Justification of the Nonlinear Schrödinger equation in spatially periodic media

Kurt Busch[1] [2], Guido Schneider [3], Lasha Tkeshelashvili[1] [4], Hannes Uecker[3]

October 12, 2005

Abstract

The dynamics of the envelopes of spatially and temporarily oscillating wave packets advancing in spatially periodic media can approximately be described by solutions of a Nonlinear Schrödinger equation. Here we prove estimates for the error made by this formal approximation using Bloch wave analysis, normal form transformations, and Gronwall's inequality.

Contents

[1]Department of Physics and College of Optics & Photonics: CREOL & FPCE, University of Central Florida, Orlando FL 32816, USA

[2]new adress: Institut für Theoretische Festkörperphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany, kurt@tfp.physik.uni-karlsruhe.de

[3]Mathematisches Institut I, D-76128 Universität Karlsruhe, guido.schneider@math.uni-karlsruhe.de hannes.uecker@math.uni-karlsruhe.de

[4]Institut für Nanotechnologie, Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft, Postfach 3640, D-76021 Karlsruhe, Germany, lasha@tfp.uni-karlsruhe.de

1   Introduction

Partial differential equations represent the main mathematical tool for researchers in many diverse fields of science and technology. Since, in general, the laws of nature are nonlinear, the corresponding governing equations of motion are nonlinear as well. Moreover, the equations often involve multiple scales. In these cases approximate methods for dealing with nonlinear partial differential equations become very important. Through appropriate schemes, one can reduce the original equations to simpler ones which describe the behaviour on the large scales. In fact, this approach allows one to reduce the description of vastly different physical phenomena to only a few universal nonlinear partial differential equations, for which, moreover, often a method of integration is known.

There is, however, an important issue concerning the solutions of the reduced nonlinear partial differential equations. This is the problem of the structural stability of the mathematical models which stems from the approximations that have been made (i) in the derivation of the equations of motion and (ii) in their reduction to an integrable mathematical model. The neglected terms, even if they remain small, may significantly affect the behavior of the solutions on different time and length scales. Therefore, it is not guaranteed that the solutions of the reduced model are of any relevance for the original physical problem [FW98]. As a result, it becomes very important to estimate the error introduced by these approximations and to accurately determine the time and length scales for which the reduced model remains valid.

In this paper, we discuss one of the most widely studied universal mathematical models, the Nonlinear Schr¨odinger (NLS) equation and provide answers to the question of its validity in the context of spatially periodic media. The NLS equation plays an important role in plasma physics [DEGM82], statistical and condensed matter physics [BKT80], nonlinear optics [Ag01], hydrodynamics [ZK97], and magnetic systems [DHS01]. Recently, the NLS equation has been derived as the modulation equation for wave packets in spatially periodic media such as photonic band gap materials and Bose-Einstein condensates [dSS88, DGPS99]. This provides our motivation to concentrate on wave propagation processes in periodic media from a rigorous mathematical point of view. Without considering any particular physical problem, this paper provides error estimates and justification of the validity of the NLS equation in spatially periodic systems with either quadratic or cubic

nonlinearities. In the presence of quadratic nonlinearities, the NLS equation is a relevant model only when additional non-resonance conditions are fulfilled so that the nonlinear response of system effectively becomes cubic [Lv94]. Clearly, this is the mathematically more challenging case.

1.1 The main results

In order to keep the technicalities on a reasonable level we consider throughout the paper semilinear wave equations

$$\partial_t^2 u(x,t) = \chi_1(x)\partial_x^2 u(x,t) - \chi_2(x)u(x,t) - \chi_3(x)u^\vartheta(x,t) \tag{1}$$

with $x \in \mathbb{R}$ and $t \in \mathbb{R}$, $u = u(x,t) \in \mathbb{R}$, $\vartheta = 2$ or $\vartheta = 3$, in a spatially periodic medium. That means that the smooth coefficient functions $\chi_j = \chi_j(x)$ satisfy $\chi_j(x) = \chi_j(x + L)$ for $j = 1, 2, 3$. We assume $\chi_1(x) > 0$ and $\chi_2(x) > 0$ for all $x \in [0, L)$, and without loss of generality, $L = 2\pi$ throughout the paper. The physically more relevant problem with cubic nonlinearities ($\vartheta = 3$) is mainly used as a warm-up for the mathematically much more involved quadratic case ($\vartheta = 2$).

The linearized problem

$$\partial_t^2 v(x,t) = \chi_1(x)\partial_x^2 v(x,t) - \chi_2(x)v(x,t)$$

is solved by the Bloch waves

$$v(x,t) = f_n(\ell, x)e^{i\ell x}e^{\pm i\omega_n(\ell)t}$$

where $n \in \mathbb{N}$, $\ell \in (-1/2, 1/2]$, with $\omega_n(\ell) \in \mathbb{R}$ determined below, satisfying $\omega_{n+1}(\ell) \geq \omega_n(\ell)$, see Fig. 1 on page 4, and $f_n(x, \ell)$ satisfying

$$f_n(\ell, x) = f_n(\ell, x + 2\pi) \quad \text{and} \quad f_n(\ell, x) = f_n(\ell + 1, x)e^{ix}.$$

