Chapter 11

Wave packets and the NLS equation

The transport of information through glass fibers by light is one of today's key technologies. Information is encoded digitally by ones and zeroes, in one approach by sending a light pulse through the optical fiber or not. From a physical point of view such a light pulse consists of an underlying electromagnetic carrier wave moving with phase velocity c_p and of a pulse-like envelope moving with group velocity c_g.

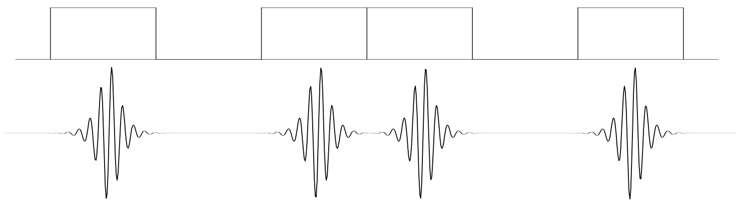

Figure 11.1. 0s and 1s are encoded physically by sending a light pulse or not; thus, for instance, the above series of electromagnetic waves encodes the sequence 101101.

The analysis of the evolution of the associated physical system is a non-trivial task. It shows linear and nonlinear dispersion and (weak) dissipation. As a result the pulses are smeared out which eventually causes an unwanted loss of information. A numerical investigation of the problem leads to a very large system due to the multiple scaling character of the problem. The wave length of the underlying carrier wave is around 10^{-7} m. Resolving such small oscillations in a fiber of 100 km $=10^5$ m gives in a one-dimensional uniform and not very accurate spatial discretization 10^{12} points even if we ignore the transverse directions and the temporal discretization. Hence, a direct simulation of Maxwell's equations which describe these electromagnetic waves is very expensive, if not impossible. Therefore, before making any numerical investigation, the system has to be analyzed and simpler, numerically more suitable, models have to be derived.

It turns out that the multiple scaling character of the problem is not only a curse, but also a blessing, since it allows to separate the dynamics of the envelope from the dynamics of the carrier wave, such that by multiple scaling analysis the NLS equation can be derived for the description of the slow modulations in time and space of the envelope of the spatially and temporarily oscillating wave packet. Due to the immense reduction of the dimension of the discretized problem by this procedure the NLS equation turned out to be a very successful model. Even though arguably its most important application is in nonlinear optics, e.g., [**Agr01**], the NLS equation has also been derived for water waves [**Zak68, Osb10**], for waves in DNA [**SH94b**] and other discrete chains, for Bose-Einstein condensates [**Pel11**], in plasma physics [**Deb05**, Chapter 10], and in many other fields as a universal envelope or modulation equation, cf. also [**Mil06**, Chapter 10]. In this chapter we explain its justification by approximation theorems for model problems. We explain its universal character, give an overview about approximation results, and explain a few applications.

11.1. Introduction

The Nonlinear Schrödinger (NLS) equation

$$\partial_T A = \mathrm{i}\nu_1 \partial_X^2 A + \mathrm{i}\nu_2 A\,|A|^2\,, \tag{11.1}$$

with $T \in \mathbb{R}$, $X \in \mathbb{R}$, $\nu_1, \nu_2 \in \mathbb{R}$, and $A(X,T) \in \mathbb{C}$ is a universal modulation equation which can be derived via multiple scaling analysis in order to describe slow modulations in time and space of the envelope of a spatially and temporarily oscillating wave packet. For instance, for the nonlinear wave equation

$$\partial_t^2 u = \partial_x^2 u - u - u^3, \qquad (x \in \mathbb{R},\, t \in \mathbb{R},\, u(x,t) \in \mathbb{R}), \tag{11.2}$$

also called the cubic Klein-Gordon equation, the ansatz for the derivation of the NLS equation is

$$\varepsilon\psi_{\mathrm{NLS}} = \varepsilon A\left(\varepsilon(x - c_g t), \varepsilon^2 t\right) \mathrm{e}^{\mathrm{i}(k_0 x + \omega_0 t)} + \mathrm{c.c.}, \tag{11.3}$$

where $0 < \varepsilon \ll 1$ is a small perturbation parameter, where c_g is the group velocity, and where the basic temporal and basic spatial wave number ω_0 and k_0 are related by the linear dispersion relation $\omega_0^2 = k_0^2 + 1$. We obtain that the envelope A of the underlying carrier wave $e^{i(k_0 x+\omega_0 t)}$ has to satisfy in lowest order the NLS equation

$$2i\omega_0 \partial_T A = (1-c_g^2)\partial_X^2 A - 3A\,|A|^2\,. \tag{11.4}$$

The dynamics of the NLS equation has been discussed in §8.1. The pulse solutions found in §8.1.1 correspond to modulating pulse solutions in the original system, cf. Figure 11.2.

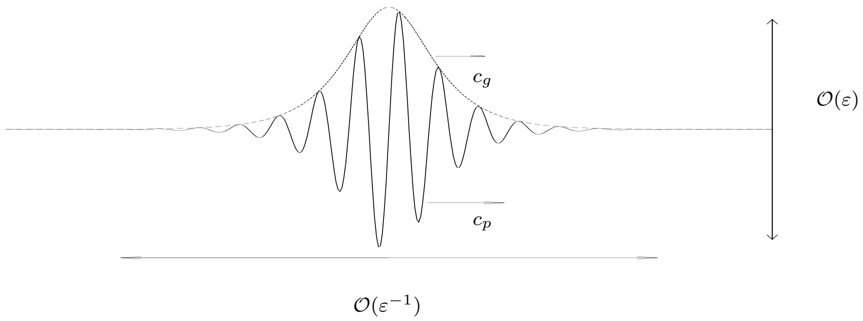

Figure 11.2. A modulating pulse described by the NLS equation. The envelope advancing with group velocity c_g in the laboratory frame modulates the underlying carrier wave $e^{i(k_0 x+\omega_0 t)}$ advancing with group velocity c_p. The envelope evolves approximately as a solution of the NLS equation.

Here we explain mathematical results which justify this formal approximation and show that the NLS equation makes correct predictions about the behavior of the solutions in the original system. In case of no quadratic terms in the original system the proof of error estimates turns out to be rather easy. The estimates follow by a simple application of Gronwall's inequality. This is the situation as it occurs in nonlinear optics due to symmetries. A complete proof of the estimates in this situation is given in §11.2. The proof of the approximation property in case of quadratic terms is presented in §11.4. In this case there are serious difficulties due to the fact that solutions of order $\mathcal{O}(\varepsilon)$ have to be bounded on the long $\mathcal{O}(1/\varepsilon^2)$ time scale w.r.t. t, which corresponds to an $\mathcal{O}(1)$ time scale w.r.t. $T = \varepsilon^2 t$. However, if a non-resonance condition is satisfied, then by averaging or normal form techniques the quadratic terms can be eliminated and this case can be brought back to the situation discussed in §11.2. In §11.5 we explain how the theory from the previous sections can be extended to the situation of additional resonances and to quasilinear systems. In §11.6 we transfer the theory to problems with spatially periodic coefficients and explain why standing light pulses can theoretically occur in photonic crystals. In this chapter we concentrate on nonlinear wave equations as original systems.

Nevertheless, the NLS equation is a universal modulation equation. This is explained in §11.3 and an in §11.7 the connection to nonlinear optics is given.

It turns out that the formal derivation of the NLS equation, its universality, and for cubic nonlinearities also its justification, has many similarities with the derivation, justification and universality of the GL equation for the simple model problems in §10.2 and §10.3. Nevertheless, we repeat most details to better be able to explain the crucial differences, which mainly lie in different suitable phase spaces, in the role of so called non-resonance conditions, and in the lack of attractivity properties of the set of modulated waves as in §10.8.

11.2. Justification in case of cubic nonlinearities

In this section we explain how to justify the NLS equation in case of cubic nonlinearities in the original system. For expository reasons we restrict ourselves to the cubic Klein-Gordon equation (11.2) as original system. Our main purpose is to prove that its solutions behave as predicted by the associated NLS equation (11.4). The NLS approximation (11.3) is formally a good approximation if the terms which do not cancel after inserting $\varepsilon\psi_{\mathrm{NLS}}$ into (11.2) are small. They are collected in the residual

$$\mathrm{Res}(u) = -\partial_t^2 u + \partial_x^2 u - u - u^3. \tag{11.5}$$

If $\mathrm{Res}(u) = 0$, then u is an exact solution of (11.2). With the abbreviation $\mathbf{E} = \mathrm{e}^{\mathrm{i}(k_0 x + \omega_0 t)}$ we find

$$\begin{aligned}
\mathrm{Res}(\varepsilon\psi_{\mathrm{NLS}}) \quad = \quad & \varepsilon\mathbf{E}\left((\omega_0^2 - k_0^2 - 1)A\right) \\
& +\varepsilon^2\mathbf{E}\left((2\mathrm{i}k_0 - 2\mathrm{i}c_g\omega_0)\partial_X A\right) \\
& +\varepsilon^3\mathbf{E}\left((-2\mathrm{i}\omega_0\partial_T A + (1-c_g^2)\partial_X^2 A - 3A\,|A|^2\right) \\
& +\varepsilon^3\mathbf{E}^3(-A^3) \\
& +\varepsilon^4\mathbf{E}(2c_g\partial_X\partial_T A) \\
& +\varepsilon^5\mathbf{E}(-\partial_T^2 A) + \text{c.c.}.
\end{aligned}$$

By choosing $\omega = \omega_0$ and $k = k_0$ to satisfy the linear dispersion relation

$$\omega^2 = k^2 + 1,$$

by choosing c_g to be the linear group velocity

$$c_g = \frac{\mathrm{d}}{\mathrm{d}k}\omega\Big|_{k=k_0,\omega=\omega_0} = \frac{k_0}{\omega_0},$$

and by choosing A to satisfy the NLS equation

$$2\mathrm{i}\omega_0\partial_T A = (1-c_g^2)\partial_X^2 A - 3A\,|A|^2\,, \tag{11.6}$$

the first three lines in the residual cancel. However, we still have $\mathrm{Res}(\varepsilon\psi_{\mathrm{NLS}}) = \mathcal{O}(\varepsilon^3)$.

Formal smallness of the residual. It turns out that by adding higher order terms to the approximation $\varepsilon\psi_{\mathrm{NLS}}$ the residual can be made arbitrarily small, i.e., for arbitrary, but fixed $n \in \mathbb{N}$ with $n \geq 3$ there exists an approximation $\varepsilon\psi_n$ with $\varepsilon\psi_n - \varepsilon\psi_{\mathrm{NLS}} = \mathcal{O}(\varepsilon^3)$ and $\mathrm{Res}(\varepsilon\psi_n) = \mathcal{O}(\varepsilon^n)$. Since $\varepsilon\psi_n - \varepsilon\psi_{\mathrm{NLS}} = \mathcal{O}(\varepsilon^3)$ the approximation $\varepsilon\psi_n$ makes the same predictions as $\varepsilon\psi_{\mathrm{NLS}}$ about the behavior of the solutions u of the original system. We will show $\mathrm{Res}(\varepsilon\psi_n) = \mathcal{O}(\varepsilon^n)$ for $n = 4, 5$. With these two examples the general situation can be understood.

In order to obtain

$$\mathrm{Res}(\varepsilon\psi_4) = \mathcal{O}(\varepsilon^4) \tag{11.7}$$

we define

$$\varepsilon\psi_4 = \varepsilon\psi_{\mathrm{NLS}} + \left(\varepsilon^3 A_3\left(\varepsilon(x - c_g t), \varepsilon^2 t\right)\mathbf{E}^3 + \text{c.c.}\right).$$

We find

$$\mathrm{Res}(\varepsilon\psi_4) = \varepsilon^3\mathbf{E}^3\left(-A^3 - (9\omega_0^2 - 9k_0^2 - 1)A_3\right) + \mathcal{O}(\varepsilon^4).$$

Due to the non-resonance $9\omega_0^2 - 9k_0^2 - 1 = 9(k_0^2 + 1) - 9k_0^2 - 1 = 8 \neq 0$ we can choose $A_3 = -(9\omega_0^2 - 9k_0^2 - 1)^{-1}A^3$ in order to achieve (11.7). In order to achieve

$$\mathrm{Res}(\varepsilon\psi_5) = \mathcal{O}(\varepsilon^5) \tag{11.8}$$

we define

$$\varepsilon\psi_5 = \varepsilon\psi_4 + \left(\varepsilon^2 A_{12}(\varepsilon(x - c_g t), \varepsilon^2 t)\mathbf{E} + \varepsilon^4 A_{32}(\varepsilon(x - c_g t), \varepsilon^2 t)\mathbf{E}^3 + \text{c.c.}\right)$$

where A_{12} and A_{32} are new functions to be chosen below. We find

$$\begin{aligned}\mathrm{Res}(\varepsilon\psi_5) =& \varepsilon^4\mathbf{E}\left(-2\mathrm{i}\omega_0\partial_T A_{12} + (1 - c_g^2)\partial_X^2 A_{12}\right)\\ &+ \varepsilon^4\mathbf{E}\left(-3A^2\overline{A_{12}} - 6\,|A|^2 A_{12} - 2c_g\partial_X\partial_T A\right)\\ &+ \varepsilon^4\mathbf{E}^3((9\omega_0^2 - 9k_0^2 - 1)A_{32} - 3A^2 A_{12}) + \mathcal{O}(\varepsilon^5) + \text{c.c.}.\end{aligned}$$

By choosing A_{12} to satisfy the linearized NLS equation

$$-2\mathrm{i}\omega_0\partial_T A_{12} + (1 - c_g^2)\partial_X^2 A_{12} - 3A^2\overline{A_{12}} - 6\,|A|^2 A_{12} - 2c_g\partial_X\partial_T A = 0 \tag{11.9}$$

and A_{32} to satisfy

$$(9\omega_0^2 - 9k_0^2 - 1)A_{32} - 3A^2 A_{12} = 0$$

we achieve (11.8). In order to achieve $\mathrm{Res}(\varepsilon\psi_n) = \mathcal{O}(\varepsilon^n)$ we choose

$$\varepsilon\psi_n = \sum_{m=-N,\ldots,N}\sum_{j=1}^{\widetilde{\alpha}(m)} \varepsilon^{\alpha(m)+(j-1)} A_{mj}(X,T)\mathbf{E}^m$$

with $N = n-1$, $X = \varepsilon(x - c_g t)$, $T = \varepsilon^2 t$, and $\alpha, \widetilde{\alpha}$ chosen according to

m	0	1	2	3	$\cdots$	m	$\cdots$	N
$\alpha(m)$	2	1	2	3	$\cdots$	$\lvert\lvert m\rvert - 1\rvert + 1$	$\cdots$	N
$\widetilde{\alpha}(m)$	$N-1$	$N-2$	$N-1$	$N-2$	$\cdots$	$N + 1 - \alpha(m) - 2\delta_{\lvert m\rvert 1}$	$\cdots$	1

The mode distribution of the NLS approximation is similar to one for the GL approximation which is sketched in Figure 10.6. As before A_{11} satisfies the NLS equation, the A_{1j} for $j \geq 2$ linearized NLS equations, and the A_{mj} for $m \neq \pm 1$ algebraic equations, which are linear in A_{mj} and can be solved w.r.t. the A_{mj} due to the validity of the non-resonance conditions

$$(m\omega_0)^2 - (mk_0)^2 - 1 = (m\omega(k_0))^2 - \omega(mk_0)^2 \neq 0, \tag{11.10}$$

for $m = -N, \ldots, N$, where $\omega(k) = \sqrt{1+k^2}$. In case of an odd nonlinearity such as for (11.2) we can set $A_{mj} = 0$ for m even.

Estimates for the residual. The formal orders of the residual can be improved to estimates in norms. We find for instance

$$\|\operatorname{Res}(\varepsilon\psi_{\mathrm{NLS}})\|_{C_b^0} \leq s_1 + s_2 + s_3$$

where

$$\begin{aligned} s_1 &= 2\|\varepsilon^3 E^3 A^3\|_{C_b^0} \leq 2\varepsilon^3 \, \|A\|^3_{C_b^0}\,, \\ s_2 &= 4\|\varepsilon^4 E c_g \partial_X \partial_T A\|_{C_b^0} \leq 4\varepsilon^4 c_g \|\partial_T A\|_{C_b^1}, \\ s_3 &= 2\|\varepsilon^5 E \partial_T^2 A\|_{C_b^0} \leq 2\varepsilon^5 \|\partial_T^2 A\|_{C_b^0}. \end{aligned}$$

We can use the right-hand side of the NLS equation to estimate $\|\partial_T A\|_{C_b^1}$ and $\|\partial_T^2 A\|_{C_b^0}$. For instance we have

$$\|\partial_T A\|_{C_b^1} \leq \frac{1}{2\omega_0}\left((1-c_g^2)\|\partial_X^2 A\|_{C_b^1} + 3\|A\|^3_{C_b^1}\right) < \infty$$

if $A \in C_b^3$. Similarly, we find $\|\partial_T^2 A\|_{C_b^0} < \infty$ if $A \in C_b^4$. This is completely analogous to the GL case in §10.2.

Lemma 11.2.1. *Let $A \in C([0,T_0], C_b^4)$ be a solution of the NLS equation and $\varepsilon_0 \in (0,1]$. There exists a $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\operatorname{Res}(\varepsilon\psi_{\mathrm{NLS}}(t))\|_{C_b^0} \leq C\varepsilon^3.$$

Similarly, for every $n \in \mathbb{N}$ with $n \geq 4$ there exists an approximation $\varepsilon\psi_n$ such that the following holds. Let $A \in C([0,T_0], C_b^{\theta_A})$ with $\theta_A = 3(n-3)+1$ be a solution of the NLS equation. Then for all $\varepsilon_0 \in (0,1]$ there exists a $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$:

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\operatorname{Res}(\varepsilon\psi_n(t))\|_{C_b^0} \leq C\varepsilon^n$$

and

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon\psi_{NLS}(t) - \varepsilon\psi_n(t)\|_{C_b^0} \leq C\varepsilon^3.$$

For (11.2) we have $\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon\psi_{NLS}(t)-\varepsilon\psi_\beta(t)\|_{C_b^0} \leq C\varepsilon^3$. If we look in more detail at the approximation $\varepsilon\psi_5$ we recognize that most regularity is lost in (11.9) for A_{12}. Since $\|\partial_T^2 A_{12}(\cdot,T)\|_{C_b^0}$ has to be estimated, we need $A_{12}(\cdot,T) \in C_b^4$, and so we need $\partial_X\partial_T A(\cdot,T) \in C_b^4$, respectively $A(\cdot,T) \in C_b^7$. Since the structure of the approximation equations is the same for the next orders we lose three derivatives in each step such that the estimates are possible with $\theta_A = 3(n-3)+1$. In fact θ_A can be chosen much smaller by a number of simple tricks, cf. §11.5.1.

The equations for the error. Estimates for the residual, even in norms, are only a necessary condition for showing that the NLS equation makes correct predictions about the behavior of the original systems. By no means they are sufficient. The errors can sum up in time and there are a number of counter-examples, cf. [**Sch95b, SSZ15**], showing that formally derived modulation equations make wrong predictions about the behavior of the original system.

The error $\varepsilon^\beta R = u - \varepsilon\psi$, the difference between the solution u and the approximation $\varepsilon\psi = \varepsilon\psi_n$, with β and n suitably chosen, is estimated with the help of Gronwall's inequality. It satisfies

$$\partial_t^2 R = \partial_x^2 R - R - 3\varepsilon^2\psi^2 R - 3\varepsilon^{\beta+1}\psi R^2 - \varepsilon^{2\beta}R^3 - \varepsilon^{-\beta}\operatorname{Res}(\varepsilon\psi). \tag{11.11}$$

Although there is local existence and uniqueness for (11.11) in C_b^m-spaces by the method of characteristics, here some crucial differences to Chapter 10 arise, since these spaces and this method are not suitable for obtaining estimates on the long time scale $\mathcal{O}(1/\varepsilon^2)$.

