7
Nonlinear Schrödinger models in nonlinear optics

An important and rich area of application of nonlinear wave propagation is the field of nonlinear optics. Asymptotic methods play an important role in this field. For example, since the scales are so disparate in Maxwell's equations, long-distance transmission in fiber optic communications depends critically on asymptotic models. Hence the nonlinear Schrödinger (NLS) equation is central for understanding phenomena and detailed descriptions of the dynamics. In this chapter we will outline the derivation of the NLS equation for electromagnetic wave propagation in bulk optical media. We also briefly discuss how the NLS equation arises as a model of spin waves in magnetic media.

7.1 Maxwell equations

We begin by considering Maxwell's equations for electromagnetic waves with no source charges or currents, cf. Landau et al. (1984) and Jackson (1998)

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} \tag{7.1a}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{7.1b}$$

$$\nabla \cdot \boldsymbol{D} = 0 \tag{7.1c}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{7.1d}$$

where $\boldsymbol{H}$ is the magnetic field, $\boldsymbol{E}$ is the electromagnetic field, $\boldsymbol{D}$ is the electromagnetic displacement and $\boldsymbol{B}$ is the magnetic induction.

We first consider non-magnetic media so there is no magnetization term in $\boldsymbol{B}$. The magnetic induction $\boldsymbol{B}$ and magnetic field $\boldsymbol{H}$ are then related by

$$\boldsymbol{B} = \mu_0 \boldsymbol{H}. \tag{7.2}$$

The constant μ_0 is the magnetic permeability of free space. We allow for induced polarization of the media giving rise to the following relation

$$\boldsymbol{D} = \epsilon_0(\boldsymbol{E} + \boldsymbol{P}), \tag{7.3}$$

where ϵ_0 is the electric permittivity of free space, a constant. When we apply an electric field $\boldsymbol{E}$ to an ideal dielectric material a response of the material can occur and the material is said to become polarized. Typically in these materials the electrons are tightly bound to the nucleus and a displacement of these electrons occurs. The macroscopic effect (summing over all displacements) yields the induced polarization $\boldsymbol{P}$.

Since we are working with non-magnetic media, we can reduce Maxwell's equations to those involving only the electromagnetic field $\boldsymbol{E}$ and the polarization $\boldsymbol{P}$. To that end, we take the curl of (7.1b) and use (7.2) and (7.1a) to find

$$\begin{aligned}\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{E}) &= -\frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \boldsymbol{B}) \\ &= -\mu_0 \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \boldsymbol{H}) \\ &= -\mu_0 \frac{\partial^2}{\partial t^2}\boldsymbol{D}.\end{aligned}$$

The vector identity

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{E}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}, \tag{7.4}$$

is useful here. With this, along with relation (7.3), we find the following equations for the electromagnetic field

$$\begin{gathered}\nabla^2 \boldsymbol{E} - \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{E}) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(\boldsymbol{E} + \boldsymbol{P}(\boldsymbol{E})) \\ \boldsymbol{\nabla} \cdot (\boldsymbol{E} + \boldsymbol{P}(\boldsymbol{E})) = 0,\end{gathered} \tag{7.5}$$

where the constant $c^2 = \dfrac{1}{\mu_0 \epsilon_0}$ is the square of the speed of light in a vacuum. Notice that we have made explicit the connection between the polarization $\boldsymbol{P}$ and the electromagnetic field $\boldsymbol{E}$.

Before we investigate polarizable, non-magnetic media in detail, let us also write down the dual equation for magnetic media without polarization ($\boldsymbol{P} = 0$) satisfying

$$\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M}(\boldsymbol{H})) \tag{7.6}$$

where $\boldsymbol{M}$ is called the magnetization. We see that the magnetization vector $\boldsymbol{M}$ plays a similar role to that of the polarization $\boldsymbol{P}$, i.e., we are assuming that $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{H})$. Namely, we assume that the magnetization is related to the magnetic field. The magnetization can be permanent for ferromagnetic materials (permanent magnets such as iron). Taking the curl of (7.1a) then for non-polarized media using $\boldsymbol{D} = \epsilon_0 \boldsymbol{E}$ gives

$$\begin{aligned}\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{H}) &= \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \boldsymbol{D}) \\ &= \epsilon_0 \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \boldsymbol{E}) \\ &= -\epsilon_0 \frac{\partial^2}{\partial t^2}\boldsymbol{B}.\end{aligned}$$

Using (7.4) and (7.6) gives

$$\begin{gathered}\nabla^2 \boldsymbol{H} - \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{H}) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(\boldsymbol{H} + \boldsymbol{M}(\boldsymbol{H})) \\ \boldsymbol{\nabla} \cdot (\boldsymbol{H} + \boldsymbol{M}(\boldsymbol{H})) = 0.\end{gathered}$$

Later, in Section 7.4 we will discuss one way $\boldsymbol{M}$ can be coupled to $\boldsymbol{H}$.

7.2 Polarization

In homogeneous, non-magnetic media, matter responds to intense electromagnetic fields in a nonlinear manner. In order to model this, we use a well-known relationship between the polarization vector $\boldsymbol{P}$ and the electromagnetic field $\boldsymbol{E}$ that is a good approximation to a wide class of physically relevant media:

$$\boldsymbol{P}(\boldsymbol{E}) = \int \chi^{(1)} * \boldsymbol{E} + \int \chi^{(2)} * \boldsymbol{E}\boldsymbol{E} + \int \chi^{(3)} * \boldsymbol{E}\boldsymbol{E}\boldsymbol{E}. \tag{7.7}$$

The above equation involves tensors, so the operation $*$ is a special type of convolution that will be defined below.

