justified, and are commonly made.

To estimate for what pulse duration T_0 the derived equation (75) is valid, consider glass at $\lambda = 1.55\ \mu\mathrm{m}$, with group velocity dispersion either $k_1'' = -20\ \mathrm{ps}^2\ \mathrm{km}^{-1}$ for a normal fibre or $k_1'' = -2\ \mathrm{ps}^2\ \mathrm{km}^{-1}$ for a dispersion shifted fibre. For the non-linearity take the non-linear refractive index n_2, related to $\chi^{(3)}$ by $n_2 = 3\chi^{(3)}/8n$, to be $n_2 = 2.3 \times 10^{-22}\ \mathrm{m}^2\ \mathrm{V}^{-2}$ (values taken from [2]). This value of n_2 implies $\chi^{(3)} = 8.9 \times 10^{-22}\ \mathrm{m}^2\ \mathrm{V}^{-2}$.

First of all, in the derivation, the terms $(\boldsymbol{U}\cdot\boldsymbol{U})\boldsymbol{U}A^3\,\mathrm{e}^{3\mathrm{i}\phi}$ and $(\boldsymbol{U}\cdot\boldsymbol{U}^*)\boldsymbol{U}|A|^2 A\,\mathrm{e}^{\mathrm{i}\phi}$ have been supposed not to overlap in the ω-domain. Therefore, for the validity of (75), the spectral width of $\hat{A}$ must certainly be less than $\frac{1}{3}\omega_1$.

$\hat{E}(\omega)$ must be negligible at the resonance frequencies of the material. Glass has linear resonance frequencies at wavelengths of approximately 0.1 μm and 10 μm. Given a central frequency corresponding to $\lambda = 1.5\ \mu\mathrm{m}$, this means that $\hat{A}(\omega)$ must be negligible at $\omega = 0.85\omega_1$; so $T_0 \gg 1$ fs.

As for the non-linear perturbations, the most stringent limit on the pulse length is given by the stimulated Raman scattering. The stimulated Raman scattering has a resonance frequency corresponding to a time $\tau_1 \approx 12$ fs, see section 2.2.2. Therefore $|\boldsymbol{E}|^2$ cannot be allowed to have a spectral width larger than approximately 10^{14} Hz. This means that T_0 must be much greater than 10 fs. The other perturbation terms involving time derivatives of the non-linear terms are all much smaller.

To estimate the magnitude of the higher order non-linear perturbation terms, one needs to know the magnitude of the electric field. Since the longitudinal component of the modal profile, i.e. the z-component of $\boldsymbol{U}$, is small, $\mathscr{L}_k \boldsymbol{U}$ can be approximated by $-2k\boldsymbol{U}$, see (29). If one inserts this approximation in the expansion of (67), one can see

that the electric field will be of the order of

$$|E| \approx \sqrt{\frac{8nc|k_1''|}{3\omega_1\chi^{(3)}T_0^2}}\,|a| \approx \frac{1\text{ fs}}{T_0} \times 5 \times 10^9\text{ V m}^{-1} \times |a| \tag{76}$$

for a normal fibre and $(1\text{ fs}/T_0) \times 1 \times 10^9\text{ V m}^{-1} \times |a|$ for a dispersion shifted fibre.

To estimate the size of the fifth order polarization $P^{(5)} = \varepsilon_0\chi^{(5)}E^5$, we also need to know $\chi^{(5)}$. It is hard to find a definite figure for this parameter, but we can make an estimate if there are no resonances. In that case, the nth order susceptibility $\chi^{(n)}$ is of the order of $E_a^{-(n-1)}$, where E_a is the interval field of order $3 \times 10^{10}\text{ V m}^{-1}$ that binds the electrons to the ions [8, 20]. So take $\chi^{(5)} \approx 1 \times 10^{-42}\text{ m}^4\text{ V}^{-4}$.

With this value of $\chi^{(5)}$, the contribution of the fifth order polarization to equation (75) is $(0.02\text{ fs}^2/T_0^2)|a|^4a$ for a normal fibre and $(0.002\text{ fs}^2/T_0^2)|a|^4a$ for a dispersion shifted fibre. So for T_0 much larger than a femtosecond, this will be negligible ($|a|$ is assumed to be of order 1).