Estimates for the residual in Sobolev spaces. Sobolev spaces turn out to be more suitable for (11.11). Hence, we assume that

$A \in C([0,T_0],H^{\theta_A})$ *is a solution of the NLS equation with* $\theta_A \geq 0$ *sufficiently large.*

As a first step we have to re-estimate the residual in Sobolev spaces, taking into account the scaling properties of the L^2-norm. As in §10.2.2 we find

$$\begin{aligned}\|\operatorname{Res}(\varepsilon\psi_{\mathrm{NLS}})\|_{H^\theta} \leq & C\Big(\varepsilon^3\|A(\varepsilon\cdot)\|_{C_b^\theta}^2\|A(\varepsilon\cdot)\|_{H^\theta} \\ & \qquad + \varepsilon^4\|\partial_X\partial_T A(\varepsilon\cdot)\|_{H^\theta} + \varepsilon^5\|\partial_T^2 A(\varepsilon\cdot)\|_{H^\theta}\Big) \\ = & \mathcal{O}\left(\varepsilon^3\|A(\varepsilon\cdot)\|_{H^{\theta+4}}\right).\end{aligned}$$

However,

$$\|A(\varepsilon\cdot)\|_{L^2} = \left(\int_{\mathbb{R}} |A(\varepsilon x)|^2 \,\mathrm{d}x\right)^{1/2} = \varepsilon^{-1/2}\left(\int_{\mathbb{R}} |A(X)|^2 \,\mathrm{d}X\right)^{1/2} = \varepsilon^{-1/2}\|A\|_{L^2}$$

such that finally

$$\|\operatorname{Res}(\varepsilon\psi_{\mathrm{NLS}}(t))\|_{H^\theta} = \mathcal{O}\left(\varepsilon^{5/2}\|A\|_{H^{\theta+4}}\right).$$

It is essential that we estimate A^3 with $\|A(\varepsilon\cdot)\|^2_{C^\theta_b}\|A(\varepsilon\cdot)\|_{H^\theta}$ and not with $\|A(\varepsilon\cdot)\|^3_{H^\theta}$.

Nevertheless, ultimately this loss of $\varepsilon^{-1/2}$ is no problem, since as before, the residual can be made arbitrarily small by adding higher order terms to the approximation.

Lemma 11.2.2. *For all $n \in \mathbb{N}$ with $n \geq 4$ and $\theta \geq 1$ the following holds. Let $A \in C([0,T_0], H^{\theta_A})$ with $\theta_A = 3(n-3)+1+\theta$ be a solution of the NLS equation. Then for all $\varepsilon_0 \in (0,1]$ there exists a $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ there is an approximation $\varepsilon\psi_n$ with*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\operatorname{Res}(\varepsilon\psi_n(t))\|_{H^\theta} \leq C\varepsilon^{n-1/2}$$

and

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon\psi_{NLS}(t) - \varepsilon\psi_n(t)\|_{H^\theta} \leq C\varepsilon^{3/2}.$$

The equations for the error in Fourier space. For (11.2) simple energy estimates are possible, cf. Exercise 11.1. However, in order to have a method which also works for more general systems we use semigroup theory and write the equations for the error as first order system in Fourier space. We set $\beta = n - 5/2$, choose $\varepsilon\psi = \varepsilon\psi_n$ and find

$$\partial_t^2 \widehat{R} = -\omega^2\widehat{R} - 3\varepsilon^2\widehat{\psi}^{*2} * \widehat{R} - 3\varepsilon^{\beta+1}\widehat{\psi} * \widehat{R}^{*2} - \varepsilon^{2\beta}\widehat{R}^{*3} + \varepsilon^{-\beta}\widehat{\operatorname{Res}(\varepsilon\psi)},$$

where $\omega(k) = \sqrt{k^2+1}$. This is conveniently written as a first order system

$$\partial_t\widehat{R}_1 = \mathrm{i}\omega\widehat{R}_2,$$
$$\partial_t\widehat{R}_2 = \mathrm{i}\omega\widehat{R}_1 + \varepsilon^2\widehat{f},$$

where

$$\widehat{f} = \frac{1}{\mathrm{i}\omega}\left(-3\widehat{\psi}^{*2} * \widehat{R} - 3\varepsilon^{\beta-1}\widehat{\psi} * \widehat{R}^{*2} - \varepsilon^{2\beta-2}\widehat{R}^{*3} + \varepsilon^{-\beta-2}\widehat{\operatorname{Res}(\varepsilon\psi)}\right).$$

This system is abbreviated in the following as

$$\partial_t\widehat{\mathcal{R}}(k,t) = \Lambda(k)\widehat{\mathcal{R}}(k,t) + \varepsilon^2\widehat{F}(k,t),$$

with

$$\Lambda(k) = \begin{pmatrix} 0 & \mathrm{i}\omega(k) \\ \mathrm{i}\omega(k) & 0 \end{pmatrix}, \qquad \widehat{F}(k,t) = \begin{pmatrix} 0 \\ \widehat{f}(k,t) \end{pmatrix}.$$

We use the variation of constant formula

$$\widehat{\mathcal{R}}(k,t) = \mathrm{e}^{t\Lambda(k)}\widehat{\mathcal{R}}(k,0) + \varepsilon^2 \int_0^t \mathrm{e}^{(t-\tau)\Lambda(k)} \widehat{F}(k,\tau)\,\mathrm{d}\tau$$

in order to estimate the solutions of this system.

Lemma 11.2.3. *The semigroup* $(\mathrm{e}^{t\Lambda(k)})_{t\geq 0}$ *is uniformly bounded in every* H^0_θ, *cf. Definition 7.3.30 and Lemma 7.3.31, i.e., there exists a* $C>0$ *such that we have* $\sup_{t\in\mathbb{R}} \|\mathrm{e}^{t\Lambda}\|_{H^0_\theta \to H^0_\theta} \leq C$.

Proof. We have $\Lambda(k) = SD(k)S^{-1}$ where

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad \text{and} \qquad D(k) = \begin{pmatrix} \mathrm{i}\omega(k) & 0 \\ 0 & -\mathrm{i}\omega(k) \end{pmatrix}$$

such that $\mathrm{e}^{t\Lambda(k)} = S\mathrm{e}^{tD(k)}S^{-1}$. Hence,

$$\|\mathrm{e}^{t\Lambda}\widehat{u}\|_{H^0_\theta} \leq \sup_{k\in\mathbb{R}} \|\mathrm{e}^{t\Lambda(k)}\|_{\mathbb{C}^2\to\mathbb{C}^2} \|\widehat{u}\|_{H^0_\theta}$$

and further

$$\begin{aligned} \sup_{k\in\mathbb{R}} \|\mathrm{e}^{t\Lambda(k)}\|_{\mathbb{C}^2\to\mathbb{C}^2} &\leq \|S\|_{\mathbb{C}^2\to\mathbb{C}^2} \sup_{k\in\mathbb{R}} \|\mathrm{e}^{tD(k)}\|_{\mathbb{C}^2\to\mathbb{C}^2} \cdot \|S^{-1}\|_{\mathbb{C}^2\to\mathbb{C}^2} \\ &\leq \|S\|_{\mathbb{C}^2\to\mathbb{C}^2} \|S^{-1}\|_{\mathbb{C}^2\to\mathbb{C}^2} < \infty. \end{aligned}$$

□

Lemma 11.2.4. *For every* $\theta \geq 1$ *there is a* $C>0$ *such that for all* $\varepsilon \in (0,1]$ *we have*

$$\|\widehat{F}\|_{H^0_\theta} \leq C\left(\|\widehat{\mathcal{R}}\|_{H^0_\theta} + \varepsilon^{\beta-1}\|\widehat{\mathcal{R}}\|^2_{H^0_\theta} + \varepsilon^{2\beta-2}\|\widehat{\mathcal{R}}\|^3_{H^0_\theta} + 1 \right).$$

Proof. The estimate follows from

$$\begin{gathered} \Big\|\frac{1}{\omega}\widehat{u}\Big\|_{H^0_\theta} \leq \|\widehat{u}\|_{H^0_\theta}, \quad \|\varepsilon^{-\beta}\widehat{\mathrm{Res}(\varepsilon\psi)}\|_{H^0_\theta} \leq C, \quad \|\widehat{R}^{*3}\|_{H^0_\theta} \leq C\|\widehat{R}\|^3_{H^0_\theta}, \\ \Big\|\widehat{\psi} * \widehat{\psi} * \widehat{R}\Big\|_{H^0_\theta} \leq C\|\widehat{\psi}\|^2_{L^1_\theta}\|\widehat{R}\|_{H^0_\theta}, \quad \|\widehat{\psi} * \widehat{R} * \widehat{R}\|_{H^0_\theta} \leq C\|\widehat{\psi}\|_{L^1_\theta}\|\widehat{R}\|^2_{H^0_\theta}, \end{gathered}$$

and

$$\begin{aligned} \|\widehat{\psi}\|_{L^1_\theta} &= 2\Big\|\frac{1}{\varepsilon}A\Big(\frac{\cdot - k_0}{\varepsilon}\Big) + \text{h.o.t.}\Big\|_{L^1_\theta} \leq C\Big\|\frac{1}{\varepsilon}A\Big(\frac{\cdot}{\varepsilon}\Big)\Big\|_{L^1_\theta} + \text{h.o.t.} \\ &\leq C\|A\|_{L^1_\theta} + \text{h.o.t.} \leq C\|A\|_{H^0_{\theta+1}} + \text{h.o.t.} \end{aligned}$$

□

Remark 11.2.5. Note that $\|\widehat{\psi}\|_{H^0_\theta} = \mathcal{O}(\varepsilon^{-1/2})$ such that $\widehat{\psi}$ has to be estimated in the L^1_θ-norm, respectively ψ in the C^θ_b-norm in order to get ε^2 for the most dangerous term $\varepsilon^2\widehat{\psi} * \widehat{\psi} * \widehat{R}$. This power is necessary to obtain estimates on the natural time scale $\mathcal{O}(1/\varepsilon^2)$ w.r.t. t. ⌋

Using the previous lemmas shows that

$$\begin{aligned}\|\widehat{\mathcal{R}}(t)\|_{H^0_\theta} &\leq C\varepsilon^2 \int_0^t \left(\|\widehat{\mathcal{R}}(\tau)\|_{H^0_\theta} + \varepsilon^{\beta-1}\|\widehat{\mathcal{R}}(\tau)\|^2_{H^0_\theta} + \varepsilon^{2\beta-2}\|\widehat{\mathcal{R}}(\tau)\|^3_{H^0_\theta} + 1\right) \mathrm{d}\tau \\ &\leq C\varepsilon^2 \int_0^t \left(\|\widehat{\mathcal{R}}(\tau)\|_{H^0_\theta} + 2\right) \mathrm{d}\tau \leq 2CT_0 + C\varepsilon^2 \int_0^t \|\widehat{\mathcal{R}}(\tau)\|_{H^0_\theta} \,\mathrm{d}\tau,\end{aligned}$$

which holds as long as

$$\varepsilon^{\beta-1}\|\widehat{\mathcal{R}}(\tau)\|^2_{H^0_\theta} + \varepsilon^{2\beta-2}\|\widehat{\mathcal{R}}(\tau)\|^3_{H^0_\theta} \leq 1. \tag{11.12}$$

Gronwall's inequality then yields

$$\|\widehat{\mathcal{R}}(t)\|_{H^0_\theta} \leq 2CT_0 \mathrm{e}^{C\varepsilon^2 t} \leq 2CT_0 \mathrm{e}^{CT_0} = M$$

for all $t \in [0, T_0/\varepsilon^2]$. Choosing $\varepsilon_0 > 0$ such that $\varepsilon_0^{\beta-1}M^2 + \varepsilon_0^{2\beta-2}M^3 \leq 1$ we have satisfied the condition (11.12) and so proved the following approximation result.

Theorem 11.2.6. *For all $n \in \mathbb{N}$ with $n \geq 4$ and $\theta \geq 1$ the following holds: Let $A \in C([0, T_0], H^{\theta_A})$ with $\theta_A = 3(n-3) + 1 + \theta$ be a solution of the NLS equation* (11.6). *Then there exists an $\varepsilon_0 > 0$ and a $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there are solutions u of the original system* (11.2) *which can be approximated by $\varepsilon\psi_n$ with*

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|u(t) - \varepsilon\psi_n(t)\|_{H^\theta} < C\varepsilon^{n-5/2}$$

and as a consequence

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|u(t) - \varepsilon\psi_{NLS}(t)\|_{H^\theta} < C\varepsilon^{3/2}.$$

Remark 11.2.7. We have local existence and uniqueness of the solutions of the nonlinear wave equation (11.2), respectively the equations for the error (11.11), in the spaces where we proved the error estimates: Fix $\theta \geq 1$ and let $(u_0, u_1) \in H^{\theta+1} \times H^\theta$. Then there exists a $t_0 > 0$ such that (11.2) possesses a unique solution $u \in C([-t_0, t_0], H^{\theta+1})$ with $u|_{t=0} = u_0$ and $\partial_t u|_{t=0} = u_1$.

In order to construct solutions of (11.2) we use the formula

$$u(x,t) = \frac{1}{2}(u_0(x+t) + u_0(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} u_1(\xi)\,\mathrm{d}\xi + \int_0^t \int_{x-(t-s)}^{x+(t-s)} f(y,s)\,\mathrm{d}y\,\mathrm{d}s$$

with $f(x,t) = -u(x,t) - u(x,t)^3$ which is based on the solution formula for the inhomogeneous wave equation. For $t_0 > 0$ sufficiently small the right-hand side $F(u)$ is a contraction in the space $C([-t_0, t_0], H^{\theta+1})$. Thus, there exists a unique fixed point $u^* = F(u^*)$ which is a classical solution of (11.2) if $m \geq 2$.

The solutions exist as long as the norm of the solutions stay bounded. By using the error estimates as a priori estimates we can guarantee that the solutions stay bounded for $t \in [0, T_0/\varepsilon^2]$ and so we can apply the local existence and uniqueness again and again to guarantee the existence and uniqueness of the solutions of the error equations which are obtained from (11.2) by a smooth change of variables. $\rfloor$

11.3. The universality of the NLS equation

As already said, the NLS approximation can be derived in various systems. In order to explain why this is the case, and why the NLS equation plays such an important role, we review the derivation of the NLS equation from (11.2) from a different point of view. This derivation will explain why the NLS equation occurs as a universal modulation equation describing the evolution of modulated wave packets. The underlying system will condense in the values of the coefficients ν_1 and ν_2 in (11.1).

The Fourier transformed system. As for the GL approximation in §10.3 it turns out that Fourier transform is the key for the understanding of the universality. Hence, we consider (11.2) in Fourier space. The Fourier transform $\widehat{u}$ satisfies

$$\partial_t^2 \widehat{u}(k,t) = -\omega^2(k)\widehat{u}(k,t) - \widehat{u}^{*3}(k,t), \tag{11.13}$$

where $\omega(k) = \sqrt{k^2+1}$. By introducing $\widehat{w}(k) = (\widehat{u}(k), \frac{1}{\mathrm{i}\omega(k)}\partial_t \widehat{u}(k))$ we rewrite (11.13) into the first order system

$$\partial_t \widehat{w}(k,t) = \widehat{M}(k)\widehat{w}(k,t) + \widehat{N}(\widehat{w})(k,t), \tag{11.14}$$

where

$$\widehat{M}(k) = \begin{pmatrix} 0 & \mathrm{i}\omega(k) \\ \mathrm{i}\omega(k) & 0 \end{pmatrix}, \quad \widehat{N}(\widehat{w})(k,t) = \begin{pmatrix} 0 \\ \frac{-1}{\mathrm{i}\omega(k)}\widehat{u}^{*3}(k,t) \end{pmatrix}.$$

This system is diagonalized for fixed wave number k. For (11.14) the associated transformation $\widehat{S} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is independent of k and unitary, i.e., $\widehat{S}^{-1} = \widehat{S}^*$. The transformed variable $\widehat{z} = \widehat{S}^*\widehat{w}$ satisfies the diagonalized system

$$\partial_t \widehat{z} = \widehat{\Lambda}\widehat{z} + \widehat{S}^*\widehat{N}(\widehat{S}\widehat{z}), \tag{11.15}$$

with $\widehat{\Lambda}(k) = \mathrm{diag}(i\omega(k), -i\omega(k))$. It turns out that the NLS equation can be derived whenever the original system can be transformed in a system of this form.

General dispersive wave systems. The nonlinear wave equation (11.2) is an example of a dispersive wave system

$$\partial_t U = LU + N_2(U,U) + N_3(U,U,U) + \dots$$

for an unknown function $U = U(x,t)$ with values in $\mathbb{C}^N$ and $x, t \in \mathbb{R}$. The operator L is linear and skew symmetric, and the terms N_j are j-multilinear and w.l.o.g. symmetric in their arguments, and do not depend explicitly on $x \in \mathbb{R}$. Due to this translation invariance w.r.t. $x \in \mathbb{R}$ the linearization $\partial_t U = LU$ around $U^* = 0$ possesses solutions of the form

$$U(x,t) = f_n(k) e^{ikx} e^{i\omega_n(k)t},$$

with $k \in \mathbb{R}$ and $n \in I$, where I is some finite index set and $f_n(k) \in \mathbb{C}^N$. The case of $U(\cdot,t) : \Omega \times \mathbb{R} \to \mathbb{R}^N$, where $\Omega = \mathbb{R} \times \Sigma$, $\Sigma \subset \mathbb{R}^m$ a bounded cross section with $m \geq 1$ leads to a countable index set $I = \mathbb{N}$ or $I = \mathbb{Z}$ and to $f_n(k) \in L^2(\Sigma, \mathbb{C}^N)$. We sort the eigenvalues for fixed k as $\omega_n \leq \omega_{n+1}$. For real-valued systems we assume $\omega_n = -\omega_{-n}$. After an expansion

$$U(x,t) = \int_{-\infty}^{\infty} \sum_{n \in I} c_n(k,t) f_n(k) e^{ikx} \, dk$$

in eigenfunctions, the coefficients $c_n(k,t)$ satisfy

$$\begin{aligned}
&\partial_t c_n(k,t) \\
&= i\omega_n(k) c_n(k,t) + \sum_{n_1,n_2 \in I} \int_{-\infty}^{\infty} s_{2,n,n_1,n_2}(k, k-l, l) c_{n_1}(k-l) c_{n_2}(l) \, dl \\
&\quad + \sum_{n_1,n_2 \in I} \int_{-\infty}^{\infty} s_{3,n,n_1,n_2,n_3}(k, k-l, l-m, m) c_{n_1}(k-l) c_{n_2}(l-m) c_{n_3}(m) \, dl \, dm + \ldots .
\end{aligned}$$

with complex-valued kernels $s_{2,n,n_1,n_2}(k, k-l, l)$, $s_{3,n,n_1,n_2,n_3}(k, k-l, l-m, m)$, etc. We have for instance

$$s_{2,n,n_1,n_2}(k, k-l, l) = \langle f_n^*(k), e^{-ikx} N_2[f_{n_1}(k-l) e^{i(k-l)x}, f_{n_2}(l) e^{ilx}] \rangle_{L^2}$$

where $f_n^*(k)$ is the associated adjoint eigenfunction w.r.t. the scalar product $\langle \cdot, \cdot \rangle_{L^2}$.