For notational purposes, we will write the polarization vector as follows:

$$\boldsymbol{P} = \begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} \equiv \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}.$$

Notice that we can break up (7.7) into a linear and nonlinear part

$$\boldsymbol{P} = \boldsymbol{P}_{\mathrm{L}} + \boldsymbol{P}_{\mathrm{NL}}.$$

In glass, and hence fiber optics, the quadratic term $\chi^{(2)}$ is zero. This is cubically nonlinear and is a so-called "centro-symmetric material", cf. Agrawal (2001), Boyd (2003) and also Ablowitz et al. (1997, 2001a) and included references for more information. We write the linear and nonlinear components of the polarization vector, and explicitly define the convolution mentioned above, as follows:

$$P_{\mathrm{L},i} = (\chi^{(1)} * \boldsymbol{E})_i = \sum_{j=1}^{3} \int_{-\infty}^{\infty} \chi_{ij}^{(1)}(t-\tau)E_j(\tau)\,d\tau \tag{7.8}$$

$$P_{\mathrm{NL},i} = \sum_{j,k,l} \int_{-\infty}^{\infty} \chi_{ijkl}^{(3)}(t-\tau_1, t-\tau_2, t-\tau_3)E_j(\tau_1)E_k(\tau_2)E_l(\tau_3)\,d\tau_1 d\tau_2 d\tau_3. \tag{7.9}$$

In (7.9), the sums are over the indices $\{1,2,3\}$ and the integration is over all of $\mathbb{R}^3$. The matrix $\chi^{(1)}$ is called the linear susceptibility and the tensor $\chi^{(3)}$ is the third-order susceptibility. If the material is "isotropic", $\chi_{ij}^{(1)} = 0$, $i \neq j$ (exhibiting properties with the same values when measured along axes in all directions), then the matrix $\chi^{(1)}$ is diagonal. We will also usually identify subscripts $i = 1,2,3$ as $i = x,y,z$. In cubically nonlinear or "Kerr" materials (such as glass), it turns out that the only important terms in the $\chi^{(3)}$ tensor correspond to $\chi_{xxxx}^{(3)}$, which is equal to $\chi_{yyyy}^{(3)}$ and $\chi_{zzzz}^{(3)}$.

From now on, we will suppress the superscript in χ when context makes the choice clear. For example $\chi_{xx}^{(1)} = \chi_{xx}$ and $\chi_{xxxx}^{(3)} = \chi_{xxxx}$ follow due to the number of entries in the subscript.

We can now pose the problem of determining the electromagnetic field $\boldsymbol{E}$ in Kerr media as solving (7.5) subject to (7.8) and (7.9). Consider the asymptotic expansion

$$\boldsymbol{E} = \varepsilon \boldsymbol{E}^{(1)} + \varepsilon^2 \boldsymbol{E}^{(2)} + \varepsilon^3 \boldsymbol{E}^{(3)} + \cdots, \; |\varepsilon| \ll 1. \tag{7.10}$$

Then, to leading order, we assume that the electromagnetic field is initially polarized along the x-axis; it propagates along the z-axis, and for simplicity we assume no transverse y variations. So

$$\boldsymbol{E}^{(1)} = \begin{pmatrix} E_x^{(1)} \\ 0 \\ 0 \end{pmatrix}, \qquad E_x^{(1)} = A(X,Z,T)e^{i\theta} + \text{c.c.}, \tag{7.11}$$

where $\theta = kz - \omega t$ and the amplitude A is assumed to be slowly varying in the x, z, and t directions:

$$X = \varepsilon x, \quad Z = \varepsilon z, \quad T = \varepsilon t.$$

In order to simplify the calculations, we assume $\boldsymbol{E}$ and hence A are independent of y (i.e., $Y = \varepsilon y$). Derivatives are replaced as follows

$$\begin{aligned} \partial_x &= \varepsilon\partial_X \\ \partial_t &= -\omega\partial_\theta + \varepsilon\partial_T \\ \partial_z &= k\partial_\theta + \varepsilon\partial_Z. \end{aligned} \tag{7.12}$$

Substituting in the asymptotic expansion for $\boldsymbol{E}$ we note the important simplification that the assumption of slow variation has on the polarization $\boldsymbol{P}$:

$$\begin{aligned} P_{\mathrm{L},x} &= \sum_{j=1}^{3} \int_{-\infty}^{\infty} \chi_{xj}^{(1)}(t-\tau)E_j(\tau)\,d\tau \\ &= \varepsilon \int_{-\infty}^{\infty} \chi_{xx}(t-\tau)\left(A(X,Z,T)e^{i(kz-\omega\tau)} + \text{c.c.}\right)d\tau + O(\varepsilon^2) \\ &= \varepsilon \int_{-\infty}^{\infty} \chi_{xx}(t-\tau)e^{i\omega(t-\tau)}\left(A(X,Z,T)e^{i(kz-\omega t)} + \text{c.c.}\right)d\tau + O(\varepsilon^2). \end{aligned}$$

Now make the substitution $t - \tau = u$ to get

$$P_{\mathrm{L},x} = \varepsilon \int_{-\infty}^{\infty} \chi_{xx}(u)e^{i\omega u}\left(A(X,Z,\varepsilon t - \varepsilon u)e^{i(kz-\omega t)} + \text{c.c.}\right)du + O(\varepsilon^2).$$

We expand the slowly varying amplitude A around the point εt

$$\begin{aligned} P_{\mathrm{L},x} =& \varepsilon \int_{-\infty}^{\infty} du\, \chi_{xx}(u)e^{i\omega u}\times \\ & \left[\left(1 - \varepsilon u\frac{\partial}{\partial T} + \frac{(\varepsilon u)^2}{2}\frac{\partial^2}{\partial T^2} + \cdots\right)A(X,Z,T)e^{i(kz-\omega t)} + \text{c.c.}\right] + O(\varepsilon^2). \end{aligned}$$