Slow modulations in time and space of such a Bloch mode (indexed with n_0) are described by the ansatz

$$u(x,t) = \varepsilon A(\varepsilon(x + c_g t), \varepsilon^2 t)f_{n_0}(\ell_0, x)e^{i\ell_0 x}e^{i\omega_{n_0}(\ell_0)t} + \text{cc} + \text{h.o.t}, \tag{2}$$

where cc means complex conjugate, where h.o.t means terms of order ε^2 and higher, where $0 < \varepsilon \ll 1$ is a small parameter, where $c_g = \partial_\ell \omega_{n_0}(\ell_0)$ is the negative group velocity, and where A is the slowly varying envelope. Plugging the ansatz into (1) one finds that A has to satisfy a NLS equation

$$\partial_T A = i\nu_1 \partial_X^2 A + i\nu_2 A|A|^2 \tag{3}$$

with coefficients $\nu_1 \in \mathbb{R}$ and $\nu_2 \in \mathbb{R}$. This describes via the complex valued amplitude $A(X,T) \in \mathbb{C}$ slow modulations in time $T = \varepsilon^2 t$, and space $X = \varepsilon(x + c_g t)$, of the underlying
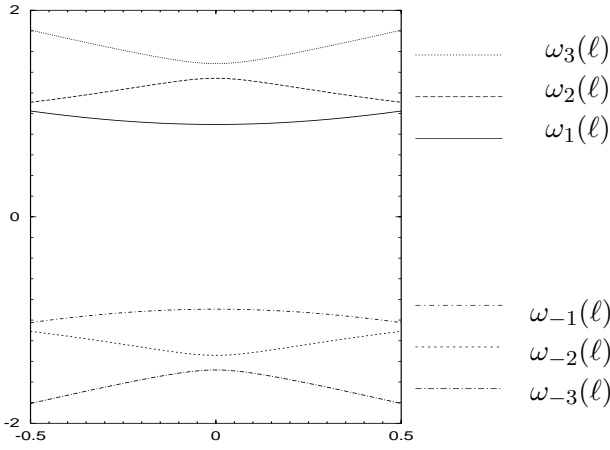
3

Figure 1: The curve of eigenvalues in the spatially periodic case over the Bloch wave numbers, with $\omega_{-n} = -\omega_n$ for later use. Some spectral gaps occur.

wave $f_{n_0}(\ell_0, x)\mathrm{e}^{\mathrm{i}\ell_0 x}\mathrm{e}^{\mathrm{i}\omega_{n_0}(\ell_0)t}$. Here we show that to a solution A of (3) given for $T \in [0, T_0]$, the difference between the formal approximation (2) and exact solutions of (1) stays small for all t in the long time interval $[0, T_0/\varepsilon^2]$ under the validity of a number of assumptions to be checked. For the sake of clarity the formulation of these rather technical non resonance conditions is shifted behind the formulation of the approximation results.

Notation. Fourier transform is defined by $\hat{u}(k) = (\mathcal{F}u)(k) = \frac{1}{2\pi}\int \mathrm{e}^{-\mathrm{i}kx}u(x)\,\mathrm{d}x$. The Sobolev space H^s is equipped with the norm

$$\|u\|_{H^s} = \left(\int |\hat{u}(k)|^2(1 + |k|^2)^s\,\mathrm{d}k\right)^{1/2}.$$

Moreover, let $\|u\|_{C_b^n} = \sum_{j=0}^n \|\partial_x^j u\|_{C_b^0}$, where $\|u\|_{C_b^0} = \sup_{x\in\mathbb{R}}|u(x)|$. We use the spaces $L^1(m)$ with $u \in L^1(m) \Leftrightarrow u\rho \in L^1$, where $\rho(k) = (1 + k^2)^{1/2}$, the weighted Sobolev spaces $H^s(m)$ equipped with the norm $\|u\|_{H^s(m)} = \|u\rho^m\|_{H^s}$, and $L^2(m) = H^0(m)$.

For $\vartheta = 3$ we have the following approximation result.

Theorem 1.1 *Let $s \geq 1$, $s_A \geq \max\{3, s\}$, and assume that the non resonance condition (4) below holds. Then for all C_1 and $T_0 > 0$ there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A})$ of (3) with*

$$\sup_{T\in[0,T_0]} \|A(\cdot, T)\|_{H^{s_A}} \leq C_1$$

the following holds. For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions $u \in C([0, T_0/\varepsilon^2], H^s)$ of (1) with

$$\sup_{t\in[0,T_0/\varepsilon^2]} \left\|u(\cdot, t) - \left(\varepsilon A(\varepsilon(\cdot + c_g t), \varepsilon^2 t)f_{n_0}(\cdot, \ell_0)\mathrm{e}^{\mathrm{i}\ell_0\cdot}\mathrm{e}^{\mathrm{i}\omega_{n_0}^+(\ell_0)t} + \mathrm{cc}\right)\right\|_{H^s} \leq C_2\varepsilon^{3/2} .$$

For $\vartheta = 2$ we have the following approximation result.

4

Theorem 1.2 *Assume that the χ_j are chosen such that the non resonance conditions (5)–(7) below hold with $\alpha > 1/2$, and let $s \in (1/2, \alpha + 1/2)$, $s_A \geq \max\{4, s\}$. Then for all C_1 and $T_0 > 0$ there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A})$ of (3) with*

$$\sup_{T \in [0,T_0]} \|A(\cdot, T)\|_{H^{s_A}} \leq C_1$$

the following holds. For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions $u \in C([0, T_0/\varepsilon^2], H^s)$ of (1) with

$$\sup_{t \in [0,T_0/\varepsilon^2]} \left\| u(\cdot, t) - \left(\varepsilon A(\varepsilon(\cdot + c_g t), \varepsilon^2 t) f_{n_0}(\cdot, \ell_0) \mathrm{e}^{\mathrm{i}\ell_0 \cdot} \mathrm{e}^{\mathrm{i}\omega_{n_0}^+(\ell_0)t} + \mathrm{cc} \right) \right\|_{H^s} \leq C_2 \varepsilon^{3/2} \ .$$

Remark 1.3 The error of order $\mathcal{O}(\varepsilon^{3/2})$ in Theorem 1.1 and Theorem 1.2 is much smaller than the solution and the approximation which are both of order $\mathcal{O}(\varepsilon)$ in L^∞ and of order $\mathcal{O}(\varepsilon^{1/2})$ in H^s. The long time interval $[0, T_0/\varepsilon^2]$ for (1) is necessary to cover nontrivial dynamics of (3). For instance, Theorems 1.1 and 1.2 show that the soliton dynamics [ZS72] of the NLS equation (3) approximately can be found in the original systems, too. ⌋