Derivation of the NLS equation for (11.15). Taking the Fourier transform of the ansatz in physical space leads to the ansatz

$$\begin{aligned}
\widehat{z}(k,t) &= \varepsilon\varepsilon^{-1} \widehat{A}_1 \left(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t \right) e^{i\omega(k_0)t} e^{ic_g(k-k_0)t} \vec{e}_1 \\
&\quad + \varepsilon\varepsilon^{-1} \widehat{A}_{-1} \left(\frac{k+k_0}{\varepsilon}, \varepsilon^2 t \right) e^{-i\omega(k_0)t} e^{ic_g(k+k_0)t} \vec{e}_2
\end{aligned} \tag{11.16}$$

for (11.15), where

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The notation $\varepsilon\varepsilon^{-1}$ refers to the amplitude scaling ε and the wave number scaling ε^{-1}, but in the following we shorten this to $\varepsilon\varepsilon^{-1} = 1$. Since the Fourier modes of the wave packet are concentrated in an $\mathcal{O}(\varepsilon)$ neighborhood of the basic wave numbers $\pm k_0$ the evolution of the wave packet will be strongly determined by the curves $\pm\omega$ at $\pm k_0$. At $e^{i\omega(k_0)t}e^{ic_g(k-k_0)t}\vec{e}_1$ we find

$$\begin{aligned}
&i\omega(k_0)\widehat{A}_1 + i\varepsilon c_g K\widehat{A}_1 + \varepsilon^2\partial_T\widehat{A}_1 \\
&\quad = i\omega(k_0)\widehat{A}_1 + i\varepsilon\partial_k\omega(k_0)K\widehat{A}_1 + \frac{i}{2}\varepsilon^2\partial_k^2\omega(k_0)K^2\widehat{A}_1 \\
&\qquad + \varepsilon^2\frac{3i}{4\omega(k_0)}\widehat{A}_1 * \widehat{A}_1 * \widehat{A}_{-1} + \mathcal{O}(\varepsilon^3),
\end{aligned}$$

where $k = k_0 + \varepsilon K$, $\widehat{A}_1 = \widehat{A}_1(K,T)$, and where we used

$$\begin{aligned}
&\int_{\mathbb{R}} \widehat{A}\left(\frac{k-l-k_0}{\varepsilon}, T\right) e^{i\omega(k_0)t}e^{ic_g(k-l-k_0)t}\widehat{A}\left(\frac{l-k_0}{\varepsilon}, T\right) e^{i\omega(k_0)t}e^{ic_g(l-k_0)t}\,dl \\
&\qquad = \varepsilon\int_{\mathbb{R}} \widehat{A}\left(\frac{k-2k_0}{\varepsilon} - m, T\right)\widehat{A}(m,T)e^{2i\omega(k_0)t}e^{ic_g(k-2k_0)t}\,dm\,.
\end{aligned}$$

At ε^0 and ε^1 we obtain the linear dispersion relation and the linear group velocity. At ε^2 we obtain a NLS equation. Undoing the transformation $w = Sz$ and $u = w_1$ gives a multiple of the original approximation

$$\begin{aligned}
\widehat{u}(k,t) = &\frac{1}{\sqrt{2}}\left(\widehat{A}_1\left(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t\right) e^{i\omega(k_0)t}e^{ic_g(k-k_0)t}\right. \\
&\left. + \widehat{A}_{-1}\left(\frac{k+k_0}{\varepsilon}, \varepsilon^2 t\right) e^{-i\omega(k_0)t}e^{ic_g(k+k_0)t}\right).
\end{aligned}$$

Derivation of the NLS equation in the general situation. By this procedure it is clear that the NLS equation occurs as a modulation equation for dispersive wave systems whenever the Fourier transform of the initial condition is strongly concentrated at a wave number $k_0 \neq 0$ and when the concentration and the amplitude are of the correct order. It also occurs in case $k_0 = 0$ if $\omega(0) \neq 0$, cf. Exercise 11.1.

W.l.o.g. let us assume that we derive the NLS equation for the curve ω_1 and that we have a real-valued system. We make the ansatz

$$\begin{aligned}
c_1(k,t) = &\widehat{A}_{1,1}(\varepsilon^{-1}(k-k_0), \varepsilon^2 t)e^{i\omega(k_0)t}e^{ic_g(k-k_0)t} \\
&+ \varepsilon\widehat{A}_{1,-2}(\varepsilon^{-1}(k+2k_0), \varepsilon^2 t)e^{-2i\omega(k_0)t}e^{ic_g(k+2k_0)t} \\
&+ \varepsilon\widehat{A}_{1,0}(\varepsilon^{-1}k, \varepsilon^2 t)e^{ic_g kt} \\
&+ \varepsilon\widehat{A}_{1,2}(\varepsilon^{-1}(k-2k_0), \varepsilon^2 t)e^{2i\omega(k_0)t}e^{ic_g(k-2k_0)t}, \\
c_{-1}(k,t) = &\widehat{A}_{-1,-1}(\varepsilon^{-1}(k+k_0), \varepsilon^2 t)e^{-i\omega(k_0)t}e^{ic_g(k+k_0)t}) \\
&+ \varepsilon\widehat{A}_{-1,-2}(\varepsilon^{-1}(k+2k_0), \varepsilon^2 t)e^{-2i\omega(k_0)t}e^{ic_g(k+2k_0)t} \\
&+ \varepsilon\widehat{A}_{-1,0}(\varepsilon^{-1}k, \varepsilon^2 t)e^{ic_g kt}
\end{aligned}$$

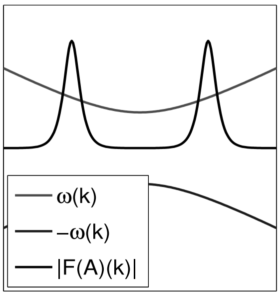

Figure 11.3. The derivation of the NLS equation is based on the concentration of the Fourier modes at a certain wave number. The figure shows the curves of eigenvalues and the concentration of the Fourier modes. Hence, for the evolution of these modes only the curves of eigenvalues close to these wave numbers play a role.

$$
\begin{aligned}
&+\varepsilon\widehat{A}_{-1,2}(\varepsilon^{-1}(k-2k_0),\varepsilon^2 t)\mathrm{e}^{2\mathrm{i}\omega(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k-2k_0)t},\\
c_n(k,t)=&\varepsilon\widehat{A}_{n,-2}(\varepsilon^{-1}(k+2k_0),\varepsilon^2 t)\mathrm{e}^{-2\mathrm{i}\omega(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k+2k_0)t}\\
&+\varepsilon\widehat{A}_{n,0}(\varepsilon^{-1}k,\varepsilon^2 t)\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}kt}\\
&+\varepsilon\widehat{A}_{n,2}(\varepsilon^{-1}(k-2k_0),\varepsilon^2 t)\mathrm{e}^{2\mathrm{i}\omega(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k-2k_0)t},
\end{aligned}
$$

for $|n|\geq 2$ and complex valued functions $\widehat{A}_{n,j}$, where $A_{n,-j}=\overline{A_{n,j}}$ in physical space. With $k-k_0=\varepsilon K$ we find at ε^2 for the modes concentrated at k_0 that

$$
\begin{aligned}
&\partial_T\widehat{A}_{1,1}(K,T)\\
&=i\partial_k^2\omega_1(k_0)K^2\widehat{A}_{1,1}(K,T)/2\\
&\quad+2\sum_{n\in\mathbb{Z}}s_{2,1,1,n}(k_0,k_0,0)\int_{-\infty}^{\infty}\widehat{A}_{1,1}(K-\kappa,T)\widehat{A}_{n,0}(\kappa,T)\,\mathrm{d}\kappa\\
&\quad+2\sum_{n\in\mathbb{Z}}s_{2,1,1,n}(k_0,-k_0,2k_0)\int_{-\infty}^{\infty}\widehat{A}_{1,-1}(K-\kappa,T)\widehat{A}_{n,2}(\kappa,T)\,\mathrm{d}\kappa\\
&\quad+3s_{3,1,1,1,1}(k_0,k_0,k_0,-k_0)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\widehat{A}_{1,1}(K-\kappa_1,T)\\
&\qquad\qquad\times\widehat{A}_{1,1}(\kappa_1-\kappa_2,T)\widehat{A}_{1,-1}(\kappa_2,T)\,\mathrm{d}\kappa_2\,\mathrm{d}\kappa_1,
\end{aligned}
$$

where we used the symmetry of the multilinear terms in their arguments, cf. (11.21). With $k-jk_0=\varepsilon K$ we find at ε in the n-th equation for the modes concentrated at jk_c for $j=0,2$ that

$$
0=\mathrm{i}\omega_n(0)A_{n,0}(K,T)
$$

$$
\begin{aligned}
&+ 2s_{2n1-1}(0,-k_c,k_c)\int_{-\infty}^{\infty}\widehat{A}_{1,-1}(K-\kappa,T)\widehat{A}_{1,1}(\kappa,T)\,\mathrm{d}\kappa,\\
0 =& \mathrm{i}(\omega_n(2k_0)-2\omega_1(k_0))A_{n,2}(K,T)\\
&+ s_{2n11}(2k_c,k_c,k_c)\int_{-\infty}^{\infty}\widehat{A}_{1,1}(K-\kappa,T)\widehat{A}_{1,1}(\kappa,T)\,\mathrm{d}\kappa,
\end{aligned}
$$

see (11.22)-(11.23) for an example. Under the non-resonance conditions

$$
\omega_n(0)\neq 0, \qquad \omega_n(2k_0)-2\omega_1(k_0)\neq 0,
$$

these two algebraic relations determine $A_{n,0}$ and $A_{n,2}$ in terms of $A_{1,1}$ and $A_{1,-1}$ such that $A_{1,1}$ finally satisfies the NLS equation in Fourier space, namely

$$
\begin{aligned}
\partial_T\widehat{A}_{1,1}(K,T) =& -\mathrm{i}\nu_1K^2\widehat{A}_{1,1}(K,T)\\
&+ \mathrm{i}\nu_2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\widehat{A}_{1,1}(K-\kappa,T)\widehat{A}_{1,1}(\kappa-\widetilde{\kappa},T)\widehat{A}_{1,-1}(\widetilde{\kappa},T)\,\mathrm{d}\widetilde{\kappa}\,\mathrm{d}\kappa,
\end{aligned}
$$

with coefficients

$$
\begin{aligned}
\nu_1 &= -\partial_k^2\omega_1(k_0)/2,\\
\nu_2 &= 4\sum_{n\in\mathbb{Z}}s_{2,1,1,n}(k_0,k_0,0)s_{2,n,1,-1}(0,-k_c,k_c)/\omega_n(0)\\
&\quad +2\sum_{n\in\mathbb{Z}}s_{2,1,1,n}(k_0,-k_0,2k_0)s_{2,n,1,1}(2k_0,k_0,k_0)/\omega_n(2k_0)\\
&\quad -3s_{3,1,1,1,1}(k_0,k_0,k_0,-k_0).
\end{aligned}
$$

Therefore, the NLS equation is the universal modulation equation describing slow modulations in time and space of a propagating wave packet in dispersive systems

$$
\partial_t\widehat{c}_j(k,t) = \mathrm{i}\omega_j(k)\widehat{c}_j(k,t) + \text{ nonlinear terms,} \tag{11.17}
$$

where j is in some index set and where the nonlinear terms have some convolution structure.

Remark 11.3.1. We have seen that for (11.2) the valdity of the non-resonance condition (11.10) is necessary for the construction of higher order approximations. For general dispersive wave systems (11.17) for the validity of a counterpart to Lemma 11.2.1 the validity of the non-resonance condition

$$
(m\omega_1(k_0))^2-\omega_j(mk_0)^2\neq 0, \tag{11.18}
$$

is necessary for $m=-N,\dots,N$, where $N=n-1$, and j in some index set. ⌋

11.4. Quadratic Nonlinearities

In this section we explain how to justify the NLS approximation in case of quadratic nonlinearities. For expository reasons we restrict ourselves to

$$\partial_t^2 u = \partial_x^2 u - u + u^2, \tag{11.19}$$

with $x \in \mathbb{R}$, $t \in \mathbb{R}$, and $u(x,t) \in \mathbb{R}$ as original system. The ansatz for the derivation of the NLS equation is then given by

$$\begin{aligned} \varepsilon\psi_{\mathrm{NLS}} =& \varepsilon A_1\left(\varepsilon(x-c_g t), \varepsilon^2 t\right) \mathrm{e}^{\mathrm{i}(k_0 x + \omega_0 t)} + \text{c.c.} \\ &+ \varepsilon^2 A_2\left(\varepsilon(x-c_g t), \varepsilon^2 t\right) \mathrm{e}^{2\mathrm{i}(k_0 x + \omega_0 t)} + \text{c.c.} \\ &+ \varepsilon^2 A_0\left(\varepsilon(x-c_g t), \varepsilon^2 t\right). \end{aligned} \tag{11.20}$$

We find as before at εE the linear dispersion relation and at $\varepsilon^2 E$ the condition for the linear group velocity c_g. At $\varepsilon^3 E$ we find

$$2\mathrm{i}\omega_0 \partial_T A_1 = (1-c_g^2)\partial_X^2 A_1 + 2A_1 A_0 + 2A_2 A_{-1}. \tag{11.21}$$

The algebraic relations which are found at $\varepsilon^2 E^0$ and $\varepsilon^2 E^2$

$$\varepsilon^2 E^0: \quad 0 = -A_0 + 2A_1 A_{-1}, \tag{11.22}$$

$$\varepsilon^2 E^2: \quad 0 = -(-4\omega_0^2 + 4k_0^2 + 1)A_2 + A_1^2, \tag{11.23}$$

can be solved w.r.t. A_0 and A_2 since

$$-4\omega_0^2 + 4k_0^2 + 1 = -(2\omega(k_0))^2 + \omega(2k_0)^2 \neq 0.$$

Inserting the solution for A_0 and A_2 into the equation for A_1 finally yields the NLS equation

$$2\mathrm{i}\omega_0 \partial_T A_1 = (1-c_g^2)\partial_X^2 A_1 + \gamma A_1 |A_1|^2, \tag{11.24}$$

with

$$\gamma = 4 + \frac{2}{-4\omega_0^2 + 4k_0^2 + 1}.$$

Like in case of cubic nonlinearities the residual

$$\mathrm{Res}(u) = -\partial_t^2 u + \partial_x^2 u - u + u^2$$

can be made arbitrarily small by adding higher order terms, i.e., we have

Lemma 11.4.1. *For all $n \in \mathbb{N}$ with $n \geq 4$ and $\theta \geq 1$ the following holds. For $A \in C([0,T_0], H^{\theta_A})$ with $\theta_A = 3(n-3)+1+\theta$ and $\varepsilon_0 \in (0,1]$ there exists a $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ there is an approximation $\varepsilon\psi_n$ with*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\mathrm{Res}(\varepsilon\psi_n(t))\|_{H^\theta} \leq C\varepsilon^{n-1/2}$$

and

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\varepsilon\psi_{\mathrm{NLS}}(t) - \varepsilon\psi_n(t)\|_{C_b^0} \leq C\varepsilon^2.$$

In order to prove that the solution A_1 of the NLS equation (11.24) predicts the behavior of the solutions u of the original system correctly we estimate as before the difference $\varepsilon^\beta R = u - \varepsilon\psi$ between the correct solution u and its approximation, and as before we choose $\varepsilon\psi = \varepsilon\psi_n$ with $\beta = n - 5/2$. This difference satisfies

$$\partial_t^2 R = \partial_x^2 R - R + 2\varepsilon\psi R + \varepsilon^\beta R^2 + \varepsilon^{-\beta}\,\mathrm{Res}(\varepsilon\psi).$$

As a first order system in Fourier space this yields

$$\begin{aligned} \partial_t \widehat{R}_1 &= \mathrm{i}\omega \widehat{R}_2, \\ \partial_t \widehat{R}_2 &= \mathrm{i}\omega \widehat{R}_1 + \tfrac{1}{\mathrm{i}\omega}\left(2\varepsilon\widehat{\psi} * \widehat{R}_1 + \varepsilon^\beta \widehat{R}_1^{*2} + \varepsilon^{-\beta}\widehat{\mathrm{Res}(\varepsilon\psi)}\right), \end{aligned} \tag{11.25}$$

where again $\omega(k) = \sqrt{k^2+1}$. The simple argument of the last section no longer works because of the new $\mathcal{O}(\varepsilon)$-term $2\varepsilon\widehat{\psi} * \widehat{R}_1$. In principle this term can give some exponential growth of order $\mathcal{O}(\exp(\varepsilon t))$ which is not $\mathcal{O}(1)$-bounded on the time scale of order $\mathcal{O}(1/\varepsilon^2)$. However, this term is oscillatory in time and can be eliminated by averaging or a normal form transformation such that it finally has an $\mathcal{O}(1)$-influence on the size of the solutions. This observation goes back to [**Kal88**].

Remark 11.4.2. Normal form transformations have already been considered in Part I of this book. For PDEs the idea is very similar. For the abstract evolutionary system

$$\partial_t u = Au + N_Q(u) + N_c(u),$$

where N_Q stands for quadratic and N_c for the higher order terms, we seek a near identity change of coordinates $v = u - K(u)$ to eliminate the quadratic terms $N_Q(u)$ and to transfer them into higher order terms. We find

$$\begin{aligned} \partial_t v =& \partial_t u - K'(u)\partial_t u \\ =& Au + N_Q(u) + N_c(u) - K'(u)Au - K'(u)N_Q(u) - K'(u)N_c(u) \\ =& Av + AK(u) - K'(u)Au + N_Q(u) + N_c(u) - K'(u)N_Q(u) - K'(u)N_c(u). \end{aligned}$$

In order to eliminate the quadratic terms we choose K to satisfy

$$AK(u) - K'(u)Au + N_Q(u) = 0$$

such that after the transformation

$$\partial_t v = Av + N_c(u) - K'(u)N_Q(u) - K'(u)N_c(u).$$

As an instructive example [**Str89**, Page 38] we consider

$$i\partial_t u = \Delta u - \nabla u \cdot \nabla u \tag{11.26}$$

and choose $v = u - \frac{1}{2}u^2$, i.e., $K(u) = \frac{1}{2}u^2$, cf. Example 11.4.3 for an explanation how to find this transformation. We have

$$AK(u) - K'(u)Au + N_Q(u) = -\mathrm{i}\Delta(\frac{1}{2}u^2) - u(-\mathrm{i}\Delta u) + \mathrm{i}\nabla u \cdot \nabla u = 0$$

such that

$$\partial_t v = Av - K'(u)N_Q(u) = -\mathrm{i}\Delta u - u\mathrm{i}\nabla u \cdot \nabla u$$

In order to obtain an evolutionary problem for v we have to invert the transformation $v = u - K(u)$. For small v this is possible by applying the implicit function theorem. ⌋

We apply this idea to the equation for the error and eliminate the $\mathcal{O}(\varepsilon)$-term $2\varepsilon\widehat{\psi} * \widehat{R}_1$ with a normal form transformation. In order to do so we first diagonalize (11.25) by introducing

$$\mathcal{R} = S^{-1}\begin{pmatrix} R_1 \\ R_2 \end{pmatrix}, \quad \widehat{\Psi} = S^{-1}\begin{pmatrix} \widehat{\psi} \\ \frac{1}{\mathrm{i}\omega(k)}\partial_t\widehat{\psi}(k)) \end{pmatrix}, \text{ where } S = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

We find

$$\partial_t\mathcal{R} = \Lambda\mathcal{R} + 2\varepsilon B(\Psi,\mathcal{R}) + \varepsilon^{\beta}B(\mathcal{R},\mathcal{R}) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi), \tag{11.27}$$

with Λ a symmetric linear map and $B(\cdot,\cdot)$ a bilinear map, which are given in Fourier space by

$$\begin{aligned} \widehat{\Lambda} &= \begin{pmatrix} \mathrm{i}\omega & 0 \\ 0 & -\mathrm{i}\omega \end{pmatrix}, & \mathrm{RES}(\varepsilon\Psi) &= S^{-1}\begin{pmatrix} 0 \\ \mathrm{Res}(\varepsilon\psi) \end{pmatrix}, \\ \widehat{B}(\widehat{U},\widehat{U}) &= \tfrac{1}{\mathrm{i}\omega}S^{-1}\widetilde{B}(S\widehat{U},S\widehat{U}), & \widetilde{B}(\widehat{U},\widehat{V}) &= \begin{pmatrix} 0 \\ \widehat{U}_1 * \widehat{V}_1 \end{pmatrix}, \end{aligned}$$

where $\widehat{U} = (\widehat{U}_1, \widehat{U}_2)$. Then we make a near identity change of variables

$$w = \mathcal{R} + \varepsilon Q(\Psi,\mathcal{R}) \tag{11.28}$$

with Q an autonomous bilinear map. This gives

$$\begin{aligned} \partial_t w =& \partial_t\mathcal{R} + \varepsilon Q(\partial_t\Psi,\mathcal{R}) + \varepsilon Q(\Psi,\partial_t\mathcal{R}) \\ =& \Lambda\mathcal{R} + 2\varepsilon B(\Psi,\mathcal{R}) + \varepsilon^{\beta}B(\mathcal{R},\mathcal{R}) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi) + \varepsilon Q(\partial_t\Psi,\mathcal{R}) \\ &+ \varepsilon Q(\Psi, \Lambda\mathcal{R} + 2\varepsilon B(\Psi,\mathcal{R}) + \varepsilon^{\beta}B(\mathcal{R},\mathcal{R}) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi)) \\ =& \Lambda w - \varepsilon\Lambda Q(\Psi,\mathcal{R}) + 2\varepsilon B(\Psi,\mathcal{R}) + \varepsilon^{\beta}B(\mathcal{R},\mathcal{R}) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi) \\ &+ \varepsilon Q(\partial_t\Psi,\mathcal{R}) + \varepsilon Q(\Psi, \Lambda\mathcal{R} + 2\varepsilon B(\Psi,\mathcal{R}) + \varepsilon^{\beta}B(\mathcal{R},\mathcal{R}) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi)), \end{aligned}$$

and so

$$\begin{aligned} \partial_t w \quad = \quad & \Lambda w + \varepsilon\left(-\Lambda Q(\Psi,\mathcal{R}) + Q(\partial_t\Psi,\mathcal{R})\right. \\ & \left. + Q(\Psi,\Lambda\mathcal{R}) + 2B(\Psi,\mathcal{R})\right) + \mathcal{O}(\varepsilon^2). \end{aligned} \tag{11.29}$$