Recall that the Fourier transform of χ_{xx}, written as $\hat{\chi}_{xx}$, and its derivatives are

$$\begin{aligned} \hat{\chi}_{xx}(\omega) &= \int_{-\infty}^{\infty} \chi_{xx}(u)e^{i\omega u}\,du \\ \hat{\chi}'_{xx}(\omega) &= \int_{-\infty}^{\infty} iu\chi_{xx}(u)e^{i\omega u}\,du \\ \hat{\chi}''_{xx}(\omega) &= \int_{-\infty}^{\infty} -u^2\chi_{xx}(u)e^{i\omega u}\,du. \end{aligned}$$

Then we can write the linear part of the polarization as

$$\begin{aligned} P_{\mathrm{L},x} &= \varepsilon\left(\hat{\chi}_{xx}(\omega) + \hat{\chi}'_{xx}(\omega)i\varepsilon\partial_T - \hat{\chi}''_{xx}(\omega)\frac{(\varepsilon\partial_T)^2}{2} + \cdots\right)(Ae^{i\theta} + \text{c.c.}) + O(\varepsilon^2) \\ &= \varepsilon\hat{\chi}_{xx}(\omega + i\varepsilon\partial_T)(Ae^{i\theta} + \text{c.c.}) + O(\varepsilon^2). \end{aligned}$$

The last line is in a convenient notation where the term $\hat{\chi}_{xx}(\omega + i\varepsilon\partial_T)$ is an operator that we can expand around ω to get the previous line.

In the nonlinear polarization equation (7.9), there will be interactions due to the leading-order mode $Ae^{i\theta}$, for example, the nonlinear term includes E_x^3, which leads to terms such as $A^3e^{3i\theta}$. Then we denote the corresponding linear polarization term (containing interaction terms) as

$$P_{\mathrm{L},x}^{\text{interactions}} = \hat{\chi}_{xx}(\omega_m + i\varepsilon\partial_T)(B_m e^{im\theta} + \text{c.c.}),$$

where $\omega_m = m\omega$ and B_m contains the nonlinear terms generated by the interaction (e.g. $B_3 = A^3$).

Similarly there are nonlinear polarization terms. For example, a typical term in the nonlinear polarization is

$$\begin{aligned} P_{\mathrm{NL},x} = {} & \varepsilon^3\hat{\chi}_{xxxx}(\omega_m + i\varepsilon\partial_{T_1}, \omega_n + i\varepsilon\partial_{T_2}, \omega_l + i\varepsilon\partial_{T_3}) \\ & \times B_m(T_1)B_n(T_2)B_l(T_3)\big|_{T_1=T_2=T_3=T}\, e^{i(m+n+l)\theta} + \text{c.c.} \end{aligned} \tag{7.13}$$

All other terms are of smaller order $O(\varepsilon^4)$.

7.3 Derivation of the NLS equation

So far, we have derived expressions for the polarization with a cubic nonlinearity (third-order susceptibility). Now we will use (7.5) to derive the NLS equation. This will give the leading-order equation for the slowly varying amplitude of the x-component of the electromagnetic field E_x.

We begin the derivation of the NLS equation by showing that $E_z \ll E_x$. Recall (7.1c) with (7.3),

$$\nabla \cdot (\boldsymbol{E} + \boldsymbol{P}) = 0.$$

Since there is no y dependence, we get

$$\partial_x(E_x + P_x) + \partial_z(E_z + P_z) = 0. \tag{7.14}$$

We assume

$$\begin{aligned} E_x &= \varepsilon A(X, Z, T)e^{i\theta} + \text{c.c.} + O(\varepsilon^2) \\ E_z &= \varepsilon A_z(X, Z, T)e^{i\theta} + \text{c.c.} + O(\varepsilon^2), \end{aligned} \tag{7.15}$$

where as before c.c. denotes the complex conjugate of the preceding term. A further note on notation. Because the x-component of the electromagnetic field E_x appears so often, we label its slowly varying amplitude with A, whereas for the z-component of the E-field, E_z, we label the slowly varying amplitude with A_z (note: A_z does not mean derivative with respect to z).

Substituting the derivatives (7.12) into (7.14) we find

$$(k\partial_\theta + \varepsilon\partial_Z)(E_z + P_z) = -\varepsilon\partial_X(E_x + P_x),$$

which, when we use the ansatz (7.15), yields

$$\begin{aligned}&\varepsilon(k\partial_\theta + \varepsilon\partial_Z)\Big(A_z e^{i\theta} + \text{c.c.} + \hat{\chi}_{zz}(\omega + i\varepsilon\partial_T)(A_z e^{i\theta} + \text{c.c.})\Big)\\ &= -\varepsilon\partial_X\Big(\varepsilon A e^{i\theta} + \text{c.c.} + \hat{\chi}_{xx}(\omega + i\varepsilon\partial_T)(\varepsilon A e^{i\theta} + \text{c.c.})\Big) + \cdots .\end{aligned}$$

The leading-order equation is

$$\begin{aligned}\varepsilon i k A_z(1 + \hat{\chi}_{zz}(\omega)) &= -\varepsilon^2 \frac{\partial A}{\partial X}(1 + \hat{\chi}_{xx}(\omega)) \qquad \text{implying}\\ A_z &= -\varepsilon \frac{1 + \hat{\chi}_{xx}(\omega)}{ik(1 + \hat{\chi}_{zz}(\omega))}\frac{\partial A}{\partial X}\\ &= \frac{-\varepsilon}{ik}\frac{\partial A}{\partial X} = A_X, \quad \text{if} \quad \hat{\chi}_{xx} = \hat{\chi}_{zz}.\end{aligned}$$

Assuming $\partial A/\partial X$ is $O(1)$, this implies $A_z = O(\varepsilon)$, which implies $E_z = O(\varepsilon^2)$. While we used the divergence equation in (7.5) with (7.3) to calculate this relationship, we could have proved the same result with the z-component of the dynamic equation in (7.5). We also claim that from the dynamic equation of motion in (7.5), if we had studied the y-component, then it would have followed that $E_y = O(\varepsilon^3)$. Also, $P_{\text{NL},z} = O(\varepsilon^4)$, which is obtained from the symmetry of the third-order susceptibility tensor $\chi^{(3)}$ since $\chi_{zxxx} = 0$, etc. (e.g., for glass).