Remark 1.4 Such approximation results are not obvious. There are counterexamples [Sch95, GS01], where formally correctly derived modulation equations similar to (3) make wrong predictions about the dynamics of the original system. For the spatially homogeneous hyperbolic case, approximation results can be found in [Kal88]; see also [KSM92, Sch98, La98, Co02]. See [Schn04] for new ideas to handle systems with resonances. ⌋

Remark 1.5 Locally in time the solutions of (1) can be chosen arbitrarily smooth. However, long time existence and the error estimates are only guaranteed in the space H^s with $s \in (1/2, \alpha + 1/2)$. The lower bounds on s_A are explained in the proofs, see, e.g., the proof of Lemma 4.3 where $A \in H^3$ is used in the cubic case. ⌋

Remark 1.6 We refrain from greatest generality and restrict ourselves to some model problems which already turned out in the spatially homogeneous case to cover all fundamental non-technical difficulties of general systems. As noted above, there is an important difference between the cubic case ($\vartheta = 3$ in (1)) and the quadratic case ($\vartheta = 2$ in (1)). For cubic nonlinearities the proof is based on a simple application of Gronwall's inequality whereas in the quadratic case before the application of Gronwall's inequality the quadratic terms have to be eliminated by a normal form transform. Fourier analysis used in the spatially homogeneous case is replaced in the spatially periodic case by Bloch wave analysis. A special difficulty in case of spatially periodic medium lies in the fact that the approximation equation lives in a spatially homogeneous domain whereas the original system lives in a spatially periodic domain. Hence, Fourier analysis and Bloch wave analysis have to be linked. ⌋

1.2 The non-resonance conditions

The non resonance condition

$$\inf_{n\in\mathbb{Z}\setminus\{0\},j\in\{\pm3,\pm1\},(n,j)\neq\pm(n_0,1)} |\omega_n(j\ell_0) - j\omega_{n_0}(\ell_0)| > 0 \tag{4}$$

is needed in case of cubic nonlinearities ($\vartheta=3$) to derive the NLS equation and to obtain a sufficiently small so called residual. This smallness of the residual is necessary in the proof of the approximation result given in Section 4. In order to compute $\omega_n(\ell)$ for $\ell \notin (-1/2, 1/2]$, the $\omega_n(\ell)$ are extended periodically in (4), i.e. $\omega_n(\ell)=\omega_n(\ell+1)$.

Section 5 contains the proof of the approximation result for (1) in case of a quadratic nonlinearity ($\vartheta=2$). In order to transfer the proof from cubic nonlinearities to quadratic nonlinearities and especially to prove estimates on the long time scale of order $\mathcal{O}(1/\varepsilon^2)$ the quadratic terms are eliminated by a normal form transform. This leads to the non-resonance conditions

$$\inf_{n\in\mathbb{Z}\setminus\{0\},|j|\leq4,(n,j)\notin\{-(n_0,1),(n_0,1)\}} |\omega_n(j\ell_0) - j\omega_{n_0}(\ell_0)| > 0, \tag{5}$$

$$\inf_{r,n\in\mathbb{Z}\setminus\{0\},\ell,m\in(-\frac{1}{2},\frac{1}{2}],|\ell-m-\ell_0|<\delta} |-\omega_r(\ell) - \omega_{n_0}(\ell - m) + \omega_n(m)| > 0, \tag{6}$$

for a $\delta > 0$. Moreover, we need the following assumption on the quadratic interaction of the Bloch modes $f_n(\ell)$: there exists an $\alpha > 1/2$ and a $C > 0$ such that for all $j, j_1, j_2 \in \mathbb{N}$ and $\ell_1, \ell_2, \ell_3 \in (-1/2, 1/2]$ we have

$$|\langle f_j(\ell_1, \cdot), \chi_3(\cdot) f_{j_1}(\ell_2, \cdot) f_{j_2}(\ell_3, \cdot)\rangle_{\chi_1}| \leq \left(\frac{C}{1 + |j - j_1 - j_2|}\right)^\alpha, \tag{7}$$

where

$$\langle \tilde{u}(\ell, \cdot), \tilde{v}(\ell, \cdot)\rangle_{\chi_1} = \frac{1}{2\pi} \int_0^{2\pi} \tilde{u}(\ell, x)\overline{\tilde{v}(\ell, x)} \frac{1}{\chi_1(x)} \, dx . \tag{8}$$

This assumption is verified in appendix A with an $\alpha = 2 - \delta$ with an arbitrary $\delta > 0$ in case that χ_1 is independent of x and that $\chi_2 \in C^0(0, 2\pi)$ and $\chi_3 \in C^2(0, 2\pi)$.

Conditions (4)-(6) are similar to ones of the spatially homogeneous case, see Remarks 2.1 and 4.2 for further comments. Assumption (7) trivially holds in the spatially homogeneous case due to the orthogonality properties of trigonometric polynomials resulting there in

$$\langle f_j(\ell_1, \cdot), \chi_3(\cdot) f_{j_1}(\ell_2, \cdot) f_{j_2}(\ell_3, \cdot)\rangle_{\chi_1} = \delta(j - j_1 - j_2).$$

So far it is not clear if (7) is really necessary or only an outcome of our proof. Moreover, we are not aware of general conditions on χ_1 and χ_2 leading to the validity of (7). We will discuss in some subsequent remarks the applicability of the assumptions (4)-(6). It will turn out that the validity of (4) and (5) is easily checked due to the asymptotics $|\omega_n| \sim n$ for $|n| \to \infty$. In contrast, for the validity of (6) and also of (7) it is impossible to check the infinitely many conditions in the assumptions (4) or (5)–(7) in concrete physical problems. However, in physical problems dissipation plays a relevant role for modes above a certain n_0. Then the assumptions reduce to finitely many conditions.