In order to eliminate the dangerous term $2\varepsilon B(\Psi, w)$ we have to find a bilinear Q such that

$$-\Lambda Q(\Psi,\mathcal{R}) + Q(\partial_t\Psi,\mathcal{R}) + Q(\Psi,\Lambda w) + 2B(\Psi,\mathcal{R}) = 0. \tag{11.30}$$

In this form this equation is hard to analyze. For its simplification we first use that

$$\begin{aligned}
\varepsilon\partial_t\Psi &= \sqrt{2}\partial_t \begin{pmatrix} \varepsilon A_1\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}(k_0x+\omega_0 t)}+\mathcal{O}(\varepsilon^2) \\ \varepsilon A_{-1}\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{-\mathrm{i}(k_0x+\omega_0 t)}+\mathcal{O}(\varepsilon^2) \end{pmatrix} \\
&= \sqrt{2} \begin{pmatrix} \mathrm{i}\omega_0\varepsilon A_1\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}(k_0x+\omega_0 t)}+\varepsilon^2 G_1 \\ -\mathrm{i}\omega_0\varepsilon A_{-1}\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{-\mathrm{i}(k_0x+\omega_0 t)}+\varepsilon^2 G_{-1} \end{pmatrix}
\end{aligned}$$

with $\|\widehat{G}\|_{L^1_\theta}=\mathcal{O}(1)$. Hence, (11.30) is given in lowest order by

$$-\Lambda Q(\Psi,\mathcal{R})+Q(\Upsilon,\mathcal{R})+Q(\Psi,\Lambda w)+2B(\Psi,\mathcal{R})=0, \tag{11.31}$$

where $\Upsilon=\begin{pmatrix}(\Lambda(k_0)\Psi)_1 \\ (\Lambda(-k_0)\Psi)_{-1}\end{pmatrix}$. In Fourier space we have for the j-th component of B that

$$(\widehat{B}(\widehat{\psi},\widehat{\mathcal{R}}))_j = \sum_{m,n=1,2}\int_{\mathbb{R}}\widehat{b}^j_{mn}(k,k-l,l)\widehat{\psi}_m(k-l)\widehat{\mathcal{R}}_n(l)\,\mathrm{d}l,$$

with $\widehat{b}^j_{mn}=\widehat{b}^j_{mn}(k,k-l,l)$ some smooth kernel. Thus, we make the same ansatz for the j-th component of Q, namely

$$(\widehat{Q}(\widehat{\psi},\widehat{\mathcal{R}}))_j = \sum_{m,n=1,2}\int_{\mathbb{R}}\widehat{q}^j_{mn}(k,k-l,l)\widehat{\psi}_m(k-l)\widehat{\mathcal{R}}_n(l)\,\mathrm{d}l,$$

with $\widehat{q}^j_{mn}=\widehat{q}^j_{mn}(k,k-l,l)$ some kernel which we have to compute. Inserting these representations in (11.31) yields the relations

$$\mathrm{i}(\omega_j(k)-\omega_1(k_0)-\omega_n(l))\widehat{q}^j_{1n}(k,k-l,l)=2\widehat{b}^j_{1n}(k,k-l,l), \tag{11.32}$$

and

$$\mathrm{i}(\omega_j(k)-\omega_2(-k_0)-\omega_n(l))\widehat{q}^j_{2n}(k,k-l,l)=2\widehat{b}^j_{2n}(k,k-l,l), \tag{11.33}$$

with $\omega_{1,2}(k)=\pm\omega(k)$.

Example 11.4.3. For system (11.26), where the approximation is replaced by a general function and where only one curve of eigenvalues is involved, the associated relation would be

$$\mathrm{i}(\omega(k)-\omega(k-l)-\omega(l))\widehat{q}(k,k-l,l)=\widehat{b}(k,k-l,l),$$

with $\omega(k)=-k^2$ and $\widehat{b}(k,k-l,l)=\mathrm{i}(k-l)l$. We find

$$\omega(k)-\omega(k-l)-\omega(l)=-k^2+(k-l)^2+l^2=-2(k-l)l$$

such that $\widehat{q}(k,k-l,l)=-\frac{1}{2}$. ⌋

In Fourier space, the approximation $\varepsilon\psi$ has order one amplitude only close to the wave numbers k_0 in the first component and close to the wave numbers $-k_0$ in the second component. Therefore, only wave numbers $|k-l-k_0|\leq\delta$ for (11.32) and only wave numbers $|k-l+k_0|\leq\delta$ for (11.33) have to be taken into account for a small $\delta > 0$ independent of $0 < \varepsilon \ll 1$. For the other wave numbers we set $\widehat{q}^j_{mn} = 0$. Hence, in order to solve (11.32) w.r.t. the $\widehat{q}^j_{mn}$ we need the non-resonance conditions

$$\inf_{j,n\in\{1,2\}} \inf_{k,l\in\mathbb{R},|k-l|\leq\delta} |(\omega_j(k) - \omega_1(k_0) - \omega_n(l))| \geq C > 0 \tag{11.34}$$

and

$$\inf_{j,n\in\{1,2\}} \inf_{k,l\in\mathbb{R},|k-l|\leq\delta} |(\omega_j(k) - \omega_2(-k_0) - \omega_n(l))| \geq C > 0 \tag{11.35}$$

for this $\delta > 0$ fixed. The validity of the non-resonance conditions can be checked graphically by looking for intersections of the curves $k \mapsto \pm\omega(k)$ and $k \mapsto \omega(k_0) \pm \omega(k - k_0)$, see Figure 11.4. Since the asymptotes $k \mapsto \pm k$ and $k \mapsto \omega(k_0) \pm (k - k_0)$ to these curves are separated, for no value of k_0 a quadratic resonance occurs.

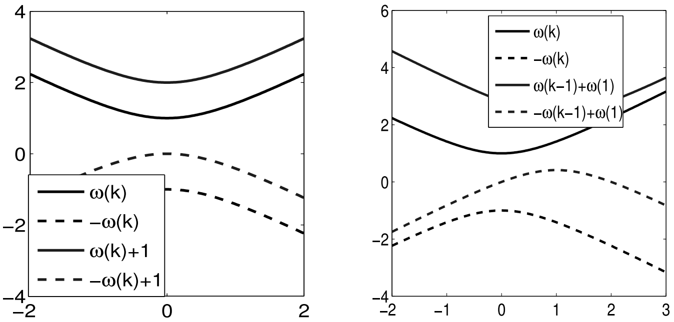

Figure 11.4. The curves $k \mapsto \pm\omega(k)$ and $k \mapsto \omega(k_0) \pm \omega(k - k_0)$ in case $k_0 = 0$ and $k_0 = 1$. There is no intersection and so no quadratic resonance.

Since

$$\sup_{j,m,n\in\{1,2\}} \sup_{k,l\in\mathbb{R},|k-l|\leq\delta} |\widehat{b}^j_{mn}(k, k-l, l)| \leq C < \infty,$$

(11.34) immediately implies

$$\sup_{j,m,n\in\{1,2\}} \sup_{k,l\in\mathbb{R},|k-l|\leq\delta} |\widehat{q}^j_{mn}(k, k-l, l)| \leq C < \infty.$$

As a consequence we obtain

$$\|(Q(\psi, \mathcal{R}))\|_{H^\theta} \leq C\|\psi\|_{C^\theta_b}\|\mathcal{R}\|_{H^\theta} \leq C\|\widehat{\psi}\|_{L^1_\theta}\|\widehat{\mathcal{R}}\|_{L^2_\theta}.$$

Thus, the transformation (11.28) can be inverted with the help of Neumann's series for $\varepsilon > 0$ sufficiently small. We denote the inverse by $\mathcal{R} = \mathcal{R}_\varepsilon(w) = \mathcal{O}(1)$. Therefore, (11.29) transforms into

$$\partial_t w = \Lambda w + \varepsilon^2 F \tag{11.36}$$

with

$$\begin{aligned}\varepsilon^2 F &= \varepsilon^{\beta} B(\mathcal{R}_\varepsilon(w), \mathcal{R}_\varepsilon(w)) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi) + \varepsilon^2 Q(G, \mathcal{R}_\varepsilon(w)) \\ &\quad + \varepsilon Q(\Psi, 2\varepsilon B(\Psi, \mathcal{R}_\varepsilon(w)) + \varepsilon^{\beta} B(\mathcal{R}_\varepsilon(w), \mathcal{R}_\varepsilon(w)) + \varepsilon^{-\beta}\mathrm{RES}(\varepsilon\Psi)).\end{aligned}$$

Since $\varepsilon^2 F$ obeys the same estimates as the one in Lemma 11.2.4 the rest of the proof of the approximation property from §11.2 applies line for line to (11.36).

11.5. Extension of the theory

There are various physically relevant systems, especially the water wave problem, where the previous non-resonance conditions (11.38) and (11.39) are not satisfied. These are systems with trivial resonances at the wave number $k = 0$, additional non-trivial resonances, and quasilinear nonlinearities. Here we give a short overview about the strategies which have been developed in last years to overcome these difficulties. We start this section with some technical improvements of the previous analysis.

11.5.1. Some technical improvements. By two simple technical changes in the previous proofs the non-resonance conditions can be simplified and less regularity for the initial conditions of the NLS equation is needed.

Less regularity is needed. Since in Fourier space the approximation is strongly concentrated at integer multiples of the basic wave number k_0 we only make a small error if we cut-off the approximation outside small neighborhoods of $k_0\mathbb{Z}$.

The formal expansion of the curve of eigenvalues and of the kernels in the multilinear maps can be estimated with the aid of the following lemma.

Lemma 11.5.1. *Let $\theta, \theta_0 \geq 0$, $\theta_\infty \in \mathbb{R}$, and let $g(k)$ satisfy*

$$|g(k)| \leq C \min(|k - k_0|^{\theta_0}, (1 + |k|)^{\theta_\infty}).$$

Then

$$\|g(\cdot)\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}(\cdot - k_0))\|_{L^2_\theta} \leq C\varepsilon^{\theta_0 - 1/2}\|\widehat{A}\|_{L^2_{\max(\theta+\theta_\infty, \theta_0)}}.$$

Proof. This follows immediately from the fact that the left-hand side of this inequality can be estimated by

$$\leq \sup_{k\in\mathbb{R}} \left| g(k)(1+\frac{\varepsilon}{|k-k_0|})^{-\max(\theta+\theta_\infty,\theta_0)}(1+|k-k_0|)^{\theta}\right| \; \|\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}\cdot)\|_{L^2_{\max(\theta+\theta_\infty,\theta_0)}}$$
$$\leq \sup_{K\in\mathbb{R}} \left| C\min(\frac{|\varepsilon K|^{\theta_0}}{(1+|K|)^{\theta_0}}, \frac{(1+|k_0+\varepsilon K|)^{\theta+\theta_\infty}}{(1+|K|)^{\theta+\theta_\infty}})\right| \; \varepsilon^{-1/2}\|\widehat{A}(\cdot)\|_{L^2_{\max(\theta+\theta_\infty,\theta_0)}}$$

where the loss of $\varepsilon^{-1/2}$ is due to the scaling properties of the L^2-norm. $\square$

This lemma can be applied for instance to

$$g(k) = \omega(k) - \omega_0 - c_g k - \frac{1}{2}\omega''(k_0)k^2 = \mathcal{O}(|k|^3)$$

such that the residual terms can also be estimated for the diagonalized system in Fourier space.

Example 11.5.2. With the simple approach from §11.2 we need $\theta_A \geq \theta + 4$ in order to estimate the error in H^θ if we take the approximation ψ_2. The reason is as follows. The highest loss of regularity comes from $\|\partial_T^2 A_3\|_{H^\theta}$ which can be estimated by $\|\partial_T^2 A\|_{H^\theta}$. This can be estimated further via the right-hand side of the NLS equation by $\|A\|_{H^{\theta+4}}$. For the associated diagonalized system we have to expand ω at k_0 up to order three, i.e., we use

$$\begin{aligned} &|\omega(k) - (\omega(k_0) + \omega'(k_0)(k-k_0) + \frac{1}{2}\omega''(k_0)(k-k_0)^2)\chi(|k-k_0|)| \\ &\leq C \min(|k-k_0|^3, |k|) \end{aligned}$$

where χ is a C_0^∞ function with $\chi(k) \in [0,1]$, $\chi(k) = 0$ for $|k| \geq 2$ and $\chi(k) = 1$ for $|k| \leq 1$. Hence, the application of Lemma 11.5.1 allows to reduce the above value of θ_A to a value satisfying $\theta_A \geq \max(\theta + 1, 3)$ since for all other terms less derivatives are needed. ⌋

Sometimes it is possible to reduce the value for θ_A even further by using a cut-off function in Fourier space to mollify the original approximation. Let again $\chi \in C_0^\infty$ with $0 \leq \chi(k) \leq 1$, $\chi(k) = 0$ for $|k| \geq 2$, and $\chi(k) = 1$ for $|k| \leq 1$. Instead of $\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}\cdot)$ we work with $\chi(\cdot)\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}\cdot)$. Then we have

$$\|g(\cdot)\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}(\cdot))\|_{L^2_\theta} \leq C\varepsilon^{\theta-1/2}\|\widehat{A}\|_{L^2_\theta} \tag{11.37}$$

for $g(k) = 1 - \chi(k)$. Hence, the modified approximation and the original approximation are $\mathcal{O}(\varepsilon^{\theta-1/2})$ close to each other in H^θ if the ansatz functions A are in H^θ. Moreover, we have

Lemma 11.5.3. *Let $\theta, \theta_0 \geq 0$, $\theta_\infty \in \mathbb{R}$, and let $g(k)$ satisfy*

$$|g_1(k)| \leq C \min(|k-k_0|^{\theta_0}, (1+|k|)^{\theta_\infty}).$$

Then

$$\|g_1(\cdot)\chi(\cdot)\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}(\cdot - k_0))\|_{L^2_\theta} \leq C\varepsilon^{\theta_0-1/2}\|\widehat{A}\|_{L^2_{\theta_0}}.$$

Proof. This follows by applying Lemma 11.5.1 to $g(k) = g_1(k)\chi(k)$. □

Therefore, less regularity θ_A is needed for estimating the residual terms in the error equations. However, the original approximation and the cut-off approximation have to be shown to be $\mathcal{O}(\varepsilon^{3/2})$ close. Therefore, in order to have such an estimate in H^θ we need at least $A \in H^{\theta+1}$, cf. (11.37). For an approximation result optimized with this respect see [**MU16**].

Weakening of the non-resonance condition. Checking the non-resonance conditions (11.34) and (11.35) is non-trivial due to the fact that a two-dimensional function has to be bounded away from zero. The following version of Lemma 11.5.1 which avoids the loss of $\varepsilon^{-1/2}$ due to the scaling properties of the L^1-norm allows to reduce the checking of the non-resonance condition to a one-dimensional problem

Lemma 11.5.4. *Let $\theta, \theta_0 \geq 0$ and let $g(k)$ satisfy $|g(k)| \leq C|k-k_0|^{\theta_0}$. Then*

$$\|g(\cdot)\varepsilon^{-1}\widehat{A}(\varepsilon^{-1}(\cdot - k_0))\|_{L^1_\theta} \leq C\varepsilon^{\theta_0}\|\widehat{A}\|_{L^1_{\theta+\theta_0}}.$$

In order to remove l from (11.34) and (11.35) we use the concentration of the approximation at the wave numbers $k = \pm k_0$ and estimate with the help of Lemma 11.5.4

$$\begin{aligned}
& \| \int_{\mathbb{R}} (\omega_n(l) - \omega_n(\cdot - k_0))\widehat{\psi}(\cdot - l)\widehat{w}_n(l)\,\mathrm{d}l\|_{L^2_\theta} \\
= \; & \| \int_{\mathbb{R}} (\omega_n(\cdot - (\cdot - l)) - \omega_n(\cdot - k_0))\widehat{\psi}(\cdot - l)\widehat{w}_n(l)\,\mathrm{d}l\|_{L^2_\theta} \\
\leq \; & \left\| \sup_{k\in\mathbb{R}} |\omega_n(k - (\cdot - l)) - \omega_n(k - k_0)||\widehat{\psi}(\cdot - l)| \right\|_{L^1_\theta} \|w_n\|_{L^2_\theta} \leq C\varepsilon\|w_n\|_{L^2_\theta}.
\end{aligned}$$

Hence, if we replace $\omega_n(l)$ by $\omega_n(k-k_0)$ in (11.29) we produce additional terms of order $\mathcal{O}(\varepsilon^2)$ which can be included into the terms which do not make trouble in obtaining error estimates on the $\mathcal{O}(1/\varepsilon^2)$ time scale. Hence, the non-resonance condition (11.34) can be weakened to

$$\inf_{j,n\in\{1,2\}} \inf_{k\in\mathbb{R}} |\omega_j(k) - \omega_1(k_0) - \omega_n(k-k_0)| \geq C > 0 \tag{11.38}$$

and (11.35) can be weakened to

$$\inf_{j,n\in\{1,2\}} \inf_{k\in\mathbb{R}} |\omega_j(k) - \omega_2(-k_0) - \omega_n(k-k_0)| \geq C > 0. \tag{11.39}$$

11.5.2. Systems with a trivial resonance at $k = 0$. In physical systems with conserved quantities very often the eigenvalue zero occurs at the wave number $k = 0$, i.e., $\omega_j(0) = 0$ for a j in some index set. The most prominent example is the so called water wave problem, cf. §12.2.1, for which the NLS equation has been derived first [**Zak68**]. The fact $\omega_j(0) = 0$ will violate the non-resonance condition (11.34) and will always lead to quadratic resonances which at a first view will not allow to remove the quadratic terms. However, due to the fact that the eigenvalue zero is created by a conserved quantity also the nonlinear terms vanish at the wave number $k = 0$. The simplest example for such a system is the so called Boussinesq equation

$$\partial_t^2 u = \partial_x^2 u + \partial_x^2\partial_t^2 u + \partial_x^2(u^2), \tag{11.40}$$

with $x \in \mathbb{R}$, $t \in \mathbb{R}$, and $u(x,t) \in \mathbb{R}$. This model occurs as a long wave limit for the water wave problem, cf. Chapter 12. The linearized problem possesses solutions $\mathrm{e}^{\mathrm{i}kx+i\omega t}$ with dispersion relation

$$\omega^2 = \frac{k^2}{1+k^2}. \tag{11.41}$$

Hence, the non-resonance condition (11.38) is not satisfied as can be seen in Figure 11.5. The resonance at the wave number $k=0$ is trivial since the nonlinear terms vanish at the wave number $k=0$, too. However, the resonance at the wave number $k=k_0$ is non-trivial. One possibility to get rid of this second resonance is a k-dependent scaling of the error function.