Now we will use the dynamic equation in (7.5) to investigate the behavior of the x-component of the electromagnetic field via multiple scales. The dynamic equation (7.5) becomes

$$\nabla^2 E_x - \frac{\partial}{\partial x}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z}\right) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(E_x + P_x). \tag{7.16}$$

The asymptotic expansions for E_x and E_z take the form

$$\begin{aligned}E_x &= \varepsilon(A e^{i\theta} + \text{c.c.}) + \varepsilon^2 E_x^{(2)} + \varepsilon^3 E_x^{(3)} + \cdots\\ E_z &= \varepsilon^2(A_z e^{i\theta} + \text{c.c.}) + \varepsilon^3 E_z^{(2)} + \cdots .\end{aligned} \tag{7.17}$$

The polarization in the x-direction depends on the linear and nonlinear polarization terms, $P_{\mathrm{L},x}$ and $P_{\mathrm{NL},x}$, but the nonlinear term comes in at $O(\varepsilon^3)$ due to the cubic nonlinearity. Explicitly, we have

$$P_x = \varepsilon(\hat{\chi}_{xx}(\omega) + \hat{\chi}'_{xx}(\omega)i\varepsilon\partial_T - \frac{\hat{\chi}''_{xx}(\omega)}{2}\varepsilon^2\partial_T^2 + \cdots)(Ae^{i\theta} + \text{c.c.}) + P_{\mathrm{NL},x}. \quad (7.18)$$

Expanding the derivatives and using (7.12) in (7.16) gives

$$\begin{aligned}\left((k\partial_\theta + \varepsilon\partial_Z)^2 + \varepsilon^2\partial_X^2\right)E_x - \varepsilon\partial_X\Big(\varepsilon\partial_X E_x + (k\partial_\theta + \varepsilon\partial_Z)E_z\Big)\\ = \frac{1}{c^2}(-\omega\partial_\theta + \varepsilon\partial_T)^2(E_x + P_{\mathrm{L},x} + P_{\mathrm{NL},x}).\end{aligned} \quad (7.19)$$

The leading-order equation is

$$O(\varepsilon): \qquad \left(-k^2 + \left(\frac{\omega}{c}\right)^2\right)A = -\left(\frac{\omega}{c}\right)^2\hat{\chi}_{xx}(\omega)A.$$

Since we assume $A \neq 0$ the above equation determines the dispersion relation; namely solving for $k(\omega)$ we find the dispersion relation

$$\begin{aligned} k^2 &= \left(\frac{\omega}{c}\right)^2(1 + \hat{\chi}_{xx}(\omega)) \\ k(\omega) &= \frac{\omega}{c}(1 + \hat{\chi}_{xx}(\omega))^{1/2} \\ &\equiv \frac{\omega}{c}n_0(\omega). \end{aligned} \quad (7.20)$$

The term $n_0(\omega)$ is called the linear index of refraction; if it were constant then c/n_0 would be the “effective speed of light”. The slowly varying amplitude $A(X, Z, T)$ is still free and will be determined by going to higher order and removing secular terms.

From (7.19) and the asymptotic expansions for E_x and $P_{\mathrm{L},x}$ in (7.17) and (7.18) (the nonlinear polarization term, $P_{\mathrm{NL},x}$, does not come in until the next order), we find

$$\begin{aligned} O(\varepsilon^2): \qquad \left(k^2 - \left(\frac{\omega}{c}\right)^2\right)\partial_\theta^2 E_x^{(2)} = \Bigg[&-2ik\partial_Z A - \frac{2i\omega}{c^2}(1 + \hat{\chi}_{xx}(\omega))\partial_T A \\ &- \left(\frac{\omega}{c}\right)^2 i\hat{\chi}'_{xx}(\omega)\partial_T A\Bigg]e^{i\theta}. \end{aligned}$$

To remove secularities, we must equate the terms in brackets to zero. This can be written as

$$2k\frac{\partial A}{\partial Z} + \left(\frac{2\omega}{c^2}(1 + \hat{\chi}_{xx}(\omega)) + \left(\frac{\omega}{c}\right)^2\hat{\chi}'_{xx}(\omega)\right)\frac{\partial A}{\partial T} = 0. \quad (7.21)$$

Notice that if we differentiate the dispersion relation (7.20) with respect to ω we find

$$2kk' = \frac{2\omega}{c^2}(1 + \hat{\chi}_{xx}(\omega)) + \left(\frac{\omega}{c}\right)^2 \hat{\chi}'_{xx}(\omega).$$

This means that we can rewrite (7.21) in terms of the group velocity $v_g = 1/k'(\omega)$:

$$\frac{\partial A}{\partial Z} + \frac{1}{v_g}\frac{\partial A}{\partial T} = 0. \tag{7.22}$$

This is a first-order equation that can be solved by the method of characteristics giving, to this order, $A = A(T - z/v_g)$. To obtain a more accurate equation for the slowly varying amplitude A, we will now remove secular terms at the next order and obtain the nonlinear Schrödinger (NLS) equation.