1.3 Maxwell's equation

As mentioned above, this work is in part motivated by the fact that nonlinear wave equations with spatially periodic coefficients describe the nonlinear dynamics of electromagnetic radiation in so–called photonic band gap materials, i.e., periodically structured materials which exhibit a gap in their linear spectrum; we refer to [Bu02] for a review of photonic band structures. For instance, the propagation of linearly polarized light in a fiber Bragg grating with intensity dependent index of refraction (Kerr medium) can be modeled by the wave equation

$$-\partial_x^2 E(x,t) + \chi_1(x)\partial_t^2 E(x,t) = -\chi_3(x)\partial_t^2 E^3(x,t), \tag{9}$$

where $\chi_1(x)$ denotes the square of the linear index of refraction. Both $\chi_1(x)$ and the nonlinear Kerr-susceptibility $\chi_3(x)$ are spatially periodic functions. In the case of a weak nonlinearity, a slowly varying envelope approximation reduces (9) to the NLS equation, if the center frequency is not in a gap, see Remark 1.7 for this case. In contrast to homogeneous dielectric media, adjusting the periodicity of the fiber Bragg grating allows one to tailor the linear dispersion properties of this system. As a result, the dynamics of nonlinear waves can also be controlled by this tailoring. Moreover, the unique dispersion properties of fiber Bragg gratings allow novel kinds of localized excitations, the so-called gap solitons, which are simply impossible to realize within homogeneous materials. Gap solitons are stable optical pulses with central (carrier) frequencies inside the photonic band gap of a photonic band gap material [Ac00] and can exhibit propagation velocities ranging from zero all the way to the speed of light. Nonlinear interaction processes between different pulses may be utilized for controlling light with light opening novel possibilities in optical information processing [Tke03, TPB04]. The main difference between (9) and our model (1) is that (9) is quasilinear, and therefore semigroup theory and the variation of constant formula used for (1) have to be replaced by energy estimates. Such energy estimates can be found in Section 6 in case of an additional term $-\chi_2(x)E(x,t)$ on the right hand side. The validity question for (9) as original system remains open.

Remark 1.7 The NLS equation is a valid model if the nonlinear effects are sufficiently weak. However, the NLS equation fails in the case when the central frequency of a nonlinear wave lies inside a spectral gap but is sufficiently far from the band edge. In order to describe such situations, nonlinear coupled mode equations (NLCME) should be employed [Tke03]. Here, it is important to note that the NLCME allows one to account exactly for the systems linear dispersion. For sufficiently wide pulses with narrow spectral content and central frequency either near the band edge or within the allowed bands the NLCME can be reduced to the NLS equation. However, we would like to emphasize that the NLCMEs allow one to analyze the dynamics of nonlinear waves with central frequencies within and/or near a certain spectral gap only. See [GWH01, SU01] for a mathematical justification of the NLCMEs. ⌋

Remark 1.8 According to the Kramers-Kronig relations, in optics material dispersion always comes together with dissipation. However, the dispersion near a photonic band edge

induced by the spatial periodicity of a photonic band gap material is orders of magnitudes larger than the dispersion of the constituent materials themselves and is not subject to Kramers-Kronig relations. Therefore, very often dissipative effects can be neglected in such systems. In this work, we, too, ignore dissipative effects. See [SU03] for a work where dissipation has been taken into account and a complex Ginzburg-Landau equation for Maxwell's equations in homogeneous media has been justified. Similarly, material dispersion and dissipation effects within the NLS equation have been discussed in [BS01].   ⌋

Remark 1.9 An approximation result for the approximation by a linear Schrödinger equation in a nonlinear quantum mechanical context has been shown recently in [CMS04]. We refer to this paper and to [Te03] for an overview about additional results which hold in case that the original system is linear.   ⌋

Notation: Throughout this paper many different constants are denoted by C if they are independent of the small perturbation parameter $0 < \varepsilon \ll 1$. Bloch transform is defined in (27). Functions u and operators M in Fourier space and Bloch space will be denoted with $\hat{u}$, $\tilde{u}$, $\hat{M}$ and $\tilde{M}$, respectively.

2   Multiple scaling analysis

In order to understand the derivation of the NLS equation for (1) we first consider the cubic constant coefficient case

$$\partial_t^2 u(x,t) = \partial_x^2 u(x,t) - u(x,t) - u^3(x,t). \tag{10}$$

Here the linearized problems has solution $v(x,t) = \mathrm{e}^{\mathrm{i}kx}\mathrm{e}^{\pm\mathrm{i}\mu(k)t}$, where $\mu^2(k) = k^2 + 1$. The ansatz

$$u(x,t) = \varepsilon\psi_A(x,t) = \varepsilon A(\varepsilon(x + c_{\mathrm{g}}t), \varepsilon^2 t)\mathrm{e}^{\mathrm{i}k_0 x}\mathrm{e}^{\mu(k_0)t} + \mathrm{cc}, \tag{11}$$

with $0 < \varepsilon \ll 1$ a small parameter, and $k_0 > 0$ the basic spatial wavenumber describes via the complex valued amplitude $A(X,T) \in \mathbb{C}$ slow modulations in time $T = \varepsilon^2 t$, and space $X = \varepsilon(x + c_{\mathrm{g}}t)$, of the underlying wave $\mathrm{e}^{\mathrm{i}k_0 x}\mathrm{e}^{\mathrm{i}\mu(k_0)t}$. Plugging (11) into (10) and equating the coefficients in front of $\varepsilon^m\mathrm{e}^{\mathrm{i}n(k_0 x + \mu(k_0)t)}$ to zero gives for $(m,n) = (1,1)$ the

linear dispersion relation $\mu^2(k_0) = k_0^2 + 1$, for $(m, n) = (2, 1)$ the negative group velocity $c_g = \partial_k \mu(k)|_{k=k_0}$, and for $(m, n) = (3, 1)$ the NLS equation

$$\partial_T A = i\nu_1 \partial_X^2 A + i\nu_2 A|A|^2 \tag{12}$$

with coefficients $\nu_1 = -\frac{1}{2}\partial_k^2 \mu(k_0)$ and $\nu_2 = \frac{3}{2\mu(k_0)}$.