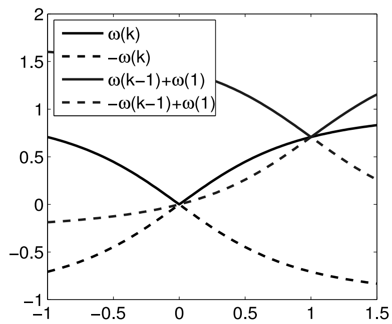

Figure 11.5. Intersection of the curves $k \mapsto \pm\omega(k)$ and $k \mapsto \omega(k_0) \pm \omega(k-k_0)$, here with $k_0 = 1$. There are two intersections at $k=0$ and $k=k_0$.

Before we explain this in more detail we derive the NLS equation for the Boussinesq model (11.40). As before we make the ansatz

$$\begin{aligned} \varepsilon\psi_{\mathrm{NLSbouss}} =& \varepsilon A_1\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}(k_0x+\omega_0 t)} + \mathrm{c.c.} \\ &+ \varepsilon^2 A_2\left(\varepsilon(x-c_g t),\varepsilon^2 t\right)\mathrm{e}^{2\mathrm{i}(k_0x+\omega_0 t)} + \mathrm{c.c.} \\ &+ \varepsilon^2 A_0\left(\varepsilon(x-c_g t),\varepsilon^2 t\right). \end{aligned} \tag{11.42}$$

We find as before at $\varepsilon\mathbf{E}$ the linear dispersion relation (11.41) and at $\varepsilon^2\mathbf{E}$ the condition for the linear group velocity c_g. At $\varepsilon^3\mathbf{E}$ we find

$$2\mathrm{i}\omega_0(1+k_0^2)\partial_T A_1 = (1-c_g^2-\omega_0^2)\partial_X^2 A_1 - 2k_0^2(A_1A_0 + A_2A_{-1}).$$

The algebraic relations which are found at $\varepsilon^4\mathbf{E}^0$ and $\varepsilon^2\mathbf{E}^2$

$$\begin{aligned} \varepsilon^4\mathbf{E}^0: &\quad 0 = (1-c_g^2)^2\partial_X^2 A_0 + 2\partial_X^2 A_1A_{-1}, \\ \varepsilon^2\mathbf{E}^2: &\quad 0 = -(-4\omega_0^2+4k_0^2+16\omega_0^2k_0^2)A_2 - 4k_0^2A_1^2, \end{aligned}$$

can be solved w.r.t. A_0 since $(1-c_g^2)^2 \neq 0$ and w.r.t. A_2 since $-4\omega_0^2+4k_0^2+16\omega_0^2k_0^2 \neq 0$. Inserting the solutions for A_0 and A_2 into the equation for A_1 finally yields the NLS equation

$$2\mathrm{i}\omega_0(1+k_0^2)\partial_T A_1 = (1-c_g^2-\omega_0^2)\partial_X^2 A_1 + \gamma A_1 |A_1|^2, \tag{11.43}$$

with

$$\gamma = \frac{4k_0^2}{1-c_g^2} + \frac{k_0^4}{-\omega_0^2+k_0^2+4\omega_0^2k_0^2}.$$

The error $\varepsilon^\beta R = u - \varepsilon\Psi$ satisfies

$$\partial_t^2 R = \partial_x^2 R + \partial_x^2\partial_t^2 R + 2\varepsilon\partial_x^2(\Psi R) + \mathcal{O}(\varepsilon^2).$$

Writing this as a first order system gives two equations of the form

$$\partial_t \widehat{R}_j(k,t) = \mathrm{i}\omega_j(k)\widehat{R}_j(k,t) + \varepsilon\rho_j(k)\int_{\mathbb{R}} \widehat{\psi}(k-m,t)\widehat{R}_j(m,t)\,\mathrm{d}m + \mathcal{O}(\varepsilon^2),$$

with $\omega_j(0) = \rho_j(0) = 0$, but non-vanishing $\omega_j'(0)$ and $\rho_j'(0)$. The approximation $\varepsilon\psi$ is concentrated at the wave numbers $\pm k_0$. Hence, a subsystem is given by

$$\begin{aligned}
\partial_t \widehat{R}_j(k_0,t) &= \mathrm{i}\omega_j(k_0)\widehat{R}_j(k_0,t) + \varepsilon\rho_j(k_0)\,\widehat{\psi}(k_0,t)\widehat{R}_j(0,t) + h.o.t.,\\
\partial_t \widehat{R}_j(0,t) &= \mathrm{i}\omega_j(0)\widehat{R}_j(0,t) + \varepsilon\rho_j(0)\widehat{\psi}(-k_0,t)\widehat{R}_j(k_0,t) + h.o.t..
\end{aligned}$$

For the second equation we have

$$\omega_j(0) - \omega_j(-k_0) - \omega_j(k_0) = 0 + \omega_j(k_0) - \omega_j(k_0) = 0$$

such that the non-resonance condition is not satisfied for $k=0$. However, we have $\rho_j(0) = 0$ such that the nonlinear terms vanish for this resonant wave number, too. Hence, this resonance is called trivial. Since

$$\omega_j(k) - \omega_j(-k_0) - \omega_j(k+k_0) = \mathcal{O}(|k|) \qquad \text{and} \qquad \rho_j(k) = \mathcal{O}(|k|),$$

with the denominator in the normal form transform also the nominator vanishes and therefore the quadratic nonlinear terms close to $k=0$ can be eliminated. For the first of these equations we have

$$\omega_j(-k_0) - \omega_j(k_0) - \omega_j(0) = \omega_j(k_0) - 0 - \omega_j(k_0) = 0$$

such that the non-resonance condition is not satisfied for $k = -k_0$ and similarly for $k = k_0$. However, we have $\rho_j(k_0) \neq 0$ such that the nonlinear terms do not vanish for this resonant wave number. Hence, this resonance is called non-trivial. One way to get rid of this difficulty is to scale $R_j(0)$ with $\varepsilon^{\beta+1}$ instead of ε^β. Doing so we obtain

$$\begin{aligned}
\partial_t \widehat{R}_j(k_0,t) &= \mathrm{i}\omega_j(k_0)\widehat{R}_j(k_0,t) + \varepsilon^2\rho_j(k_0)\widehat{\psi}(k_0,t)\widehat{R}_j(0,t) + h.o.t.,\\
\partial_t \widehat{R}_j(0,t) &= \mathrm{i}\omega_j(0)\widehat{R}_j(0,t) + \rho_j(0)\widehat{\psi}(-k_0,t)\widehat{R}_j(k_0,t) + h.o.t..
\end{aligned}$$

Hence, in the first equation the nonlinear terms do not make problems anymore. However, in the second equation we have now terms of order $\mathcal{O}(1)$

which can be eliminated in the full system with the argument from above. The terms of order $\mathcal{O}(\varepsilon)$ resulting from this transformation in the second equation can either be eliminated by another transformation or are of long wave form, i.e., of a similar form as $\varepsilon\partial_x(B(\varepsilon x)R(x))$ in physical space. Such terms can be estimated by energy estimates to have an $\mathcal{O}(\varepsilon^2)$ influence on the dynamics, cf. Chapter 12.

This can be made rigorous for the full system by making the ansatz

$$\varepsilon^\beta \vartheta(k)\widehat{R} = u - \varepsilon\Psi$$

for the error function R, where $\vartheta(k) = \min(\varepsilon + |k|/\delta, 1)$ with $\delta > 0$ sufficiently small, but independent of $0 < \varepsilon \ll 1$. This has been carried out in [**Sch98a**] with a correction explained in [**DS06**].

The above idea has been transferred to the water wave problem without surface tension in case of finite depth in [**DSW16**]. However, the water wave problem is a quasilinear problem. Below we will explain the additional difficulties occurring for such systems. But also the so called FPU system falls into this class, cf. Exercise 11.4. In [**Sch10**] it has been explained that the proofs of the approximation theorems given for the PDE systems can be transferred almost line for line to the FPU system by looking at the Fourier transformed FPU system.

11.5.3. Stable and unstable non-trivial resonances. For the water wave problem with small surface tension, i.e., for surface tension parameter $\sigma \in (0, 1/3)$ additional resonances are present in the system, i.e., there exist spatial wave numbers k_1, k_2, k_3 and associated temporal wave numbers ω_1, ω_2, ω_3 satisfying

$$k_1 + k_2 + k_3 = 0 \qquad \text{and} \qquad \omega_1 + \omega_2 + \omega_3 = 0,$$

cf. Exercise 11.6 and §12.2.1. The same happens for dispersive wave systems with spatially periodic coefficients, or the poly-atomic FPU model. There are at least two different approaches to get rid of these resonances.

Solutions of the NLS equation which are analytic in a strip in the complex plane decay with some exponential rate in Fourier space. By nonlinear interaction, the solutions of the original system have a Fourier mode distribution which is strongly localized at integer multiples of the basic wave number k_0. In between, the solutions will be exponentially small, i.e., the original system can be solved in a weighted L^1-space equipped with the norm

$$\|\widehat{u}\|_{L^1_w} = \int_{\mathbb{R}} |\widehat{u}(k)| w(k)\, \mathrm{d}k$$

with

$$1/w(k) = \sup_{m\in\mathbb{Z}} |\mathrm{e}^{-\alpha\varepsilon^{-1}|k - mk_0|}|. \tag{11.44}$$

Hence, the Fourier modes associated to the resonant wave numbers are exponentially small initially, i.e., of order $\mathcal{O}(\exp(-r\varepsilon^{-1}))$ for an $r > 0$, independent of $0 < \varepsilon \ll 1$. The quadratic resonances will lead to growth rates $\mathcal{O}(\exp(\varepsilon t))$ for solutions of order $\mathcal{O}(\varepsilon)$. Hence, it takes a time of order $\mathcal{O}(1/\varepsilon^2)$ to have $\mathcal{O}(\exp(-r\varepsilon^{-1}))\mathcal{O}(\exp(\varepsilon t)) = \mathcal{O}(1)$, i.e., it takes the time scale of the NLS equation for the error to grow to the order of the NLS approximation. This idea can be used to prove error estimates on an $\mathcal{O}(1/\varepsilon^2)$ time scale for the validity of the NLS approximation also in case of additional non-trivial resonances, if the solutions of the NLS equation are analytic in a strip in the complex plane, and if the set of wave numbers resonant to k_0 is separated from the set of integer multiples of the basic wave number k_0. This idea can be made rigorous by making the coefficient α time-dependent, i.e., by choosing $\alpha(t) = \alpha_0 - \check{\beta}\varepsilon^2 t$ in (11.44). This idea has been explained in [**Sch98c**] and carried out in [**DHSZ16**].

The second approach is based on a more detailed analysis of the resonances. Consider a basic wave number $k_1 = k_0$, resonant to wave numbers k_2 and k_3. The ansatz

$$u(x,t) = \varepsilon A_1(\varepsilon t)\mathrm{e}^{\mathrm{i}(k_1 x+\omega_1 t)}\varphi_1 + \varepsilon A_2(\varepsilon t)\mathrm{e}^{\mathrm{i}(k_2 x+\omega_2 t)}\varphi_2 + \varepsilon A_3(\varepsilon t)\mathrm{e}^{\mathrm{i}(k_3 x+\omega_3 t)}\varphi_3,$$

with vectors φ_j, then yields a three wave interaction (TWI) system

$$\partial_T A_1 = \mathrm{i}\gamma_1 \overline{A_2 A_3}, \qquad \partial_T A_2 = \mathrm{i}\gamma_2 \overline{A_1 A_3}, \qquad \partial_T A_3 = \mathrm{i}\gamma_3 \overline{A_1 A_2},$$

with coefficients $\gamma_j \in \mathbb{R}$, associated to the resonances, cf. [**Sch05**, §3.3]. In [**Sch05**, Theorem 3.8] a NLS approximation theorem has been shown in case that the subspace $\{A_2 = A_3 = 0\}$ associated to the wave number $k_1 = k_0$ is stable in the TWI system, cf. Figure 11.6. The proof is based on a mixture of normal form transforms for the non-resonant wave numbers and energy estimates for the resonant wave numbers. In [**DS06**] the ideas of [**Sch98a**] and [**Sch05**] are brought together to handle Boussinesq equations which model the water wave problem in case of small positive surface tension.

There is also a counter-example [**Sch05**, §4.1] showing that the NLS equation fails to approximate solutions in the original system in case of an unstable k_0-subspace in the associated TWI system and periodic boundary conditions in the original system. This idea has been carried out in [**SSZ15**] for the water wave problem with surface tension and periodic boundary conditions showing that there exists a continuum of wave numbers and values of surface tension where the NLS approximation does not make correct predictions.

The situation on the whole real line for an unstable resonance is still open. In this case the different group velocities $\omega'(k_j)$ at the resonant wave numbers k_j no longer can be neglected. For a thorough discussion

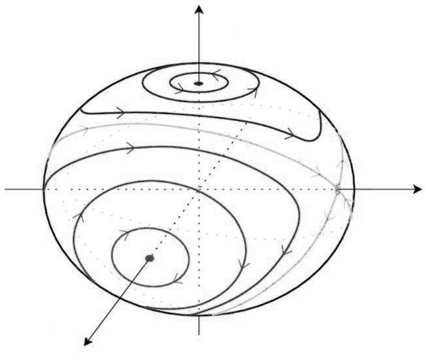

Figure 11.6. The phase portrait of the TWI system in the energy surface which is an ellipsoid since due to conservation of energy not all γ_j have the same sign. The axes are invariant subspaces associated to the wave numbers k_j. There are one unstable and two stable subspaces.

see [**Sch05**, §4.2]. A recent attempt to understand this situation can be found in [**MN13**].

11.5.4. Quasilinear quadratic nonlinearities. The first and very general NLS approximation theorem was proved in [**Kal88**] for general quasilinear dispersive wave systems. However, the occurrence of quasilinear quadratic terms has been excluded explicitly. A typical example is given by the quasilinear wave equation

$$\partial_t^2 u = \partial_x^2 u - u + \frac{1}{2}\partial_x^2(u^2) + u^3, \tag{11.45}$$

with $x \in \mathbb{R}$, $t \in \mathbb{R}$, and $u(x,t) \in \mathbb{R}$. With the NLS ansatz (11.42) we obtain the linear dispersion relation $\omega_0^2 = k_0^2 + 1$, the group velocity $c_g = k_0/\omega_0$, and the NLS equation,

$$i\nu_1\partial_T A + \nu_2\partial_X^2 A + \nu_3 A|A|^2 = 0, \tag{11.46}$$

with

$$\nu_1 = 2\omega_0, \qquad \nu_2 = (1-c^2), \qquad \nu_3 = \frac{9\alpha_c - 2k_0^4}{3}, \tag{11.47}$$

for the description of small spatio-temporal modulations of the underlying carrier wave $e^{i(k_0 x+\omega_0 t)}$. As above, the idea is to use a normal form transform to eliminate the quadratic terms. For (11.45) the quasilinear quadratic nonlinearity causes the required normal form transform to lose regularity such that the original quasilinear system can no longer be handled after applying the transformation (see §11.4). This can be seen as follows. Writing the error equations as first order system gives nonlinear terms with growth rates proportional to $|k|$ for $|k| \to \infty$. This results in the normal form transform in a nominator proportional to $|k|$ and in a denominator proportional to

$$\omega(k) - \omega(k_0) - \omega(k - k_0) = \mathcal{O}(1)$$

for $|k| \to \infty$. Thus, the normal form transform is of the form identity plus a term which is small but loses one derivative. Hence, the normal form transform can no longer be inverted with Neumann's series.

So far there are only a few quasilinear systems where approximation results for the NLS approximation could have been established. One example is where the right-hand side of the quasilinear dispersive wave system only loses half a derivative, i.e., a factor of $\sqrt{k}$ in Fourier space, as a result of the normal form transformation. In this case the elimination of the quadratic terms is still possible and the transformed system can be handled with the Cauchy-Kowalevskaya theorem [**SW11**]. The Lagrangian formulation of the water wave problem in case of finite depth and zero surface tension falls into this class, cf. [**DSW16**]. In case of zero surface tension and infinite depth such a result has been established in [**TW12, Tot15**] by finding a special transformation which allows to eliminate all quadratic terms for this particular system without loss of regularity. Another example is in the context of the KdV equation where the result can be obtained by simply applying the Miura transformation [**Sch11**], cf. Exercise 11.5. In [**CS13**] numerical evidence is given that the NLS approximation is valid for (11.45). Very recently it turned out that the solutions of the transformed quasilinear system can be estimated by more clever energy estimates [**Dül16, CW16, DH16**].

Although relevant systems can be handled with the last approach a validity theory for general dispersive wave systems with quasilinear quadratic terms is still an open problem.

11.6. Pulse dynamics in photonic crystals

One of the major goals of photonics is the construction of 'electronic' devices where the electrons are completely replaced by photons. Photonic crystals turned out to be a suitable tool for the construction of such devices. They consist of a dielectric material such as glass with a periodic structure with a period comparable to the wave length of light. Due to the periodic structure the linearized problem is no longer solved by Fourier modes, but by so called Bloch modes. The curves of eigenvalues plotted as a function over the Bloch wave numbers can now possess horizontal tangencies, i.e., vanishing group velocities. Thus, in principle, standing light pulses are possible. At the horizontal tangencies very often spectral gaps occur, i.e., there are temporal wave numbers for which the associated wave cannot travel through the photonic crystal. The first fact allows to use photonic crystals as optical storage the second fact to use them for filtering. We will explain the possibility of standing light pulses in photonic crystals by using the NLS approximation.

For simplicity we restrict ourselves again to a nonlinear wave equation

$$\partial_t^2 u = \chi_1 \partial_x^2 u - \chi_2 u + \chi_3 u^3, \tag{11.48}$$

but now with spatially periodic coefficients $\chi_j(x) = \chi_j(x+L)$ for a $L > 0$, $x \in \mathbb{R}$, $t \in \mathbb{R}$, and $u(x,t) \in \mathbb{R}$. Moreover, we assume $\chi_1(x) \geq \gamma_1 > 0$ and $\chi_2(x) \geq \gamma_2 > 0$.

11.6.1. The eigenvalue problem for photonic crystals. We start by reviewing a number of well known results, cf. [**Eas73, Sca99**]. For notational simplicity assume here $L = 2\pi$. The linear problem

$$\partial_t^2 u = \chi_1 \partial_x^2 u - \chi_2 u$$

possesses solutions $u(x,t) = \mathrm{e}^{\mathrm{i}\omega t} v(x)$ where v satisfies

$$-\omega^2 v = \chi_1 \partial_x^2 v - \chi_2 v.$$

Uniformly bounded solutions are given by Bloch waves

$$v(x) = w(x)\mathrm{e}^{\mathrm{i}\ell x},$$

with $w(x) = w(x+2\pi)$ satisfying the eigenvalue problem

$$L(\ell)w = -\chi_1(\partial_x + \mathrm{i}\ell)^2 w + \chi_2 w = \omega^2 w.$$

With $w(x) = w(x+2\pi)$ also $w(x)\mathrm{e}^{\mathrm{i}nx}$ for $n \in \mathbb{Z}$ is 2π-periodic. Hence we can restrict ourselves to $\ell \in [-\frac{1}{2}, \frac{1}{2})$, the so called Brillouin zone.

Lemma 11.6.1. *The properties of $L(\ell)$ are as follows:*

a) $L(\ell)$ is self-adjoint in $L^2_{\mathrm{per}}(\chi_1^{-1}\,\mathrm{d}x)$.

b) $L(\ell)$ is positive definite.

c) $L(\ell)$ has discrete spectrum with ∞ the only accumulation point. All eigenvalues are real, semi-simple, and non-negative.