As before, we perturb (7.22)

$$2ik\left(\frac{\partial A}{\partial Z} + \frac{1}{v_g}\frac{\partial A}{\partial T}\right) = \varepsilon f_1 + \varepsilon^2 f_2 + \cdots .$$

Using this and (7.19) to $O(\varepsilon^3)$ and then, as we have discussed in earlier chapters, choosing f_1 to remove secular terms, we find after some calculation

$$\begin{aligned} O(\varepsilon^3): \qquad & 2ik\left(\frac{\partial A}{\partial Z} + k'(\omega)\frac{\partial A}{\partial T}\right) \\ & +\varepsilon\left[\partial_X^2 + \partial_Z^2 - \frac{1}{c^2}(1 + \hat{\chi}_{xx}(\omega) + 2\omega\hat{\chi}'_{xx}(\omega) + \frac{1}{2}\omega^2\hat{\chi}''_{xx}(\omega))\partial_T^2\right]A \\ & +3\varepsilon\left(\frac{\omega}{c}\right)^2\hat{\chi}_{xxxx}(\omega,\omega,-\omega)|A|^2A = 0. \end{aligned} \tag{7.23}$$

We now see in (7.23) a term due to the nonlinear polarization $P_{\mathrm{NL},x}$ i.e., the term with $\hat{\chi}_{xxxx}(\omega,\omega,-\omega)$. In fact, there are three choices for m, n, p in the leading-order term for the nonlinear polarization [recall (7.13)] that give rise to the exponential $e^{i\theta}$ (the coefficient of $e^{i\theta}$ is the only term that induces secularity). Namely, two of these integers are $+1$ and the other one is -1. In the bulk media we are working in, the third-order susceptibility tensor is the same in all cases. This means

$$\begin{aligned} \hat{\chi}_{xxxx}(\omega,\omega,-\omega) + \hat{\chi}_{xxxx}(\omega,-\omega,\omega) + \hat{\chi}_{xxxx}(-\omega,\omega,\omega) \\ = 3\hat{\chi}_{xxxx}(\omega,\omega,-\omega) \\ \equiv 3\hat{\chi}_{xxxx}(\omega). \end{aligned}$$

Let us investigate the coefficient of A_{TT} due to the linear polarization. If we differentiate the dispersion relation (7.20) two times with respect to ω we get the following result

$$k'^2 + kk'' = \frac{1}{c^2}\left(1 + \hat{\chi}_{xx}(\omega) + 2\omega\hat{\chi}'_{xx}(\omega) + \tfrac{1}{2}\omega^2\hat{\chi}''_{xx}(\omega)\right).$$

Then we can conveniently rewrite the linear polarization term as

$$\varepsilon\left[\,\partial_X^2 + \partial_Z^2 - (k'^2 + kk'')\partial_T^2\right]A.$$

Now, make the following change of variable

$$\xi = T - k'(\omega)Z, \quad Z' = \varepsilon Z$$
$$\partial_Z = -k'\partial_\xi + \varepsilon\partial_{Z'}, \quad \partial_T = \partial_\xi.$$

Substituting this into the third-order equation (7.23), we find

$$2ik\frac{\partial A}{\partial Z'} + \frac{\partial^2 A}{\partial X^2} - kk''\frac{\partial^2 A}{\partial \xi^2} + 2k\nu|A|^2 A = 0. \tag{7.24}$$

Here we took

$$\nu = \frac{3(\omega/c)^2\hat{\chi}_{xxxx}(\omega)}{2k}. \tag{7.25}$$

It is useful to note that had we included variation in the y-direction, then the term $\partial^2 A/\partial Y^2$ would be added to (7.24) and we would have found

$$2ik\frac{\partial A}{\partial Z'} + \nabla^2 A - kk''\frac{\partial^2 A}{\partial \xi^2} + 2k\nu|A|^2 A = 0, \tag{7.26}$$

where

$$\nabla^2 A = \frac{\partial^2 A}{\partial X^2} + \frac{\partial^2 A}{\partial Y^2}.$$

If there is no variation in the X-direction then we get the $(1+1)$-dimensional (one space dimension and one time dimension) nonlinear Schrödinger (NLS) equation describing the slowly varying wave amplitude A of an electromagnetic wave in cubically nonlinear bulk media

$$i\frac{\partial A}{\partial Z} + \left(\frac{-k''(\omega)}{2}\right)\frac{\partial^2 A}{\partial \xi^2} + \nu|A|^2 A = 0, \tag{7.27}$$

where we have removed the prime from Z' for convenience. If $k'' < 0$ then we speak of anomalous dispersion. The NLS equation is then called "focusing" and gives rise to "bright" soliton solutions as discussed in Chapter 6 (note: $\nu > 0$ and $k > 0, \hat{\chi}_{xxxx} > 0$). If $k'' > 0$ then the system is said to have normal dispersion and the NLS equation is called "defocusing". This equation then admits "dark" soliton solutions (see also Chapter 6).

Next we remark about the coefficient of nonlinearity ν defined in (7.25). The previous derivation of the NLS equation was for electromagnetic waves in bulk optical media. In the field of optical communications, one is interested in waves propagating down narrow fibers. Perhaps surprisingly, the NLS arises again in exactly the same form as (7.27) with only a small change in the value of ν:

$$\nu_{\text{eff}} = \frac{\nu}{A_{\text{eff}}} = \frac{3(\omega/c)^2 \hat{\chi}_{xxxx}(\omega)}{2kA_{\text{eff}}} \;;$$

thus ν only changes from its vacuum value by an additional factor in the denominator. Here, A_{eff} corresponds to the effective cross-sectional area of the fiber; see Hasegawa and Kodama (1995) and Agrawal (2001) for more information.

We often write ν in a slightly different way,

$$\nu = \frac{\omega}{c}\tilde{n}_2(\omega),$$

so that the nonlinear index of refraction $\tilde{n}_2$ is

$$\tilde{n}_2(\omega) = \frac{3\omega}{2kc}\hat{\chi}_{xxxx}(\omega).$$

We can relate the linear and quadratic indices of refraction ($n_0(\omega)$ and $\tilde{n}_2(\omega)$) to the Stokes frequency shift as follows. We assume the wave amplitude has the following ansatz

$$A(Z,\xi) = \tilde{A}(Z)e^{-ikZ},$$

called a CW or continuous wave, i.e., this is the electromagnetic analog of a Stokes water wave. Substituting this into (7.27) gives

$$\begin{aligned} -i\frac{\partial \tilde{A}}{\partial Z} &= k\tilde{A} + \nu|\tilde{A}|^2\tilde{A} \\ &= \frac{\omega}{c}(n_0 + \tilde{n}_2|\tilde{A}|^2)\tilde{A}. \end{aligned} \tag{7.28}$$