Remark 2.1 In order to show an approximation result for the description of solutions of (10) via (11) and (12) we need to make the so called residual

$$\text{Res}(\varepsilon\psi) = \varepsilon(-\partial_t^2 \psi + \partial_x^2 \psi - \psi) - \varepsilon^3 \psi^3$$

sufficiently small. Therefore we must eliminate the remaining $\mathcal{O}(\varepsilon^3)$ terms obtained at $(m, n) = (3, 3)$ in the above calculation. This is achieved by adding

$$\varepsilon^3 A_3(\varepsilon(x + c_g t), \varepsilon^2 t)e^{3i(k_0 x + \mu(k_0)t)} + \text{cc}$$

to the ansatz (11) and using the non resonance condition

$$|\mu(3k_0) - 3\mu(k_0)| > 0 \tag{13}$$

to define $A_3 = (\mu(3k_0) - 3\mu(k_0))^{-1}A^3$. Hence, in the constant coefficient case we only have one non resonance condition instead of the (formally) infinitely many in (4). ⌋

The classical approach in the (cubic) spatially periodic case proceeds the same way. The main difference is that the carrier wave is no longer a Fourier mode $e^{ik_0 x}e^{\mu(k_0)t}$ but a Bloch mode $f_{n_0}(\ell_0, x)e^{i\ell_0 x}e^{i\omega_{n_0}(\ell_0)t}$ for some $n_0 \in \mathbb{Z}\backslash\{0\}$, $\ell_0 \in (-1/2, 1/2]$, and that the calculations become somewhat more complicated. We use the normalization $\|f_{n_0}\|_{L^2_{\chi_1}((0,2\pi]]} = 1$ and make the ansatz

$$u(x, t) = \varepsilon\psi_A(x, t) = \varepsilon A(\varepsilon(x + c_g t), \varepsilon^2 t)f_{n_0}(\ell_0, x)e^{i\ell_0 x}e^{i\omega_{n_0}(\ell_0)t} + \text{cc}, \tag{14}$$

where again $T = \varepsilon^2 t$, $X = \varepsilon(x + c_g t)$, and $A(X, T) \in \mathbb{C}$. Substituting this into (1) and equating the coefficients in front of $\varepsilon^m e^{in(\ell_0 x + i\omega_{n_0}(\ell_0)t)}f_{n_0}(\ell_0, x)$ to zero we find the linear dispersion relation, the negative group velocity $c_g = \partial_\ell \omega_{n_0}(\ell_0)$, and finally that A has to solve the NLS equation

$$\partial_T A = i\nu_1 \partial_X^2 A + i\nu_2 A|A|^2 \tag{15}$$

with

$$\nu_1 = -\frac{1}{2}\partial_\ell^2 \omega_{n_0}(\ell_0) \quad \text{and} \quad \nu_2 = \frac{3}{2\omega_{n_0}(\ell_0)}\int_0^{2\pi} \frac{\chi_3(x)}{\chi_1(x)}|f_{n_0}(\ell_0, x)|^4 \, \mathrm{d}x. \tag{16}$$

These calculations are based on projection onto $\text{span}\{f_{n_0}\}$, explicitly given via the inner product (8), which usually has to be evaluated numerically (see [dSS88, HFBW01] for further discussions). Since our approach to (1) is slightly different we skip here the details and refer to the subsequent sections.

To explain our approach to (1) we review the derivation of (12) from (10) from another point of view. Essentially we want to consider (1) and (10) as first order systems. In order to do so we first consider (10) in Fourier space. The Fourier transform $\hat{u}$ fulfills

$$\partial_t^2 \hat{u}(k,t) = -\mu^2(k)\hat{u}(k,t) - \hat{u}^{*3}(k,t), \quad \mu(k) = \sqrt{k^2+1}, \tag{17}$$

where $\hat{u}^{*p}$ denotes p times convolution, i.e. $(\hat{u} * \hat{v})(k) = \int_{-\infty}^{\infty} \hat{u}(k-m)\hat{v}(m)\,\mathrm{d}m$. To rewrite this second order equation as a first order system we set $\hat{w}(k) = (\hat{u}(k), \frac{1}{\mu(k)}\partial_t\hat{u}(k))$. Then

$$\partial_t \hat{w}(k,t) = \hat{M}(k)\hat{w}(k,t) + \hat{N}(\hat{w})(k,t), \tag{18}$$

where

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

The eigenvalues of $\hat{M}(k)$ are $\pm \mathrm{i}\mu(k)$, with eigenvectors $\phi_1 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ \mathrm{i} \end{smallmatrix}\right)$ and $\phi_2 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ -\mathrm{i} \end{smallmatrix}\right)$ independent of k. Hence, using the (k-independent) unitary transformation $\hat{z} = \hat{U}^*\hat{w}$, with $\hat{U} = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{smallmatrix}\right)$, i.e. $\hat{U}^{-1} = \hat{U}^*$, we can diagonalize (18) to

$$\partial_t \hat{z} = \hat{\Lambda}\hat{z} + \hat{U}^*\hat{N}(\hat{U}\hat{z}), \tag{19}$$

with $\hat{\Lambda}(k) = \mathrm{diag}(\mathrm{i}\mu(k), -\mathrm{i}\mu(k))$.