Proof. a) follows from

$$\begin{aligned}
(L(\ell)w_1, w_2)_{L^2(\chi_1^{-1}\,\mathrm{d}x)} &= \int_0^{2\pi} -\chi_1((\partial_x+\mathrm{i}\ell)^2 w_1)\overline{w}_2 \frac{1}{\chi_1}\,\mathrm{d}x + \int_0^{2\pi} (\chi_2 w_1)\overline{w}_2 \frac{1}{\chi_1}\,\mathrm{d}x \\
&= \int_0^{2\pi} ((\partial_x+\mathrm{i}\ell)w_1)\overline{(\partial_x+\mathrm{i}\ell)w_2}\,\mathrm{d}x + \int_0^{2\pi} \chi_2 w_1 \overline{w}_2 \frac{1}{\chi_1}\,\mathrm{d}x \\
&= (w_1, L(\ell)w_2)_{L^2(\chi_1^{-1}\,\mathrm{d}x)}.
\end{aligned}$$

b) follows from

$$(L(\ell)w, w)_{L^2(\chi_1^{-1}\,\mathrm{d}x)} = \int_0^{2\pi} |(\partial_x + \mathrm{i}\ell)w|^2 + \frac{\chi_2}{\chi_1}|w|^2\,\mathrm{d}x > 0, \qquad (w \neq 0).$$

c) follows from the fact that for fixed ℓ the operator $L(\ell) : H^2_{\mathrm{per}}(\mathbb{R}/2\pi\mathbb{Z}) \to L^2_{\mathrm{per}}(\mathbb{R}/2\pi\mathbb{Z})$ is elliptic, i.e., $L(\ell)$ has a compact resolvent and so $L(\ell)$ has discrete spectrum with ∞ the only accumulation point. Due to the self-adjointness all eigenvalues are real and semi-simple. Due to the positive definiteness all eigenvalues are non-negative. □

Lemma 11.6.2. *Except of intersection points the curves of eigenvalues $\ell \mapsto \omega_n(\ell)$ of $L(\ell)$ are smooth.*

Proof. This follows by a simple perturbation argument, cf. [**Kat95**]. □

11.6.2. The computation of the eigenvalues. Here and the subsequent example we assume $L = 1$, hence $\ell \in [-\pi, \pi)$. The eigenvalue problem

$$\chi_1 \partial_x^2 u - \chi_2 u = -\lambda u \tag{11.49}$$

is a second order scalar ODE with spatially periodic coefficients. We write (11.49) as first order system

$$\begin{array}{rcl} \partial_x u(x) & = & v, \\ \partial_x v(x) & = & -s(x)\lambda u(x) + q(x)u(x), \end{array} \tag{11.50}$$

where $s(x) = 1/\chi_1(x)$ and $q(x) = \chi_2(x)/\chi_1(x)$. The fundamental matrix of (11.50) is denoted by $\Phi_\lambda = \Phi_\lambda(x, x_0)$ where $\Phi_\lambda(x_0, x_0) = I$. Floquet's theorem 2.1.17 shows that

$$\Phi_\lambda(x, x_0) = P_\lambda(x, x_0)\mathrm{e}^{(x-x_0)M_\lambda},$$

with $P_\lambda(x, x_0) = P_\lambda(x+1, x_0)$ and a matrix M_λ, independent of x and x_0. Note that M_λ is not unique since $\mathrm{e}^{2\pi \mathrm{i} n} = 1$ for $n \in \mathbb{Z}$. The eigenvalues of M_λ are the Floquet exponents. The two eigenvalues ρ_- and ρ_+ of the monodromy matrix $C_\lambda = \mathrm{e}^{M_\lambda}$ are the Floquet multipliers. Since the trace of the linear vector field on the right-hand side of (11.50) vanishes we have the conservation of the phase volume for (11.50) leading to $\rho_+\rho_- = 1$, cf. the proof of Theorem 4.1.3. Hence, the Floquet multipliers can be computed via the characteristic polynomial

$$\rho^2 - D(\lambda)\rho + 1 = 0$$

and are given by

$$\rho_\pm(\lambda) = \frac{1}{2}D(\lambda) \pm \frac{1}{2}\sqrt{(D(\lambda))^2 - 4},$$

where the trace of the monodromy matrix, $D(\lambda) = \operatorname{trace} C_\lambda$, is called the discriminant. We find that

a) if $|D(\lambda)| > 2$ then the Floquet multipliers $\rho_\pm(\lambda)$ are real. As a consequence the solutions have exponential growth or decay w.r.t. x.

b) if $|D(\lambda)| < 2$ then the Floquet multipliers $\rho_\pm(\lambda)$ are on the complex unit circle. As a consequence the solutions are uniformly bounded w.r.t. x.

c) if $|D(\lambda)| = 2$ then the Floquet multipliers $\rho_\pm(\lambda)$ are 1 or -1. In this case we have at most linear growth

Example 11.6.3. We consider the eigenvalue problem (11.49) with

$$s(x) = \chi_{[0,6/13]} + 16\chi_{(6/13,7/13)} + \chi_{[7/13,1]}(x \bmod 1)$$

and $q(x) = \mu \in \mathbb{R}$ from [**BCBLS11**]. For this choice the ODE can be solved and the monodromy matrix can be computed explicitly. We find for the discriminant

$$D(\lambda) = \frac{25}{8}\cos\left(\frac{16}{13}\sqrt{\lambda+\mu}\right) - \frac{9}{8}\cos\left(\frac{8}{13}\sqrt{\lambda+\mu}\right). \tag{11.51}$$

See Exercise 11.7. The graph $\lambda \mapsto D(\lambda)$ of the discriminant and the associated dispersion relation can be found in Figure 11.7 for $\mu = 0$. ⌋

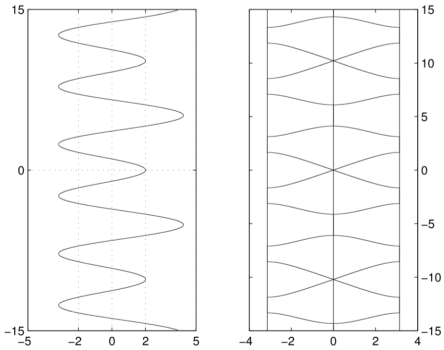

Figure 11.7. The left panel shows the multivalued inverse of the map $\lambda \mapsto D(\lambda)$ and the right panel the associated curves of eigenvalues $k \mapsto \omega^2(k)$ for Example 11.6.3.

For frequencies ω with $\omega^2 = \lambda$ which fall into a spectral gap the incoming wave is damped in the photonic crystal with some exponential rate w.r.t. the depth of penetration x. Hence, photonic crystals can be used as filters. This is one of the reasons why the wings of a butterfly show their colorful appearance.

Remark 11.6.4. There is a relation between the regularity of the coefficients χ_1 and χ_2 on one side and the size of the spectral gaps in

$$\{\omega_n(l) : l \in (-1/2, 1/2],\ n \in \mathbb{Z} \setminus \{0\}\} \subset \mathbb{R}$$

on the other side. For continuous or even smoother χ_1 and χ_2 the spectral gaps become smaller as n increases. According to [**Eas73, Nti76**], the size of the gaps which are found at $\omega \sim n$ decays at least with $1/n^{\theta+1}$ for $n \to \infty$ if $\chi_1 \in C_b^\theta$ and $\chi_2 \in C_b^{\theta-2}$, i.e., the more regular χ_1 and χ_2 are, the faster the gaps close.

Remark 11.6.5. Spectral gaps can be obtained from a spatially homogeneous situation by adding small spatially periodic perturbations to the coefficients. As an example consider the eigenvalue problem

$$(\partial_x + \mathrm{i}\ell)^2 w(x) + \omega^2(1 + 2\varepsilon \cos(2x))w(x) = 0,$$

with $w(x) = w(x + 2\pi)$, $\ell \in [-\frac{1}{2}, \frac{1}{2})$, and $0 \le \varepsilon \ll 1$. For $\varepsilon = 0$ the problem is given by

$$(\partial_x + \mathrm{i}\ell)^2 w + \omega^2 w = 0,$$

and can be solved by $w(x) = \mathrm{e}^{\mathrm{i}nx}$ with $n \in \mathbb{Z}$ and associated eigenvalues $\omega_n^2(\ell) = (n + \ell)^2$. Hence, at $(\ell, \omega) = (0, 1)$ there is a crossing of the curves of eigenvalues. All single eigenvalues ω vary smoothly w.r.t. small ε since $2\varepsilon \cos(2x)\cdot$ is a small perturbation of the operator $(\partial_x + \mathrm{i}\ell)^2$. Hence, the smooth curves $\ell \mapsto \omega_n(\ell)$ will only vary slightly w.r.t. ε. However, at the crossing points the curves can split. As an example we consider the point $(\ell, \omega) = (0, 1)$. We use $2\cos 2x = \mathrm{e}^{2\mathrm{i}x} + \mathrm{e}^{-2\mathrm{i}x}$, make the ansatz $w(x) = \sum_{n \in 2\mathbb{Z}+1} c_n \mathrm{e}^{\mathrm{i}nx}$, and set $\omega^2 = 1 + \widetilde{\omega}^2$. We obtain

$$-(1+\ell)^2 c_1 + (1 + \widetilde{\omega}^2)[c_1 + \varepsilon(c_3 + c_{-1})] = 0, \tag{11.52}$$

$$-(-1+\ell)^2 c_{-1} + (1 + \widetilde{\omega}^2)[c_{-1} + \varepsilon(c_1 + c_{-3})] = 0, \tag{11.53}$$

$$-(3+\ell)^2 c_3 + (1 + \widetilde{\omega}^2)[c_3 + \varepsilon(c_1 + c_5)] = 0, \tag{11.54}$$

$$\vdots$$

The equations (11.54)-... can be solved w.r.t. $c_3, c_{-3}, c_5, \ldots$, if $\ell, \widetilde{\omega}^2, \varepsilon$ are small, i.e., there exist functions

$$c_j = c_j(\ell, \widetilde{\omega}^2, \varepsilon, c_{-1}, c_1) = \mathcal{O}(|\ell| + |\widetilde{\omega}^2| + |\varepsilon|)\mathcal{O}(|c_1| + |c_{-1}|)$$

which are linear w.r.t. c_1 and c_{-1}. Inserting this into (11.52)-(11.53) gives

$$0 = (\mp 2\ell - \ell^2 + \widetilde{\omega}^2)c_{\pm 1} + \varepsilon(1 + \widetilde{\omega}^2)c_{\mp 1} + \varepsilon\mathcal{O}(|\ell| + |\widetilde{\omega}|^2 + |\varepsilon|)\mathcal{O}(|c_1| + |c_{-1}|),$$

or equivalently,

$$\begin{pmatrix} -2\ell - \ell^2 + \widetilde{\omega}^2 + \text{h.o.t.} & (1 + \widetilde{\omega}^2)\varepsilon + \text{h.o.t.} \\ (1 + \widetilde{\omega}^2)\varepsilon + \text{h.o.t.} & 2\ell - \ell^2 + \widetilde{\omega}^2 + \text{h.o.t.} \end{pmatrix} \begin{pmatrix} c_1 \\ c_{-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In order to have non-trivial solutions we need a vanishing determinant. We find

$$(\widetilde{\omega}^2)^2 - \varepsilon^2 + \text{h.o.t.} = 0,$$

i.e., $(\omega^2)_{1/2} = \pm\varepsilon + \text{h.o.t.}$, i.e., there is a splitting of the eigenvalues. In higher space dimensions spectral gaps cannot be obtained by a small perturbation of the spatially homogeneous situation. For examples of spectral gaps in 2D and 3D see [**Kuc93, BFL$^+$07**]. ⌋

11.6.3. Bloch transform. To derive and justify the NLS equation for (11.48) we follow [**BSTU06**] and adapt the Fourier space approach for the constant coefficient problem (11.2) from §11.2 to Bloch space. For Schwartz functions $u \in \mathcal{S}$, the Bloch transform is defined by

$$\widetilde{u}(\ell, x) = (\mathcal{T}u)(\ell, x) = \sum_{j \in \mathbb{Z}} \mathrm{e}^{\mathrm{i}jx} \widehat{u}(\ell + j), \tag{11.55}$$

and its inverse by

$$u(x) = (\mathcal{T}^{-1}\widetilde{u})(x) = \int_{-1/2}^{1/2} \mathrm{e}^{\mathrm{i}\ell x} \widetilde{u}(\ell, x) \, \mathrm{d}\ell. \tag{11.56}$$

By construction we have

$$\widetilde{u}(\ell, x) = \widetilde{u}(\ell, x + 2\pi) \quad \text{and} \quad \widetilde{u}(\ell, x) = \widetilde{u}(\ell + 1, x)\mathrm{e}^{\mathrm{i}x}. \tag{11.57}$$

The Bloch transform turns out to be an isomorphism between $H^\theta(\mathbb{R}, \mathbb{C})$ and $L^2((-1/2, 1/2], H^\theta([0, 2\pi), \mathbb{C}))$, cf. [**RS75b, Sca99**], where

$$\|\widetilde{u}\|_{L^2((-1/2,1/2],H^\theta([0,2\pi),\mathbb{C}))} = \left(\int_{-1/2}^{1/2} \|\widetilde{u}(\ell, \cdot)\|^2_{H^\theta[0,2\pi]} d\ell \right)^{1/2}.$$

Multiplication $u(x)v(x)$ in physical space corresponds in Bloch space to the operation

$$(\widetilde{u} \star \widetilde{v})(\ell, x) = \int_{-1/2}^{1/2} \widetilde{u}(\ell - m, x)\widetilde{v}(m, x) \, \mathrm{d}m, \tag{11.58}$$

where (11.57) has to be used for $|\ell - m| > 1/2$. However, if $\chi : \mathbb{R} \to \mathbb{R}$ is 2π-periodic w.r.t. x, then

$$\mathcal{T}(\chi u)(\ell, x) = \chi(x)(\mathcal{T}u)(\ell, x).$$

Applying the Bloch transform to (11.48) gives

$$\partial_t^2 \widetilde{u}(\ell, x) = -\widetilde{L}(\ell, \partial_x)\widetilde{u}(\ell, x) + \chi_3(x)\widetilde{u}^{\star 3}(\ell, x), \tag{11.59}$$

where the Bloch operators $\widetilde{L}(\ell, \partial_x) : H^2([0, 2\pi)) \to L^2([0, 2\pi))$ are given by

$$\widetilde{L}(\ell, \partial_x)\widetilde{u}(\ell, \cdot)(x) = -\chi_1(x)(\partial_x + \mathrm{i}\ell)^2 \widetilde{u}(\ell, x) + \chi_2(x)\widetilde{u}(\ell, x).$$

According to Lemma 11.6.1 for fixed ℓ these operators are self-adjoint and positive definite in the space $L^2(\chi_1^{-1} \, \mathrm{d}x)$. The induced norm $\| \cdot \|_{L^2(\chi_1^{-1} \, \mathrm{d}x)}$ and the usual L^2-norm are equivalent since $\chi_1(x) \geq \gamma_1 > 0$ for a constant γ_1 independent of x by assumption. Thus, for each fixed ℓ there exists a Schauder base $(f_j(\ell, \cdot))_{j \in \mathbb{N}}$ of $L^2([0, 2\pi))$ of eigenfunctions of $\widetilde{L}(\ell, \partial_x)$ with strictly positive eigenvalues $\lambda_j(\ell) > 0$, i.e., $\widetilde{L}(\ell, \partial_x) f_j(\ell, \cdot) = \lambda_j(\ell) f_j(\ell, \cdot)$. We make the ansatz

$$\widetilde{u}(\ell, x, t) = \sum_{j \in \mathbb{N}} \widetilde{u}_j(\ell, t) f_j(\ell, x).$$

Since $\widetilde{L}(\ell, \partial_x)$ is self-adjoint in $L^2(\chi_1^{-1}\,\mathrm{d}x)$, the eigenfunctions $(f_j(\ell, \cdot))_{j\in\mathbb{N}}$ can be chosen to form an orthonormal basis of $L^2(\chi_1^{-1}\,\mathrm{d}x)$ for each fixed ℓ. Hence,

$$\widetilde{u}_j(\ell, t) = \langle f_j(\ell, \cdot),\, \widetilde{u}(\ell, \cdot, t)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)},$$

and therefore

$$\begin{aligned} \partial_t^2 \widetilde{u}_j(\ell, t) &= -\lambda_j(\ell)\widetilde{u}_j(\ell, t) + \langle f_j(\ell, \cdot),\, \chi_3(\cdot)\widetilde{u}^{\star 3}(\ell, \cdot, t)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)} \\ &= -\lambda_j(\ell)\widetilde{u}_j(\ell, t) + \sum_{j_1, j_2, j_3 \in \mathbb{N}} \int_{-1/2}^{1/2}\int_{-1/2}^{1/2} b^j_{j_1 j_2 j_3}(\ell, \ell-\ell_1, \ell_1-\ell_2, \ell_2) \\ &\qquad \times \widetilde{u}_{j_1}(\ell-\ell_1, t)\widetilde{u}_{j_2}(\ell_1-\ell_2, t)\widetilde{u}_{j_3}(\ell_2, t)\,\mathrm{d}\ell_2\,\mathrm{d}\ell_1, \end{aligned} \tag{11.60}$$

where

$$\begin{aligned} & b^j_{j_1 j_2 j_3}(\ell, \ell-\ell_1, \ell_1-\ell_2, \ell_2) \\ & \quad = \langle f_j(\ell, \cdot), \chi_3(\cdot) f_{j_1}(\ell-\ell_{j_1}, \cdot) f_{j_2}(\ell_{j_1}-\ell_{j_2}, \cdot) f_{j_3}(\ell_{j_2}-, \cdot)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)}. \end{aligned}$$

Since the nonlinear terms have some convolution structure we have a system as in §11.3. Thus, we can proceed exactly as in §11.3 in order to derive a NLS equation. However, due to the special structure of (11.60) we keep the second order system and proceed as in §11.2 and make the ansatz

$$\widetilde{u}_{n_0}(\ell, t) = \varepsilon\varepsilon^{-1}\widetilde{A}_1\left(\frac{\cdot - \ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^1 + \varepsilon\varepsilon^{-1}\widetilde{A}_{-1}\left(\frac{\cdot + \ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-1}$$

for a $n_0 \in \mathbb{N}$ where $\mathbf{E}^j = \mathrm{e}^{j\mathrm{i}\omega_{n_0}(\ell_0)t}\mathrm{e}^{\mathrm{i}\omega'_{n_0}(\ell_0)(\ell - j\ell_0)t}$. In physical space the ansatz corresponds to

$$u(x,t) = \varepsilon\psi_A(x,t) = \varepsilon A(\varepsilon(x+c_{\mathrm{g}}t), \varepsilon^2 t) f_{n_0}(\ell_0, x)\mathrm{e}^{\mathrm{i}\ell_0 x}\mathrm{e}^{\mathrm{i}\omega_{n_0}(\ell_0)t} + \text{c.c.}, \tag{11.61}$$

where again $T = \varepsilon^2 t$, $X = \varepsilon(x + c_{\mathrm{g}}t)$, and $A(X,T) \in \mathbb{C}$. We have a cancelation at $\varepsilon\varepsilon^{-1}\mathbf{E}$ and $\varepsilon^2\varepsilon^{-1}\mathbf{E}$, and at $\varepsilon^3\varepsilon^{-1}\mathbf{E}$ we find the NLS equation

$$\begin{aligned} 2\mathrm{i}\omega_{n_0}(\ell_0)\partial_T \widetilde{A}_1 &= -(\lambda''_{n_0}(\ell_0) - 2(\omega'_{n_0}(\ell_0))^2)\kappa^2\widetilde{A}_1/2 \\ &\quad + \gamma\widetilde{A}_1 * \widetilde{A}_1 * \widetilde{A}_{-1}, \end{aligned} \tag{11.62}$$

where $T = \varepsilon^2 t$, $\kappa = \varepsilon^{-1}(\ell - \ell_0)$, and $\gamma = 3b^{n_0}_{n_0 n_0 n_0}(\ell_0, 0, 2\ell_0, -\ell_0)$, i.e.,

$$\gamma = \frac{3}{2\omega_{n_0}(\ell_0)}\int_0^{2\pi}\frac{\chi_3(x)}{\chi_1(x)}|f_{n_0}(\ell_0, x)|^4\,\mathrm{d}x \in \mathbb{R}, \tag{11.63}$$

while $A_{-1} = \mathcal{F}^{-1}\widetilde{A}_{-1}$ satisfies the complex conjugate equation. In order to obtain (11.62), in the convolution term we use

$$\begin{aligned}&\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)}\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} b^{n_0}_{n_0n_0n_0}(\ell_0+\varepsilon\kappa, \varepsilon(\kappa-\kappa_1), 2\ell_0+\varepsilon(\kappa-\kappa_1), -\ell_0+\varepsilon\kappa_2))\\ &\qquad\qquad \times \widetilde{A}_1(\kappa-\kappa_1)\widetilde{A}_1(\kappa_1-\kappa_2)\widetilde{A}_{-1}(\kappa_2)\,\mathrm{d}\kappa_2\,\mathrm{d}\kappa_1\\ &\to b^{n_0}_{n_0n_0n_0}(\ell_0, 0, 2\ell_0, -\ell_0)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\widetilde{A}_1(\kappa-\kappa_1)\widetilde{A}_1(\kappa_1-\kappa_2)\widetilde{A}_{-1}(\kappa_2)\,\mathrm{d}\kappa_2\,\mathrm{d}\kappa_1\end{aligned}$$

and the symmetry of the kernel. The derivation from (11.62) is consistent with the derivation from the associated first order system since for instance

$$-(\lambda''_{n_0}(\ell_0) - 2(\omega'_{n_0}(\ell_0))^2)/(4\mathrm{i}\omega_{n_0}) = -\left((\omega_{n_0}(\ell_0)^2)'' - 2(\omega'_{n_0}(\ell_0))^2\right)/(4\mathrm{i}\omega_{n_0})$$

which equals $\mathrm{i}\omega''(\ell_0)/2$.