If we multiply (7.28) by $\tilde{A}^*$ then add that to the conjugate of (7.28) multiplied by $\tilde{A}$, we see that

$$\frac{\partial}{\partial Z}|\tilde{A}|^2 = 0 \quad \Rightarrow \quad |\tilde{A}(Z)|^2 = |\tilde{A}(0)|^2.$$

Now we can solve (7.28) directly to find

$$\tilde{A}(Z) = \tilde{A}(0) \exp\left(i\frac{\omega}{c}(n_0 + \tilde{n}_2|\tilde{A}(0)|^2)Z\right).$$

The additional "nonlinear" frequency shift is $\tilde{n}_2|\tilde{A}(0)|^2$. This term is often referred to as self-phase-modulation. Hence the wavenumber, k, is now a function of the initial amplitude and the frequency

$$k(\omega, |\tilde{A}(0)|) = \frac{\omega}{c}(n_0 + \tilde{n}_2|\tilde{A}(0)|^2).$$

In the literature, the index of refraction combining both the linear and quadratically nonlinear indices is often written in terms of the initial electric field $\boldsymbol{E}$

$$n(\omega, \boldsymbol{E}(0)) = n_0(\omega) + n_2(\omega)|E_x(0)|^2.$$

Since we assumed the form

$$E_x(0) = \varepsilon(A(0)e^{i\theta} + \text{c.c.}) = \varepsilon(\tilde{A}(0)e^{-i\omega t} + \text{c.c.}) = 2\varepsilon|\tilde{A}(0)|\cos(\omega t + \phi_0),$$

for the electric field, we see that

$$|\tilde{A}(0)| = \frac{|\boldsymbol{E}_x(0)|}{2} \quad \Rightarrow \quad n_2 = \frac{\tilde{n}_2}{4}.$$

Including quadratic nonlinear media; NLS with mean terms

We will briefly discuss the effects of including a non-zero quadratic nonlinear polarization. This implies that the second-order susceptibility tensor $\chi^{(2)}$ in the nonlinear polarization is non-zero.

Recall the definition of the nonlinear polarization for quadratic and cubic media is written schematically as,

$$\boldsymbol{P}_{NL}(\boldsymbol{E}) = \int \chi^{(2)} : \boldsymbol{EE} + \int \chi^{(3)} \vdots \boldsymbol{EEE} + \ldots$$

where the symbols ":" and "$\vdots$" represent tensor notation. The ith component of the quadratic term in the nonlinear polarization is written

$$P^{(2)}_{\text{NL},i} = \sum_{j,k} \int_{-\infty}^{\infty} \chi_{ijk}(t - \tau_1, t - \tau_2) E_j(\tau_1) E_k(\tau_2)\, d\tau_1 d\tau_2.$$

In the derivation of the nonlinear Schrödinger equation with cubic nonlinearity, the nonlinear terms entered the calculation at third order. The multi-scale procedure for $\chi^{(2)} \neq 0$ follows the same lines as before:

$$\boldsymbol{E} = \varepsilon \boldsymbol{E}^{(1)} + \varepsilon^2 \boldsymbol{E}^{(2)} + \varepsilon^3 \boldsymbol{E}^{(3)} + \cdots$$

$$\boldsymbol{E}^{(1)} = \begin{pmatrix} E_x^{(1)} \\ 0 \\ 0 \end{pmatrix}$$

$$E_x^{(1)} = A(X, Y, Z, T)e^{i\theta} + \text{c.c.},$$

see (7.10) and (7.11); to be general, we allow the amplitude A to depend on $X = \varepsilon x, Y = \varepsilon y, Z = \varepsilon z$ and $T = \varepsilon t$. The new terms of interest due to the quadratic nonlinearity are (Ablowitz et al., 1997, 2001a):

$$E_x^{(2)} = 4\left(\frac{\omega}{c}\right)^2 \frac{\hat{\chi}_{xxx}}{\Delta(\omega)}\left(A^2 e^{2i\theta} + \text{c.c.}\right) + \phi_x$$

$$E_y^{(2)} = \phi_y$$

$$E_z^{(2)} = i\frac{n_x^2}{kn_z^2}\frac{\partial A}{\partial X}e^{i\theta} + \text{c.c.} + \phi_z$$

$$n_x^2 = 1 + \hat{\chi}_{xx}(\omega), \quad n_z^2 = 1 + \hat{\chi}_{zz}(\omega), \quad k = \frac{\omega}{c}n_x(\omega)$$

$$\Delta(\omega) = (2k(\omega))^2 - (k(2\omega))^2 \neq 0.$$

We assume that $\chi_{yy} = \chi_{xx}$. Notice the inclusion of "mean" terms ϕ_x, ϕ_y, ϕ_z. If $|\Delta(\omega)| \ll 1$ then second harmonic resonance generation occurs (Agrawal, 2002). We will assume $\Delta(\omega) \neq 0$ and $\Delta(\omega)$ is not small.

In our derivation of NLS for cubic media, the only term at third order that gave rise to secularity was $e^{i\theta}$. Now, at third order, we have two sources for secularity, $e^{i\theta}$ and the mean term e^{i0} (see also the derivation of NLS from KdV in Chapter 6). Removing these secular terms in the usual way as before leads to the following equations:

$$O(\varepsilon^3): \quad e^{i\theta} \to 2ik\partial_Z A + \left(\partial_X^2 + \partial_Y^2 - kk''\partial_\xi^2\right)A$$
$$+ (M_1|A|^2 + M_0\phi_x)A = 0 \tag{7.29}$$

$$e^{i0} \to \left(\alpha_x\partial_X^2 + \partial_Y^2 + s_x\partial_\xi^2\right)\phi_x - \left(N_1\partial_\xi^2 - N_2\partial_X^2\right)|A|^2 = 0. \tag{7.30}$$