Fourier transform of the ansatz (11) for (10) yields

$$\hat{u}(k,t) = \varepsilon\varepsilon^{-1}\hat{A}\left(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k-k_0)t} + \mathrm{ccf}, \tag{20}$$

where ccf denotes the Fourier transform of the complex conjugate, i.e., $\hat{u}(k) + \mathrm{ccf} = \hat{u}(k) + \overline{\hat{u}}(-k)$, and where we do not cancel $\varepsilon\varepsilon^{-1}$ into order to display the different origins of ε (from the original amplitude scaling) and ε^{-1} (from the long wave scaling and the scaling laws for Fourier transform). Then for (19) we use the ansatz

$$
\begin{aligned}
\hat{z}(k,t) &= \varepsilon\varepsilon^{-1}\hat{A}_1\left(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k-k_0)t}\vec{e}_1 \\
&\quad + \varepsilon\varepsilon^{-1}\hat{A}_{-1}\left(\frac{k+k_0}{\varepsilon}, \varepsilon^2 t\right)\mathrm{e}^{-\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k+k_0)t}\vec{e}_2
\end{aligned}, \tag{21}
$$

where $\vec{e}_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and $\vec{e}_2 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Hence the Fourier modes of the wave packet are concentrated in an $\mathcal{O}(\varepsilon)$ neighborhood of the basic wave numbers $\pm k_0$. Therefore, the evolution of the wave packet will be strongly determined by $\pm\mu$ at $\pm k_0$. In detail we find for $\mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_{\mathrm{g}}(k-k_0)t}\vec{e}_1$ with $k = k_0 + \varepsilon K$ that

$$
\begin{aligned}
\mathrm{i}\mu(k_0)\hat{A}_1 + \mathrm{i}\varepsilon c_{\mathrm{g}}K\hat{A}_1 + \varepsilon^2\partial_T\hat{A}_1 &= \mathrm{i}\mu(k_0)\hat{A}_1 + \mathrm{i}\varepsilon\partial_k\mu(k_0)K\hat{A}_1 + \frac{\mathrm{i}}{2}\varepsilon^2\partial_k^2\mu(k_0)K^2\hat{A}_1 \\
&\quad + \varepsilon^2\frac{3\mathrm{i}}{4\mu(k_0)}\hat{A}_1 * \hat{A}_1 * \hat{A}_{-1} + \mathcal{O}(\varepsilon^3),
\end{aligned}
$$

10

where we used

$$\int \hat{A}\left(\frac{k-\ell-k_0}{\varepsilon}\right) \mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_\mathrm{g}(k-\ell-k_0)t} \hat{A}\left(\frac{\ell-k_0}{\varepsilon}\right) \mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_\mathrm{g}(\ell-k_0)t}\,\mathrm{d}\ell$$

$$= \varepsilon \int \hat{A}\left(\frac{k-2k_0}{\varepsilon}-m\right)\hat{A}(m)\mathrm{e}^{2\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_\mathrm{g}(k-2k_0)t}\,\mathrm{d}m\ . \quad (22)$$

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 which is the same as (12) except for an additional factor $1/2$ in ν_2 which is due to the normalization

$$\hat{u}(k,t) = (\hat{U}\hat{z}(k,t))_1 = \frac{1}{\sqrt{2}}(\hat{z}_1(k,t)+\hat{z}_2(k,t))$$

$$= \frac{1}{\sqrt{2}}\left[\hat{A}_1\left(\frac{k-k_0}{\varepsilon},\varepsilon^2 t\right)\mathrm{e}^{\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_\mathrm{g}(k-k_0)t}+\hat{A}_{-1}\left(\frac{k+k_0}{\varepsilon},\varepsilon^2 t\right)\mathrm{e}^{-\mathrm{i}\mu(k_0)t}\mathrm{e}^{\mathrm{i}c_\mathrm{g}(k+k_0)t}\right]$$

$$(23)$$

Exactly the same thing happens in the spatially periodic case. The Bloch modes of the ansatz are strongly concentrated in an $\mathcal{O}(\varepsilon)$ neighborhood of the basic Bloch wave numbers $\pm\ell_0$, hence the evolution of the wave packet will be determined by the associated curves $\omega_{\pm n_0}$ at $\pm\ell_0$. This point of view is not new, but it helps to understand the following sections. Note that we need $k_0 > 0$, but not $\ell_0 > 0$.

Remark 2.2 The diagonalization (19) yields little advantage for the derivation of the NLS equation. However, it turns out to be very useful for treating the error equation by semigroup methods, and essentially, for removing quadratic terms by normal form transforms in case of a quadratic nonlinearity; see below. ⌋

3 The diagonalization

In this section we recall some basic facts about scalar Sturm–Liouville operators with periodic coefficients and provide the functional analytic tools used in the subsequent sections 4-6, in particular the diagonalization of our model (1) with respect to the Bloch modes. However, we start with an introductory section discussing the spectral propertias when the spatial periodic operator differs only by a small perturbation from the spatially homogenous one.

3.1 Spectral properties

For fixed Bloch wave number ℓ the Bloch modes $f_n(\ell,x)$ satisfy the spatially periodic eigenvalue problem

$$-\tilde{L}(\ell,\partial_x)f_n(\ell,\cdot) = \chi_1(\cdot)(\partial_x+\mathrm{i}\ell)^2 f_n(\ell,\cdot) - \chi_2(\cdot)f_n(\ell,\cdot) = -(\omega_n(\ell))^2 f_n(\ell,\cdot). \quad (24)$$

11

Since the operator $\tilde{L}(\ell, \partial_x)$ on the left hand side is elliptic in the bounded domain $[0, 2\pi]$ we have for fixed ℓ countable many eigenvalues λ_n, $n \in \mathbb{N}$. In L^2 equipped with the inner product (8) the operator $\tilde{L}(\ell, \partial_x)$ is positive definite and self adjoint such that the eigenvalues $\lambda_n(\ell) = \omega_n^2(\ell)$ are real and positive. They are ordered in such a way that $\lambda_n(\ell) \leq \lambda_{n+1}(\ell)$.