11.6.4. An approximation result. The derivation of the NLS equation, the fact that the NLS equation possesses standing pulses, and the possibility of vanishing group velocities, in principle gives the possibility of standing light pulses. In order to show that these standing light pulses exist at least on an $\mathcal{O}(1/\varepsilon^2)$ time interval we again prove an approximation result.

We consider the following situation. We assume that one of the curves of eigenvalues, namely λ_{n_0}, has a horizontal tangency and that there is no other curve in a neighborhood of this point in the (ℓ,λ)-plane. Hence, $\lambda' = 2\omega\omega' = 0$ in this point and so we have a vanishing group velocity for an associated modulated wave packet. For simplicity we assume that $\lambda_1'(0) = 0$ with $\lambda_1(0) > 0$.

In order to bring together the NLS equation whose solutions in Fourier space are given on the complete real line with the Bloch wave representation of the nonlinear wave equation we introduce a cut-off operator $\chi \in C_0^\infty$ with $\chi(\ell) \in [0,1]$, $\chi(\ell) = 1$ for $\ell \in [-1/5, 1/5]$, and $\chi(\ell) = 0$ for $|\ell| \geq 2/5$, and an extension operator $\mathcal{P}$ which extends a function with length of support less than 1 to a function on the complete real axis with period 1. Then we make the modified ansatz

$$\widetilde{u}_1(\ell,t) = \varepsilon\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_1\left(\frac{\cdot}{\varepsilon},\varepsilon^2 t\right)\mathbf{E}^1)(\ell) + \varepsilon\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_{-1}\left(\frac{\cdot}{\varepsilon},\varepsilon^2 t\right)\mathbf{E}^{-1})(\ell)$$

and find the NLS equation

$$2\mathrm{i}\omega_1(0)\partial_T\widetilde{A}_1 = -(\lambda_1''(0) - 2(\omega_1'(0))^2)\kappa^2\widetilde{A}_1/2 + \gamma\widetilde{A}_1 * \widetilde{A}_1 * \widetilde{A}_{-1}, \tag{11.64}$$

where now $\gamma = 3\widetilde{b}^1_{111}(0,0,0,0) \in \mathbb{R}$, while $A_{-1} = \mathcal{F}^{-1}\widetilde{A}_{-1}$ satisfies again the complex conjugate equation. In case of a cubic nonlinearity for the proof of the approximation result it is not necessary to expand the problem in eigenfunctions, and so we set

$$\widetilde{u}(\ell, x, t) = \widetilde{u}_1(\ell, t) f_1(\ell, x) + \widetilde{u}^{\perp}(\ell, x, t),$$

with $\langle (f_1(\ell, \cdot), \widetilde{u}^{\perp}(\ell, \cdot, t) \rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)} = 0$. Moreover, the separation will be made in such a way that the support of $\widetilde{u}_1(\ell, t)$ is contained in $[-2/5, 2/5]$. The two functions $\widetilde{u}_1(\ell, t)$ and $\widetilde{u}^{\perp}(\ell, x, t)$ are defined to satisfy

$$\begin{aligned}
\partial_t^2 \widetilde{u}_1(\ell, t) &= -\lambda_1(\ell)\widetilde{u}_1(\ell, t) + E_c(\ell)\langle f_1(\ell, \cdot),\, \chi_3(\cdot)\widetilde{u}^{\star 3}(\ell, \cdot, t)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)}, \\
\partial_t^2 \widetilde{u}^{\perp}(\ell, x, t) &= -\widetilde{L}(\ell, \partial_x)\widetilde{u}^{\perp}(\ell, x, t) + \chi_3(x)\widetilde{u}^{\star 3}(\ell, x) \\
&\quad - E_c(\ell)\langle f_1(\ell, \cdot),\, \chi_3(\cdot)\widetilde{u}^{\star 3}(\ell, \cdot, t)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)} f_1(\ell, x),
\end{aligned}$$

where $E_c \in C_0^\infty$, with $E_c(\ell) \in [0, 1]$, with $E_c(\ell) = 1$ for $\ell \in [-1/5, 1/5]$, and $\chi(\ell) = 0$ for $|\ell| \geq 2/5$. Note that χ is defined on the Fourier wave numbers, where E_c is defined on the Bloch wave numbers.

We add higher order terms to the ansatz to make the residual smaller, i.e., we consider

$$\begin{aligned}
\widetilde{u}_1(\ell, t) =& \varepsilon\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_1\left(\frac{\cdot}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^1)(\ell) + \varepsilon\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_{-1}\left(\frac{\cdot}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-1})(\ell) \\
&+ \varepsilon^3\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_3\left(\frac{\cdot}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^3)(\ell) + \varepsilon^3\varepsilon^{-1}\mathcal{P}(\chi(\cdot)\widetilde{A}_{-3}\left(\frac{\cdot}{\varepsilon}, \varepsilon^3 t\right)\mathbf{E}^{-3})(\ell), \\
\widetilde{u}^{\perp}(\ell, x) =& \varepsilon^3\varepsilon^{-1}\widetilde{u}_1^{\perp}(\frac{\ell}{\varepsilon}, x, \varepsilon^2 t)\mathbf{E}^1 + \varepsilon^3\varepsilon^{-1}\widetilde{u}_{-1}^{\perp}(\frac{\ell}{\varepsilon}, x, \varepsilon^2 t)\mathbf{E}^{-1} \\
&+ \varepsilon^3\varepsilon^{-1}\widetilde{u}_3^{\perp}(\frac{\ell}{\varepsilon}, x, \varepsilon^2 t)\mathbf{E}^3 + \varepsilon^3\varepsilon^{-1}\widetilde{u}_{-3}^{\perp}(\frac{\ell}{\varepsilon}, x, \varepsilon^2 t)\mathbf{E}^{-3}.
\end{aligned}$$

As before we find $\widetilde{A}_1$ as a solution of the NLS equation (11.64) and $\widetilde{A}_{-1}$ as a solution of the complex conjugate equation. Moreover we choose

$$\begin{aligned}
&-9\omega_1^2(0)\widetilde{A}_3(\kappa, T) \\
&\qquad = -\lambda_1(0)\widetilde{A}_3(\kappa, T) + \langle f_1(0, \cdot),\, \chi_3(\cdot)(f_1(0, \cdot))^3\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)}\widetilde{A}_1^{*3}(\kappa, T), \\
&-\omega_1^2(0)\widetilde{u}_1^{\perp}(\ell, x, t) \\
&\qquad = -\widetilde{L}(\varepsilon\kappa, \partial_x)\widetilde{u}_1^{\perp}(\kappa, x, T) + 3\chi_3(x)(\widetilde{A}_1 f_1)^{*2} * \widetilde{A}_{-1} f_{-1})(\kappa, x) \\
&\qquad\quad - E_c(\ell)\langle f_1(\varepsilon\kappa, \cdot),\, 3\chi_3(\cdot)(\widetilde{A}_1 f_1)^{*2} * \widetilde{A}_{-1} f_{-1})(\kappa, \cdot, T)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)} f_1(\varepsilon\kappa, x), \\
&-9\omega_1^2(0)\widetilde{u}_3^{\perp}(\ell, x, t) \\
&\qquad = -\widetilde{L}(\varepsilon\kappa, \partial_x)\widetilde{u}_3^{\perp}(\kappa, x, T) + \chi_3(x)(\widetilde{A}_1 f_1 * \widetilde{A}_1 f_1 * \widetilde{A}_1 f_1)(\kappa, x) \\
&\qquad\quad - E_c(\ell)\langle f_1(\varepsilon\kappa, \cdot),\, \chi_3(\cdot)(\widetilde{A}_1 f_1)^{*3}(\kappa, \cdot, T)\rangle_{L^2(\chi_1^{-1}\,\mathrm{d}x)} f_1(\varepsilon\kappa, x),
\end{aligned}$$

and $\widetilde{A}_{-3}$, $\widetilde{u}_{-1}^{\perp}$, and $\widetilde{u}_{-3}^{\perp}$ as the solutions of the associated complex conjugate equations. By this choice we eliminate all nonlinear terms w.r.t. $\widetilde{A}_1$ and $\widetilde{A}_{-1}$ in the part of the residual belonging to the $u^{\perp}$ equation. This has the advantage that we do not need more information about the spectrum and the associated eigenfunctions, especially we avoid to estimate derivatives of the eigenfunctions w.r.t. ℓ.

In order to solve the equations for $\widetilde{A}_3$, $\widetilde{u}_1^\perp$, and $\widetilde{u}_3^\perp$ we need the non-resonance conditions

$$\begin{aligned} -9\omega_1^2(0) &\neq -\lambda_1(0), \\ -\omega_1^2(0) &\notin \operatorname{spec}(-\widetilde{L}(\varepsilon\kappa, \partial_x))|_{\{f_1(\varepsilon\kappa,\cdot)\}^\perp}, \\ -9\omega_1^2(0) &\notin \operatorname{spec}(-\widetilde{L}(\varepsilon\kappa, \partial_x))|_{\{f_1(\varepsilon\kappa,\cdot)\}^\perp}. \end{aligned}$$

If these conditions are not satisfied for $\ell = \varepsilon\kappa$ for which $E_c(\ell) = 1$ we can make the support of E_c smaller to satisfy the non-resonance conditions. This is possible if

$$\begin{aligned} -\omega_1^2(0) &\notin \operatorname{spec}(-\widetilde{L}(0, \partial_x))|_{\{f_1(0,\cdot)\}^\perp}, \\ -9\omega_1^2(0) &\notin \operatorname{spec}(-\widetilde{L}(0, \partial_x))|_{\{f_1(0,\cdot)\}^\perp}. \end{aligned}$$

These formal calculations can be made rigorous, as in §11.2 it can be shown that the residual is of order $\mathcal{O}(\varepsilon^{7/2})$ in H^1 in physical space.

The error estimates can be proved in physical space. Let $\varepsilon\Psi$ be the approximation in physical space satisfying

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\operatorname{Res}(\varepsilon\Psi)\|_{H^1} \leq C\varepsilon^{7/2}$$

where

$$\operatorname{Res}(u) = -\partial_t^2 u + \chi_1\partial_x^2 u - \chi_2 u + \chi_3 u^3.$$

The error $\varepsilon^{3/2}R = u - \varepsilon\psi$ satisfies

$$\partial_t^2 R = \chi_1\partial_x^2 R - \chi_2 R + 3\varepsilon^2\chi_3\psi^2 R + 3\varepsilon^{5/2}\chi_3\psi R^2 + \varepsilon^3\chi_3 R^3 + \varepsilon^{-3/2}\operatorname{Res}(\varepsilon\Psi).$$

Multiplying this equation with $\chi_1^{-1}\partial_t R$ and performing the integration $\int_{-\infty}^{\infty} \dots dx$ yields

$$\partial_t E \leq C_1\varepsilon^2 E + C_2\varepsilon^{5/2}E^{3/2} + C_3\varepsilon^3 E^2 + C_4\varepsilon^2,$$

with ε-independent constants $C_1, \dots, C_4$ and where

$$E = \int_{-\infty}^{\infty} \chi_1^{-1}(\partial_t R)^2 + (\partial_x R)^2 + \chi_1^{-1}\chi_2 R^2 dx.$$

A simple application of Gronwall's inequality yields as in §11.2 the following approximation theorem.

Theorem 11.6.6. *Let the Fourier transform of $A \in C([0,T_0],H^4)$ be a solution of the NLS equation* (11.64). *Then there exists an $\varepsilon_0 > 0$ and a $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ there are solutions u of the original system* (11.48) *which can be approximated by $\varepsilon\Psi$ w.r.t. the H^1-norm such that*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(t) - \varepsilon\Psi(t)\|_{H^1} < C\varepsilon^{3/2}.$$

If the NLS equation (11.64) is a focussing one, then it possesses standing time-periodic pulse solutions, see §8.1.1. This in combination with Theorem 11.6.6 shows the existence of approximate standing light pulses on an $\mathcal{O}(1/\varepsilon^2)$ time scale. Since we have a finite speed of propagation for (11.48) with the help of spatial dynamics and invariant manifold theory, cf. Chapter 13, the existence of the standing modulating pulses can even be shown on $\mathcal{O}(1/\varepsilon^n)$ time scales for fixed $n > 2$ if certain non-resonance conditions are satisfied [**LBCB$^+$09**]. If the coefficients χ_1 and χ_2 are chosen in a special way, then standing modulating pulses can be shown [**BCBLS11**] to exist for all $t \in \mathbb{R}$.

11.6.5. Gap solitons. Another popular class of models for light propagation in photonic crystals starts directly with a Nonlinear Schrödinger (or Gross–Pitaevsky) equation in the form

$$\mathrm{i}E_t = -\Delta E + V(x)E + \sigma|E|^2 E, \quad V : \mathbb{R}^2 \to \mathbb{R},\ x \in \mathbb{R}^2,\ t \in \mathbb{R}, \tag{11.65}$$

with $E = E(x,t) \in \mathbb{C}$, $\sigma = \pm 1$, and where w.l.o.g. the potential V is 2π–periodic in each coordinate. Here E is the “out–of–plane” electric field, i.e., perpendicular to the 2D periodicity structure, and the amplitude of V is called the contrast of the material. An interesting problem then is to search for so called gap solitons $E(x,t) = \phi(x)\mathrm{e}^{-\mathrm{i}\omega t}$, where $\phi \in \mathbb{C}$ is a localized solution of the stationary problem

$$(-\Delta + V(x) - \omega)\phi + \sigma|\phi|^2\phi = 0. \tag{11.66}$$

Localized means $|\phi(x)| \to 0$ exponentially as $|x| \to 0$, and this implies that ω has to lie in a gap of the essential spectrum of the operator $L := -\Delta + V(x)$, hence the name “gap soliton”. These have been discussed the physics literature since the early 1990ties, see, e.g., [**Ace00**] for a review. In 1D, where gaps open for small contrasts as discussed above, they are typically described by so called coupled mode equations, cf. Exercise 11.8. In two and more space dimensions gaps are more difficult to open [**Kuc93**], and in particular one needs a finite contrast. Rigorous proofs and asymptotics for 2D gap solitons, based on the reduction of (11.65) to systems of homogeneous NLS equations, here also called coupled mode equations, can be found for instance in [**DPS09**] (for the case of a separable potential) and in [**DU09**] (for general V).

11.7. Nonlinear Optics

It is the purpose of this section to explain how the previous analysis is related to the motivation given at the beginning of this Chapter, namely the transport of information by light pulses through glass fibers. Moreover, we explain why the rate of information transported through the fiber can be increased by multiplexing, i.e., by taking simultaneously pulses with different carrier waves.

11.7.1. Maxwell's equations in glass fibers. Electromagnetic waves are described by Maxwell's equations. For $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $t \in \mathbb{R}$ they are given by

$$\nabla \cdot E = \rho, \tag{11.67}$$

$$\nabla \times E = -\partial_t B, \tag{11.68}$$

$$\nabla \times B = \partial_t E + j, \tag{11.69}$$

$$\nabla \cdot B = 0, \tag{11.70}$$

where $E = E(x,t) \in \mathbb{R}^3$ is the electric field, $B = B(x,t) \in \mathbb{R}^3$ the magnetic field, $j = j(x,t) \in \mathbb{R}^3$ an electric current density, and $\rho = \rho(x,t) \in \mathbb{R}$ an electric charge density. There is conservation of charges, i.e., additionally we have $\partial_t \rho + \nabla \cdot j = 0$. There are no magnetic monopoles, i.e., (11.67)-(11.70) are not symmetric, since there is no magnetic current density and no magnetic charge density. We put the vacuum velocity of light and the vacuum electric permittivity to one.

Remark 11.7.1. For some purposes, alternative formulations of Maxwell's equations are useful. For time-independent solutions the system decouples into electrostatics

$$\nabla \cdot E = \rho, \quad \nabla \times E = 0$$

and magnetostatics

$$\nabla \times B = j, \quad \nabla \cdot B = 0.$$

Hence, there exist a scalar potential ϕ with $E = \nabla\phi$ satisfying $\Delta\phi = \rho$ and a vector potential A with $B = \nabla \times A$ satisfying $\nabla \times \nabla \times A = j$.

The integral formulation of Maxwell's equations requires less regularity. For a volume V with surface ∂V we find

$$\int_{\partial V} E \cdot n = \int_V \rho \quad \text{and} \quad \int_{\partial V} B \cdot n = 0.$$

For a surface S with boundary ∂S we find

$$\int_{\partial S} \langle E, \mathrm{d}s \rangle = -\int_S \partial_t B \quad \text{and} \quad \int_{\partial S} \langle B, \mathrm{d}s \rangle = \int_S \partial_t E + j. \qquad \rfloor$$

Introducing the polarization. For given initial conditions $E|_{t=0}$ and $B|_{t=0}$ and given external densities $\rho = \rho(x,t)$ and $j = j(x,t)$ the solutions $E = E(x,t)$ and $B = B(x,t)$ of Maxwell's equations can be computed. In a medium, the electric field E and the magnetic field B affect the motion of the charged particles in the medium leading to another electric and magnetic field, namely the polarization and the magnetization. Glass fibers and photonic crystals are insulators, i.e., there are no free charge carriers. Moreover, there is no magnetization. However, the applied electric field yields to a shift of the electron density, i.e., to a charge density distribution ρ_{pol} in the atoms. Due to the conservation of charges there is an induced current density j_{pol} satisfying

$$\partial_t \rho_{\text{pol}} + \nabla \cdot j_{\text{pol}} = 0.$$

These induced densities on the other hand induce via (11.67) and (11.69) an electric field $P = P(x,t)$, which is called the polarization such that

$$\rho_{\text{pol}} = -\nabla \cdot P \quad \text{and} \quad \nabla \cdot j_{\text{pol}} = -\partial_t \rho_{\text{pol}} = \partial_t \nabla \cdot P,$$

leading to $\partial_t P = j_{\text{pol}}$. Thus, Maxwell's equations in glass fibers and photonic crystals are given by

$$\nabla \cdot E = -\nabla \cdot P, \tag{11.71}$$

$$\nabla \times E = -\partial_t B, \tag{11.72}$$

$$\nabla \times B = \partial_t E + \partial_t P, \tag{11.73}$$

$$\nabla \cdot B = 0. \tag{11.74}$$

This system can be simplified to a single equation. Applying $\nabla\times$ to (11.72) yields

$$\nabla \times (\nabla \times E) = -\partial_t (\nabla \times B).$$

Using

$$\nabla \times (\nabla \times E) = \nabla(\nabla \cdot E) - \Delta E$$

and (11.73) yields

$$\nabla(\nabla \cdot E) - \Delta E = -\partial_t^2 P - \partial_t^2 E. \tag{11.75}$$

For linearly polarized light, e.g. $E = u(x_3,t)\vec{e}_{x_1}$ and $P = p(x_3,t)\vec{e}_{x_1}$ we find $\nabla \cdot P = 0$, $\nabla \cdot E = 0$, and so

$$\partial_{x_3}^2 u = \partial_t^2 u + \partial_t^2 p, \tag{11.76}$$

where $u(x_3,t) \in \mathbb{R}$ and $p(x_3,t) \in \mathbb{R}$.