In this, the change of variable, $\xi = T - k'(\omega)Z$, was made and the following definitions used:

$$M_0 = 2\left(\frac{\omega}{c}\right)^2 \hat{\chi}_{xxx}(\omega, 0), \qquad M_1 = 3\left(\frac{\omega}{c}\right)^2 \hat{\chi}_{xxxx}(\omega, \omega, -\omega)$$
$$+ \frac{8(\omega/c)^4 \hat{\chi}_{xxx}(2\omega, -\omega)\hat{\chi}_{xxx}(\omega, \omega)}{\Delta(\omega)},$$

$$N_1 = \frac{2}{c^2}\hat{\chi}_{xxx}(\omega, -\omega), \qquad N_2 = \frac{c^2 N_1}{1+\hat{\chi}_{xx}(\omega)},$$
$$\alpha_x = \frac{1+\hat{\chi}_{xx}(\omega)}{1+\hat{\chi}_{zz}(\omega)}, \qquad s_x = k'(\omega)^2 - \frac{1+\hat{\chi}_{xx}(\omega)}{c^2}.$$

There are several things to notice. First, if $\hat{\chi}_{xxx} = 0$ then $N_1 = N_2 = M_0 = 0$, $M_1 = 2(\omega/c)^2\hat{\chi}_{xxx}(\omega, -\omega)$. Thus (7.29) reduces to the nonlinear Schrödinger equation obtained earlier for cubic media. Since (7.29) of NLS-type but includes a mean term ϕ_x we call it NLS with mean (NLSM). We must solve these coupled equations (7.29) and (7.30) for ϕ_x and A. It turns out, when $\chi_{xx} = \chi_{yy}$, that the mean terms ϕ_y, ϕ_z decouple from the ϕ_x equation and they can be solved in terms of ϕ_x and A (Ablowitz et al., 1997, 2001a). These equations, (7.29) and (7.30), are a three-dimensional generalization of the Benney–Roskes (BR) type equations which we discussed in our study of multi-dimensional water waves in Chapter 6. Moreover, when A is independent of time, t, i.e., independent of ξ, then (7.29) and (7.30) are $(2+1)$-dimensional and reduce to the BR form discussed in Section 6.8 and wave collapse is possible (see also Ablowitz et al., 2005).

7.4 Magnetic spin waves

In this section, we will briefly discuss a problem involving magnetic materials. As discussed in Section 7.1, in this case, we neglect the polarization, $\boldsymbol{P} = 0$, but we do include magnetization $\boldsymbol{M}(\boldsymbol{H})$.

The typical situation is schematically depicted in Figure 7.1 Kalinikos et al., 1997; Chen et al., 1994; Tsankov et al., 1994; Patton et al., 1999; Kalinikos et al., 2000; Wu et al., 2004, and references therein). Consider a slab of magnetic material with magnetization $\boldsymbol{M}_s$ of thickness d. Suppose an external magnetic field $\boldsymbol{G}_s$ is applied from above and below the slab. We wish to determine an equation for the magnetic field $\boldsymbol{H}$ inside the slab. The imposed magnetic field $\boldsymbol{G}_s$ induces a magnetization $\boldsymbol{m}$ inside the material. The total magnetization inside the slab is taken to be

$$\boldsymbol{M} = \boldsymbol{M}_s + \boldsymbol{m}.$$

We write the magnetic fields inside and outside the slab as

$$\boldsymbol{H} = \boldsymbol{H}_s + \boldsymbol{h}, \qquad \boldsymbol{G} = \boldsymbol{G}_s + \boldsymbol{g}.$$

A relation between the steady state values of the internal and external fields as well as the magnetization, $\boldsymbol{H}_s$, $\boldsymbol{G}_s$, and $\boldsymbol{M}_s$, will be determined by the boundary

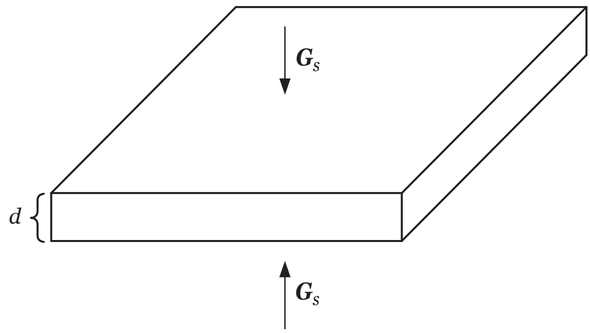

Figure 7.1 Magnetic slab.

conditions. Once a perturbation is introduced into the system, $\boldsymbol{h}$, $\boldsymbol{g}$, $\boldsymbol{m}$ are non-trivial. The goal is to find the equations governing these quantities.

Recall from Section 7.1 the dynamical equations for the magnetic field $\boldsymbol{H}$ and the magnetization $\boldsymbol{M}$ with no source currents or charge:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{H}) = \nabla^2 \boldsymbol{H} - \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{H}) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(\boldsymbol{H} + \boldsymbol{M}(\boldsymbol{H})) \tag{7.31}$$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{H} + \boldsymbol{M}(\boldsymbol{H})) = 0. \tag{7.32}$$

We also need a relation between the magnetic field $\boldsymbol{H}$ and the magnetization $\boldsymbol{M}$. The approximation we will work with, assuming no damping, is the so-called "torque" equation

$$\frac{\partial \boldsymbol{M}}{\partial t} = -\gamma \boldsymbol{M} \times \boldsymbol{H}.$$

Since $1/c^2$ is very small given the scales we are interested in, we neglect the time derivative term in (7.31) and get the following quasistatic approximation for the magnetic field inside the slab:

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{H} &= 0 \\ \boldsymbol{\nabla} \cdot (\boldsymbol{H} + \boldsymbol{M}) &= 0 \\ \frac{\partial \boldsymbol{M}}{\partial t} &= -\gamma \boldsymbol{M} \times \boldsymbol{H}. \end{aligned} \tag{7.33}$$