Example 3.1 It is instructive to consider first the constant coefficient case

$$\partial_t^2 v(x, t) = \partial_x^2 v(x, t) - v(x, t) \tag{25}$$

i.e. $\chi_j = 1$ for $j = 1, 2$ independent of x. The solutions are Fourier modes $v(x, t) = e^{ikx}e^{\pm i\mu(k)t}$, where $\mu^2(k) = k^2 + 1$. Artificially we consider the problem in a spatially periodic set-up. In a Bloch wave representation we have

$$v(x, t) = e^{inx}e^{i\ell x}e^{i\omega_n^{\pm}(\ell)t},$$

where $k = n + \ell$, with $n \in \mathbb{Z}$ here and $\ell \in (-\frac{1}{2}, \frac{1}{2}]$. The eigenvalues are related by $\omega_n^{\pm}(\ell) = \pm\mu(n + \ell)$. See Figure 2.
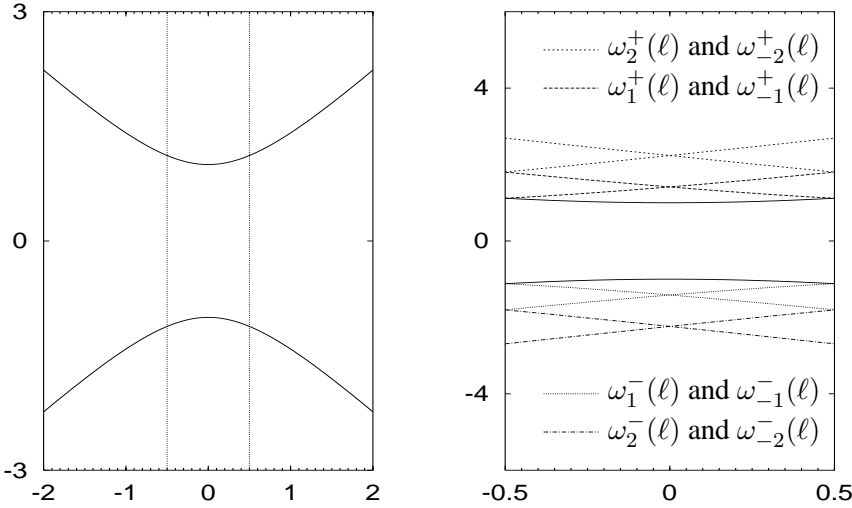

Figure 2: The curve of eigenvalues in the spatially homogeneous case over the Fourier wave numbers (left) and over the Bloch wave numbers (right).

For all $\ell \in (-1/2, 1/2]$ except of $\ell = 0, 1/2$ all eigenvalues in Example 3.1 are simple. By classical perturbation arguments [Kat66], for periodic $\chi_j = 1 + \mathcal{O}(\delta)$ the eigenvalues are smooth functions of δ and stay separated for $\delta > 0$ sufficiently small. However, for $\ell = 0, 1/2$ all eigenvalues are double and generically for small $\delta > 0$ the eigenvalues will split. This is exactly what happens in the spatially periodic case.

Example 3.2 Let $\chi_2(x) = 1 + 2\delta \cos(mx) = 1 + \delta(e^{2imx} + e^{-2imx})$ with $\delta > 0$ small and $m \in \mathbb{N}$ fixed. Setting

$$f_n(\ell, x) = \sum_{k \in \mathbb{Z}} \hat{v}_k^n(\ell)e^{ikx},$$

12

the eigenvalue problem (24) is given by the infinitely many equations

$$(1 + (k + \ell)^2)\hat{v}_k^n(\ell) + \delta(\hat{v}_{k+2m}^n(\ell) + \hat{v}_{k-2m}^n(\ell)) - \lambda_n(\ell)\hat{v}_k^n(\ell) = 0, \qquad k \in \mathbb{Z}. \qquad (26)$$

For $\delta = 0$ we have $\lambda_n^+(0) = \lambda_n^-(0)$, where $\lambda_n^\pm(\ell) = \mu(\pm n + \ell)$, i.e. a crossing of the curves of eigenvalues at $\ell = 0$. Due to the continuity of single eigenvalues or subspaces to eigenvalues separated from the rest, for small $\delta > 0$ and $\ell = 0$ the infinite dimensional eigenvalue problem for instance for $\lambda_m^\pm(0)$ can in lowest order be reduced to a two-dimensional problem for $\hat{v}_m^m$ and $\hat{v}_{-m}^m$, i.e., to

$$\det \begin{pmatrix} 1 + (-m)^2 - \lambda_m^\pm(0) & \delta \\ \delta & 1 + m^2 - \lambda_m^\pm(0) \end{pmatrix} = 0.$$

Hence $\lambda_m^\pm(0) = 1 + m^2 \pm \delta$. Thus $\lambda_m(\ell)$ and $\lambda_{-m}(\ell)$ split at the crossing $\ell = 0$ and recombine in a different way, and similar at $\ell = \frac{1}{2}$. The new curves are also denoted with $\lambda_n(\ell)$, now ordered such that $\lambda_{n+1}(\ell) \geq \lambda_n(\ell)$ and now and in the following indexed with $n \in \mathbb{N}$. We let $\omega_n(\ell) = \sqrt{\lambda_n(\ell)}$ and $\omega_{-n}(\ell) = -\omega_n(\ell) < 0$. See Figure 1.

The computations in Example 3.2 are well documented in the literature [RS78], but we recalled them here since Figure 1 is fundamental for the paper. Note that the splitting of $\lambda_m(\ell)$ and $\lambda_{-m}(\ell)$ in Example 3.2 yields a spectral gap. At the band edge we have a vanishing group velocity $\partial_\ell \lambda_m(\ell) = 0$. Usually the dispersion relation for a periodic dielectric medium has to be calculated numerically, and spectral gaps occur for suitable periodic gratings. See [Bu02] for a review of the photonic band structure computations.