Another perturbation term at order ε^5 is the term proportional to $\chi^{(3)}(E^{(1)})^2E^{(3)}$. After scaling, this contributes something of the order of $k_1|k_1''|A_{\text{core}}T_0^{-2}|a|^4a$. For a fibre core area A_{core} of a normal fibre of 50 μm^2, the scaled contribution would amount to something of the order of $(1\text{ fs}^2/T_0^2)|a|^4a$. For a dispersion shifted fibre with a core area of 3 μm^2 it would be of the order of $(0.02\text{ fs}^2/T_0^2)|a|^4a$.

All in all, the lower limit of T_0 is set by the time scale of the stimulated Raman scattering. T_0 must be larger than 30 fs. However, probably a lower lower limit on T_0 can be reached if the stimulated Raman scattering term is treated exactly, without expanding the convolution product of (9) asymptotically.

References

[1] Afanasyev V V, Vysloukh V A and Serkin V N 1990 *Opt. Lett.* **15** 489–491
[2] Agrawal G P 1989 *Nonlinear fiber optics* (New York: Academic)
[3] Agrawal G P 1992 *Contemporary nonlinear optics* (*Quantum electronics—Principles and applications*) ed G P Agrawal and R W Boyd (Boston, MA: Academic) ch 2
[4] Akhmanov S A, Vysloukh V A and Chirkin A S 1992 *Optics of Femtosecond Laser Pulses* (New York: American
Institute of Physics)
[5] Blow K J, Doran, N J and Wood D 1988 *J. Opt. Soc. Am.* B **5** 1301–4
[6] Blow K J and Wood D 1989 *IEEE J. Quantum Electron.* **QE-25** 2665–73
[7] Bullough R K and Caudrey P J (ed) 1980 *Solitons* (Berlin: Springer)
[8] Butcher P N and Cotter D 1990 *The elements of nonlinear optics.* (*Cambridge studies in modern optics* **9**) (*Cambridge*: Cambridge University Press)
[9] Drazin P G and Johnson R S 1989 *Solitons: an introduction* (Cambridge: Cambridge University Press)
[10] Eilenberger G 1981 *Solitons* (*Solid-State Sciences* **19**) (Berlin: Springer)
[11] Gardner C S, Greene J M, Kruskal M D and Miura R M 1967 *Phys. Rev. Lett.* **19** 1095–7
[12] Gordon J P 1986 *Opt. Lett.* **11** 662–4
[13] Gordon J P and Haus H A 1986 *Opt. Lett.* **11** 665–7
[14] Grudinin A B, Dianov E M, Korobkin D V, Prokhorov A M, Serkin V N and Khaidarov D V 1987 *JETP Lett.* **46** 221–5
[15] Hasegawa A 1990 *Optical Solitons in Fibres* (*Springer Tracts in Modern Physics* **116**) (Berlin: Springer)
[16] Hasegawa A and Tappert F 1973 *Appl. Phys. Lett.* **23** 142–4
[17] Hasegawa A and Tappert F 1973 *Appl. Phys. Lett.* **23** 171–2
[18] Kodama Y 1985 *J. Stat. Phys.* **39** 597–614
[19] Kodama Y and Hasegawa A 1987 *IEEE J. Quantum Electron.* **QE-23** 510–24

[20] Loudon R 1973 *The Quantum Theory of Light* (Oxford: Clarendon)
[21] Mamyshev P V and Chernikov S V 1990 *Opt. Lett.* **15** 1076–8
[22] Mitschke F M and Mollenauer L F 1986 *Opt. Lett.* **11** 659–661
[23] Mollenauer L F, Gordon J P and Evangelides S G 1992 *Opt. Lett.* **17** 1575–7
[24] Mollenauer L F, Neubelt M J, Evangelides S G, Gordon J P, Simpson J R and Cohen L G 1990 *Opt. Lett.* **15** 1203–5
[25] Mollenauer L F, Stolen R H and Gordon J P 1980 *Phys. Rev. Lett.* **45** 1095–8
[26] Nayfeh A H 1973 *Perturbation methods* (New York: Wiley)
[27] Newell A C and Moloney J V 1992 *Nonlinear Optics* (Redwood City, CA: Addison-Wesley)
[28] Sturrock P A 1957 *Proc. R. Soc.* A **242** 277–99.
[29] Yariv A 1975 *Quantum Electronics* (New York: Wiley)
[30] Zakharov V E and Shabat A B 1971 *Zh. Eksp. Teor. Fiz.* **61** 118–34 (*Sov. Phys.–JETP* **34** 62–69)