Outside the slab we have

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{G} &= 0 \\ \boldsymbol{\nabla} \cdot \boldsymbol{G} &= 0. \end{aligned} \tag{7.34}$$

The coordinate origin is taken to be in the center of the slab and the z-direction is vertical thus $\boldsymbol{G}_s = G_s \hat{\boldsymbol{z}}$. We also take $\boldsymbol{H}_s = H_s \hat{\boldsymbol{z}}$, $\boldsymbol{M}_s = M_s \hat{\boldsymbol{z}}$, with

G_s, H_s, M_s all constant. The boundary conditions (since we have no sources) are derived from the continuity of the induction and magnetic fields that can be written as

$$H_x = G_x, \quad H_y = G_y, \quad H_s + M_s = G_s \quad \text{all at } z = \pm\frac{d}{2}.$$

It is convenient to introduce the scalar potentials ψ and ϕ as

$$\boldsymbol{h} = \boldsymbol{\nabla}\psi,$$
$$\boldsymbol{g} = \boldsymbol{\nabla}\phi.$$

Thus (7.33) and (7.34) can be reduced to

$$\begin{aligned} \frac{\partial \boldsymbol{m}}{\partial t} &= -\gamma(\boldsymbol{M}_s \times \boldsymbol{h} + \boldsymbol{m} \times \boldsymbol{H}_s + \boldsymbol{m} \times \boldsymbol{h}) \\ \nabla^2\psi + \boldsymbol{\nabla}\cdot\boldsymbol{m} &= 0 \qquad \text{(inside slab)} \\ \nabla^2\phi &= 0 \qquad \text{(outside slab).} \end{aligned} \tag{7.35}$$

Suppose we assume a quasi-monochromatic wave expansion for the potentials:

$$\phi = \varepsilon(\phi^{(1)}e^{i\theta} + \text{c.c.}) + \varepsilon^2(\phi^{(0)} + \phi^{(2)}e^{2i\theta} + \text{c.c.}) + \cdots$$

$$\psi = \varepsilon(\psi^{(1)}e^{i\theta} + \text{c.c.}) + \varepsilon^2(\psi^{(0)} + \psi^{(2)}e^{2i\theta} + \text{c.c.}) + \cdots$$

$$\boldsymbol{m} = \varepsilon\begin{pmatrix} m_1^{(1)} \\ m_2^{(1)} \\ 0 \end{pmatrix}(e^{i\theta} + \text{c.c.}) + \cdots .$$

where $\theta = kx - \omega t$ and $\phi^{(j)}$, $\psi^{(j)}$, $m^{(j)}$ are slowly varying functions of x, t and are independent of y; i.e., $\phi^{(j)} = \phi^{(j)}(X, T)$, $X = \varepsilon x$, $T = \varepsilon t$, etc.

The first of (7.35) show that $m_1^{(1)}$ and $m_2^{(1)}$ satisfy

$$-i\omega m_1^{(1)} = -\omega_H m_1^{(2)} - i\omega m_1^{(2)} = -ik\omega_M\psi^{(1)} + \omega_H m_1^{(1)}$$

where $\omega_M = \gamma M_s, \omega_H = \gamma H_s$. Hence

$$m_1^{(1)} = \chi_1 ik\psi^{(1)}, \qquad m_2^{(1)} = -\chi_2 k\psi^{(1)},$$

where

$$\chi_1 = \frac{\omega_H\omega_M}{\omega_H^2 - \omega^2}, \qquad \chi_2 = \frac{\omega_M\omega}{\omega_H^2 - \omega^2}.$$

To satisfy the latter two equations in (7.35) we find

$$\phi^{(1)} = A(X,T)e^{-k\left(|z|-\frac{d}{2}\right)}\cos(k_1 d/2)$$
$$\psi^{(1)} = A(X,T)\cos(k_1 z),$$

with

$$k_1^2 = k^2(1+\chi_1).$$

Using continuity of the vertical fields, $h_z + m_z = g_z$ or $\partial_z\psi = \partial_z\phi$ at $z = \pm d/2$, we find the dispersion relation

$$(1+\chi_1)^{1/2} = \cot(k_1 d/2),$$

which implicitly determines $\omega = \omega(k)$. Finally we remark that at higher order in the perturbation expansion, from a detailed calculation, an NLS equation for $A = A(X,T)$ is determined. For thin films ($d \ll 1$), the equation is found to be (see Zvezdin and Popkov, 1983)

$$i(A_T + v_g A_X) + \varepsilon\left(\frac{\omega''(k)}{2}A_{XX} + n|A|^2 A\right) = 0,$$

where n is a suitable function of k, γ, H_s, and M_s.

Exercises

7.1 (a) Derive the NLS equation (7.23) filling in the steps in the derivation.
(b) Obtain an NLS equation in electromagnetic "bulk" media when the slowly varying amplitude A in (7.15) is independent of T; here the longitudinal direction plays the role of the the temporal variable.
(c) Obtain the NLS equation in these (bulk) media when the slowly varying amplitude A in (7.15) is a function of X, Y, Z, T. What happens when A is independent of T?

7.2 (a) Derive the "NLSM system (7.29)–(7.30) by filling in the steps in the derivation.
(b) Derive time-independent systems similar to those discussed by Ablowitz et al. (2005) (see also Chapter 6) by assuming the amplitude is steady; i.e., has no time dependence.

7.3 Assume quasi-monchromatic waves in both transverse dimensions and obtain a vector NLSM system in quadratic ($\chi^{(2)} \neq 0$) nonlinear optical media. Hint: see Ablowitz et al. (1997, 2001a).

7.4 Use (7.35), and the assumed form of solution in that subsection, to obtain the linear dispersion relation of a wave $e^{ikx-\omega t)}$ propagating in a ferromagnetic strip of width d (like that depicted in Figure 7.1) but now with the external magnetic field aligned along the:

(a) x-axis;

(b) y-axis.