Modeling the polarization. In order to have an evolutionary problem we have to close (11.75) or (11.76) by a constitutive law $P = P(E)$ or $p = p(u)$, respectively. There exist various models. Basically the law of motion of a particle with coordinates $x = x(t)$ of mass m and charge q is given by

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}x = q(E + \frac{\mathrm{d}x}{\mathrm{d}t} \times B).$$

In the simplest model the polarization is modeled as an oscillator and the influence of the magnetic field B is neglected. Thus we suppose that for an atom placed in an electric field the center $x = x(t)$ of the electron density obeys the equation

$$m\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}x + \omega_0^2 x\right) = q_{\mathrm{e}}E,$$

with q_{e} the elementary charge, ω_0 some normalized temporal wave number, and m the mass of the electron. On the continuum level in some mean-field limit for the electron position we find for the polarization

$$\partial_t^2 P + \omega_0^2 P = dE, \tag{11.77}$$

with a constant $d \in \mathbb{R}$. In case of damping by thermalization we have

$$m\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}x + \gamma\frac{\mathrm{d}}{\mathrm{d}t}x + \omega_0^2 x\right) = q_{\mathrm{e}}E$$

for a $\gamma > 0$. On the continuum level we find

$$\partial_t^2 P + \gamma\partial_t P + \omega_0^2 P = dE. \tag{11.78}$$

For larger values of E the linear oscillator has to be replaced by a nonlinear oscillator. Due to the symmetries of the problem very often there are no quadratic terms. As a nonlinear example we consider the oscillator model

$$m\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}x + \gamma\frac{\mathrm{d}}{\mathrm{d}t}x + \omega_0^2 x + r|x|^2 x\right) = q_{\mathrm{e}}E$$

for an $r \in \mathbb{R}$. On the continuum level we find

$$\partial_t^2 P + \gamma\partial_t P + \omega_0^2 P + rP|P|^2 = dE. \tag{11.79}$$

In general, matter consists of different atoms which combine to various molecules. This is modeled by considering various kind of oscillators. Hence, we finally come to a system

$$\nabla(\nabla\cdot E) - \Delta E = -\partial_t^2 P - \partial_t^2 E, \tag{11.80}$$

$$P = \sum_{j=0}^{n} P_j, \tag{11.81}$$

$$\partial_t^2 P_j + \gamma_j\partial_t P_j + \omega_j^2 P_j + r_j|P_j|^2 P_j = d_j E, \tag{11.82}$$

where the $d_j > 0$ are constants taking into account different masses, different numbers of atoms, etc., of the various kinds of atoms. In order to solve this system uniquely we need initial conditions for $E|_{t=0}, \partial_t E|_{t=0}, P_j|_{t=0}$ and $\partial_t P_j|_{t=0}$ for $j = 0, \ldots, n$. We refer to [**SU03b**] for a mathematical analysis of light pulse propagation in glass fibers with this modeling in case of damping, i.e., $\gamma_j > 0$.

Remark 11.7.2. a) (11.80)-(11.82) models isotropic media. Anisotropic media can be modeled for instance by choosing γ_j, ω_j^2, d_j, and r_j as tensors.

b) In the linear case $r_j = 0$ in (11.82), the constitutive laws can be solved explicitly. For $E = E_0 e^{i\omega t}$ we find $P = P_0 e^{i\omega t} + \widetilde{P}(t)$, with $\widetilde{P}(t) \to 0$ with some exponential rate for $t \to \infty$ and

$$P_0 = \alpha(\omega) E_0$$

where

$$\alpha(\omega) = \sum_{j=0}^{n} \frac{d_j}{-\omega^2 + i\gamma_j \omega + \omega_j^2}.$$

The numbers γ_j, d_j, and ω_j can be used to fit the constitutive law to experimental data. There is a remarkable good agreement with experimental observations.

c) More generally, in the linear case there always is a Green's function χ_1 such that P can be expressed in terms of E, i.e.,

$$P(t) = \int_{-\infty}^{t} \chi_1(t-\tau) E(\tau) \, d\tau. \tag{11.83}$$

In isotropic materials χ_1 is scalar. In glass fibers the above constants and also χ_1 only depend on the transverse variables, in photonic crystals there is a periodic dependence on the spatial variables. Although there is no Green's function for the nonlinear system, similar to the linear situation the constitutive law for the polarization is very often modeled by

$$P(t) = \int_{-\infty}^{t} \chi_1(t-\tau) E(\tau) \, d\tau + \int_{-\infty}^{t} \chi_3(t-\tau) |E(\tau)|^2 E(\tau) \, d\tau, \tag{11.84}$$

i.e., again (11.81)-(11.82) is replaced by (11.84).

d) In nonlinear optics very often time and space are interchanged. Due to the finite size of the fibers and due to the experimental data which can be measured initial conditions are posed at one end of the fiber, namely at $x = 0$, and one is interested in the solution at the end of the fiber, namely at $x = x_e$, i.e., x is considered as evolutionary variable, and t as unbounded variable. From a mathematical point of view there is no difference if no dissipation is considered. However, very often phenomenologically dissipation is added to the NLS equation. It has been explained in [**SU03b**] that then it is highly problematic if x and t is interchanged. ⌋

In a one-dimensional optical fiber without damping and one kind of oscillators in the material we have the so called Maxwell-Lorentz system

$$\begin{aligned} \partial_x^2 u &= \partial_t^2 u + \partial_t^2 p, \\ \partial_t^2 p + \omega_0^2 p - r|p|^2 p &= du, \end{aligned}$$

with coefficients ω_0, r, and d. The linearized problem

$$\partial_x^2 u = \partial_t^2 u + \partial_t^2 p, \qquad \partial_t^2 p + \omega_0^2 p = du$$

possesses solutions

$$u(x,t) = u_k e^{i(kx+\omega t)}, \quad p(x,t) = p_k e^{i(kx+\omega t)}$$

which yields

$$\begin{pmatrix} -k^2+\omega^2 & \omega^2 \\ -d & \omega_0^2-\omega^2 \end{pmatrix} \begin{pmatrix} u_k \\ p_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We have non-trivial solutions if the determinant vanishes, i.e., if

$$(-k^2+\omega^2)(\omega_0^2-\omega^2) + d\omega^2 = 0.$$

We find four curves of solutions $\omega = \omega_{1,2,3,4}(k)$ which are sketched in Figure 11.8. The Fourier transformed system can be written as first order system and then diagonalized, leading to

$$\begin{aligned} \partial_t \widehat{c}_1(k,t) &= i\omega_1(k)\widehat{c}_1(k,t) + \text{ nonlinear terms }, \\ &\vdots \\ \partial_t \widehat{c}_4(k,t) &= i\omega_4(k)\widehat{c}_4(k,t) + \text{ nonlinear terms }, \end{aligned}$$

such that the Maxwell-Lorentz system falls in the abstract class of systems considered in §11.3 for which an NLS equation can be derived.

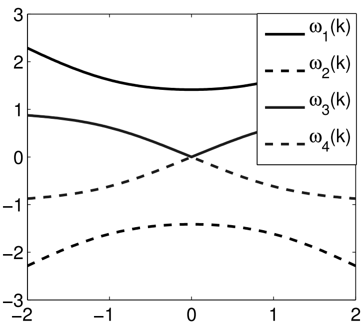

Figure 11.8. The curves of eigenvalues for the Maxwell-Lorentz system

Remark 11.7.3. Other models for the polarization are in use. The choice $\partial_t^2 p = -u - u^3$ leads to the Klein-Gordon model (11.2) for which we justified the NLS approximation in §11.2. Hence, the pulse dynamics present in the NLS equation is present in the models used for the description of the propagation of light pulses in glass fibers, too. Experimental observations confirm this approximation and modeling of reality. ⌋

11.7.2. Multiplexing. So far we considered pulses modulating one carrier wave with some basic wave number k_0. Now we use two or more carrier waves with different basic wave numbers $k_1, \ldots, k_N$. It turns out that pulses belonging to different carrier waves do not interact in lowest order, and thus the use of more than one carrier wave allows to increase the rate of information through the fiber. This concept is called multiplexing. In the following we explain the underlying ideas with the help of the NLS approximation.

Let us consider here again the nonlinear wave equation with cubic nonlinearity

$$\partial_t^2 u = \partial_x^2 u - u - u^3$$

as original system. For the situation of N different carrier waves we make the ansatz

$$\varepsilon\psi_{\mathrm{multiNLS}} = \sum_{j=1}^{N} \varepsilon A_j \left(\varepsilon(x - c_j t), \varepsilon^2 t\right) \mathrm{e}^{\mathrm{i}(k_j x + \omega_j t)} + \mathrm{c.c.}.$$

A system of coupled NLS equations is obtained, namely

$$2\mathrm{i}\omega_j \partial_T A_j = (1 - c_j^2)\partial_{X_j}^2 A_j - 3A_j |A_j|^2 - \text{coupling terms}_j$$

for $j = 1, \ldots, N$, where the jth coupling term is given by

$$6 \sum_{|n|=1,\ldots,N, n\neq j} A_j |A_n|^2 ,$$

where we assumed for simplicity that there are no resonant wave numbers $k_{j_1}, \ldots, k_{j_4}$ with $k_{j_1} + \ldots + k_{j_4} = 0$ and $\omega_{j_1} + \ldots + \omega_{j_4} = 0$. If this assumption is not satisfied there will be additional coupling terms $A_{j_2}A_{j_3}A_{j_4}$ in the equation for A_{-j_1} which however can be handled as explained below, too.

At a first view there seems to be a full coupling between all equations. However, looking more closely at the coupling terms shows that they have different arguments if $c_j \neq c_n$. We have for example

$$\begin{aligned} A_j |A_n|^2 &= A_j \left(\varepsilon(x - c_j t), \varepsilon^2 t\right) \left|A_n \left(\varepsilon(x - c_n t), \varepsilon^2 t\right)\right|^2 \\ &= A_j(X_j, T) \left|A_n \left(X_j - \varepsilon^{-1}(c_n - c_j)T, T\right)\right|^2 . \end{aligned}$$

Hence, two spatially localized functions interact only on an $\mathcal{O}(\varepsilon)$-time interval w.r.t. the T-time scale of the NLS equation if $c_j \neq c_n$. Thus, the influence of the coupling terms on the dynamics of the NLS equations is only $\mathcal{O}(\varepsilon)$. Hence, for spatially localized solutions the NLS equations decouple and the dynamics of the modulations of the carrier waves can be computed for each carrier wave individually by solving

$$2\mathrm{i}\omega_j \partial_T A_j = (1 - c_j^2)\partial_{X_j}^2 A_j - 3A_j |A_j|^2 .$$

The argument that spatially localized wave packets with different group velocities do not interact in lowest order has been made rigorous in case of the above NLS approximation first in [**PW96**]. The idea has been generalized in [**BF06**] for the interaction of wave packets in various original systems. It has also been used in case of the interaction of counter-propagating long waves [**Kal89, SW00b, SW02**], cf. §12.1. In case of dissipative systems mean-field coupled GL equations take the role of the NLS equation [**Sch97**].

The interaction of pulse solutions to different carrier waves can be described very precisely such that [**PW96**] and [**BF06**] can be improved strongly. The more detailed description is

$$\varepsilon\psi_{\mathrm{multiNLS}} = \sum_{j=1}^{N} \varepsilon A_j \left(\varepsilon(x - c_j t - \varepsilon\psi_j(x,t)), \varepsilon^2 t\right) \mathrm{e}^{\mathrm{i}(k_j x + \omega_j t + \varepsilon\Omega_j(x,t))} + \text{c.c.},$$

where the A_j satisfy decoupled equations and there are explicit formulae [**CBSU07**] for the pulse shifts $\varepsilon\psi_j$ and the phase shifts $\varepsilon\Omega_j$. The internal dynamics of the wave packets (described via the A_j) and the interaction dynamics of the wave packets (described via ψ_j and Ω_j) can be separated up to very high order. An almost complete description of the interaction of general localized NLS described wave packets can be found for the nonlinear wave equation in [**CBCSU08**] and in [**CBS12**] for general dispersive wave systems. As a consequence the shift of the underlying carrier wave and the shift of the envelope both can be shown to be of order $\mathcal{O}(\varepsilon)$ instead of $\mathcal{O}(1)$ w.r.t. the original x-variable. See also [**SUW11**] for an application to oscillator chains.

Further Reading. The description of waves in nonlinear optics by modulation equations is a very active field of research. Two main topics are: ultra-short pulses [**SW04, CdR15, PS13, New16**], see also Exercise 11.9, and the gap solitons from §11.6.5 and related phenomena. Besides gap solitons, models of type (11.65) support other solution families, for instance so called Nonlinear Bloch Modes and "Out–of-Gap" solitons, see, e.g., [**Yan10**] for a comprehensive overview, or [**DU16**] for recent rigorous results. See also [**DD13**] for related results for (simplified) nonlinear Maxwell's equations.

Moreover, the (numerical) analysis of 3D photonic crystals, and in particular their design to achieve favorable band structures are very active fields [**BFL$^+$07, STE$^+$10**]. Finally, an analysis somewhat similar to 1D and 2D photonic crystals deals with wave propagation on infinite periodic metric graphs, used as models for nanotubes or graphen, see, e.g. [**GPSar**].

NLS limits occur as singular limits of other dipersive PDE systems as well. Examples are the Zakharov system [**AA88**] or the Klein-Gordon-Zakharov system [**MN05**].

Exercises

11.1. Consider the nonlinear wave equation $\partial_t^2 u = \partial_x^2 u - u - u^3$, with $x, t, u(x,t) \in \mathbb{R}$. Make an ansatz $u(x,t) = \varepsilon A(\varepsilon x, \varepsilon^2 t)\mathrm{e}^{\mathrm{i}t} + \text{c.c.}$ to derive a NLS equation for A. Estimate the residual and prove an approximation result by considering the energy

$$E = \int_{\mathbb{R}} u^2 + (\partial_x u)^2 + (\partial_t u)^2 \,\mathrm{d}x.$$

11.2. Consider the nonlinear wave equation $\partial_t^2 u = \Delta u - u - u^3$, with $x, y, t, u(x,y,t) \in \mathbb{R}$. Make an ansatz

$$u(x,t) = \varepsilon A(\varepsilon(x - c_0 t), \varepsilon y, \varepsilon^2 t)\mathrm{e}^{\mathrm{i}(k_0 x + \omega_0 t)} + \text{c.c.}$$

to derive a 2D NLS equation for A.

11.3. Consider the lattice differential equation

$$\partial_t^2 u_n = u_{n+1} - 2u_n + u_{n-1} - 3u_n - u_n^2,$$

with $u_n = u_n(t) \in \mathbb{R}$ for $n \in \mathbb{Z}$.

a) Find solutions $u_n(t) = \mathrm{e}^{\mathrm{i}(kn - \omega t)}$ of the linearized problem. Compute the dispersion relation.

b) Check the validity of the non-resonance condition.

c) Make the ansatz $u_n(t) = \varepsilon A(\varepsilon n, \varepsilon^2 t)\mathrm{e}^{\mathrm{i}\omega_0 t} + c.c. + \mathcal{O}(\varepsilon^2)$ and derive a NLS equation for the amplitude $A = A(X,T)$.

11.4. Consider the FPU system

$$\partial_t^2 u_n = W'(u_{n+1}) - 2W'(u_n) + W'(u_{n-1}),$$

with W an analytic function, i.e., $W'(u) = au + bu^2 + \ldots$. Derive an evolution equation for $u(k,t) = \sum_{n\in\mathbb{Z}} u_n(t)\mathrm{e}^{\mathrm{i}kn}$, cf. [**Mil06**]. Find differences and similarities between this evolution equation and the Fourier transform of a nonlinear dispersive PDE.

11.5. Consider the KdV equation $\partial_t u - 6u\partial_x u + \partial_x^3 u = 0$.

a) Apply the Miura transformation $u = v^2 + \partial_x v$ and derive the mKdV equation

$$\partial_t v - 6v^2 \partial_x v + \partial_x^3 v = 0.$$

b) Plug in the ansatz $\varepsilon\psi_v(x,t) = \varepsilon A(\varepsilon(x - ct), \varepsilon^2 t)\mathrm{e}^{\mathrm{i}(kx - +\omega t)} + c.c.$ into the mKdV equation and derive a NLS equation for A.

c) Justify this approximation by energy estimates and conclude the validity of an approximation theorem for the approximation of the KdV equation by the NLS equation [**Sch11**].

11.6. The dispersion relation of the water wave problem with surface tension is given by

$$\omega^2 = (k + \sigma k^3)\tanh(k),$$

with surface tension parameter $\sigma \geq 0$. Show that for $\sigma \in (0, 1/3)$ beside $k = 0$ and $k = k_0$ there are two additional resonances k_1 and k_2 with $k_0 + k_1 + k_2 = 0$ and $\omega_0 + \omega_1 + \omega_2 = 0$.

11.7. Solve the eigenvalue problem $\partial_x u(x) = v$, $\partial_x v(x) = -s(x)\lambda u(x)$, with the 1-periodic function $s(x) = \chi_{[0,6/13]} + 16\chi_{(6/13,7/13)} + \chi_{[7/13,1]}(x \bmod 1)$. Compute the monodromy matrix C_λ and the discriminant $D(\lambda) = \operatorname{trace} C_\lambda$. Hint. On intervals where s is constant, the problem can be solved explicitly. We look for $u \in C_b^1$.

11.8. Consider the nonlinear wave equation $\partial_t^2 u = \partial_x^2 u + u + 2\varepsilon^2 \cos(2x)u - u^3$, with spatially periodic perturbed coefficients, i.e., $0 \leq \varepsilon \ll 1$. For $\varepsilon > 0$ a spectral gap of order $\mathcal{O}(\varepsilon^2)$ occurs which is too small to derive an NLS equation. In this case with the ansatz

$$u(x,t) = \varepsilon a(\varepsilon^2 x, \varepsilon^2 t)\mathrm{e}^{\mathrm{i}(x-\omega_0 t)} + \varepsilon b(\varepsilon^2 x, \varepsilon^2 t)\mathrm{e}^{-\mathrm{i}(x+\omega_0 t)} + \mathrm{c.c.}$$

derive the coupled mode system, cf. [**SU01**]

$$\begin{aligned} -2\mathrm{i}\omega_0 \partial_T a &= 2\mathrm{i}\partial_x a + \mathrm{i}b - 3a|a|^2 - 6a|b|^2, \\ -2\mathrm{i}\omega_0 \partial_T b &= -2\mathrm{i}\partial_x b + \mathrm{i}a - 3b|b|^2 - 6b|a|^2. \end{aligned}$$

11.9. If the pulses become very narrow, then the so called short pulse equation

$$\partial_\xi \partial_\tau A = A + \partial_\xi^2 (A^3),$$

can be derived. Consider the quasilinear wave equation $\partial_t^2 u = \partial_x^2 u + u + \partial_x^2(u^3)$, and make the ansatz

$$u(t,x) = 2\varepsilon A(\tau,\xi), \quad \tau = \varepsilon t, \quad \xi = \frac{x-t}{2\varepsilon}$$

to derive the short pulse equation, cf. [**PS13**].

11.10. Show that the inhomogeneous Maxwell equations

$$\partial_t E = \nabla \times B - J, \quad \nabla \cdot E = \rho, \quad \partial_t B = -\nabla \times E, \quad \nabla \cdot B = 0,$$

can be transformed to the inhomogeneous wave equations

$$\partial_t^2 u - \Delta u = \rho, \quad \partial_t^2 A - \Delta A = J,$$

where A is a vector potential of B, i.e., $\nabla \times A = B$, and u is a scalar potential of E, i.e., $E + \partial_t A = -\nabla u$, see Remark 11.7.1.
Hint. Consider a suitable gauge transform, i.e., adding $\nabla\lambda$ to A and subtracting $-\partial_t \lambda$ from u.