3.2 Bloch transform and diagonalization

To derive and justify the NLS equation for (1) we adapt the Fourier space approach for the constant coefficient problem (10) from Section 2 to Bloch space. Bloch transform is (formally) defined by

$$\begin{aligned} \tilde{u}(\ell, x) &= (\mathcal{T}u)(\ell, x) = \sum_{j \in \mathbb{Z}} e^{ijx} \hat{u}(\ell + j), \\ u(x) &= (\mathcal{T}^{-1})\tilde{u}(x) = \int_{-1/2}^{1/2} e^{i\ell x} \tilde{u}(\ell, x) \, d\ell. \end{aligned} \qquad (27)$$

By construction we have

$$\tilde{u}(\ell, x) = \tilde{u}(\ell, x + 2\pi) \quad \text{and} \quad \tilde{u}(\ell, x) = \tilde{u}(\ell + 1, x)e^{ix}. \qquad (28)$$

Bloch transform is an isomorphism between $H^s(\mathbb{R}, \mathbb{C})$ and $L^2((-1/2, 1/2], H^s([0, 2\pi), \mathbb{C}))$ [RS78, Sca99], where $\|\tilde{u}\|_{L^2((-1/2,1/2], H^s([0,2\pi),\mathbb{C}))} = \left(\int_{-1/2}^{1/2} \|\tilde{u}(\ell, \cdot)\|_{H^s[0,2\pi]}^2 d\ell \right)^{1/2}$. Multiplication $u(x)v(x)$ in x-space corresponds in Bloch space to the operation

$$(\tilde{u} \star \tilde{v})(\ell, x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \tilde{u}(\ell - m, x)\tilde{v}(m, x) \, dm, \qquad (29)$$

13

where (28) has to be used for $|\ell - m| > 1/2$. However, if $\chi : \mathbb{R} \to \mathbb{R}$ is 2π periodic, then

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

Similarly to (22) we need a rescaled version of (29). We define

$$(\tilde{u} \star_\varepsilon \tilde{v})(\ell, x) = \int\limits_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \tilde{u}(\ell - m, x)\tilde{v}(m, x) \, \mathrm{d}m,$$

and $(S_{1/\varepsilon}\tilde{u})(\ell, x) = \tilde{u}(\ell/\varepsilon, x)$ such that

$$(S_{1/\varepsilon}\tilde{u}) \star_\varepsilon (S_{1/\varepsilon}\tilde{u}) = \varepsilon S_{1/\varepsilon}(\tilde{u} \star \tilde{u}). \tag{30}$$

Applying Bloch wave transform to (1) gives

$$\partial_t^2 \tilde{u}(\ell, x) = -\tilde{L}(\ell, \partial_x)\tilde{u}(\ell, x) - \chi_3(x)\tilde{u}^{\star\vartheta}(\ell, x), \tag{31}$$

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

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

For fixed ℓ these operators are self adjoint and positive definite in the space $L_{\chi_1}^2([0, 2\pi), \mathbb{C})$ equipped with the inner product (8), i.e., $\langle \tilde{u}(\ell, \cdot), \tilde{v}(\ell, \cdot)\rangle_{\chi_1} = \int \tilde{u}(\ell, x)\overline{\tilde{v}}(\ell, x)/\chi_1(x) \, \mathrm{d}x$. The induced norm $\|\cdot\|_{L_{\chi_1}^2}$ and the usual L^2 norm are equivalent since $\chi_1(x) \geq \gamma > 0$ for a constant γ independent of x by assumption. The self-adjointness of $\tilde{L}(\ell, \partial_x)$ follows from

$$\begin{aligned}
\langle \tilde{L}(\ell, \partial_x)\tilde{u}, \tilde{v}\rangle_{\chi_1} &= \int \frac{1}{\chi_1(x)}(-\chi_1(x)(\partial_x + \mathrm{i}\ell)^2\tilde{u}(x) + \chi_2(x)\tilde{u}(x))\overline{\tilde{v}(x)} \, \mathrm{d}x \\
&= \int (-(\partial_x + \mathrm{i}\ell)\tilde{u}(x))\overline{(-(\partial_x + \mathrm{i}\ell)\tilde{v}(x))} + \frac{\chi_2(x)}{\chi_1(x)}\tilde{u}(x)\overline{\tilde{v}(x)} \, \mathrm{d}x \\
&= \langle \tilde{u}, \tilde{L}(\ell, \partial_x)\tilde{v}\rangle_{\chi_1},
\end{aligned}$$

and the positive definiteness from

$$\langle \tilde{L}(\ell, \partial_x)\tilde{u}, \tilde{u}\rangle_{\chi_1} = \int |\partial_x\tilde{u}(x)|^2 + \left(\ell^2 + \frac{\chi_2(x)}{\chi_1(x)}\right) |\tilde{u}(x)|^2 \, \mathrm{d}x > 0$$

for $\tilde{u} \neq 0$ in $L_{\chi_1}^2$. 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 $\tilde{L}(\ell, \partial_x)$ with strictly positive eigenvalues $\lambda_j(\ell) > 0$, i.e. $\tilde{L}(\ell, \partial_x)f_j(\ell, \cdot) = \lambda_j(\ell)f_j(\ell, \cdot)$. Moreover, the $(f_j(\ell, \cdot))$ can be chosen smoothly in ℓ since $\tilde{L}(\ell, \partial_x)$ is a relatively bounded perturbation of $-\chi_1(\cdot)\partial_x^2 + \chi_2(\cdot)$. We shall use this repeatedly in the following, see e.g. Lemma 3.3.

We make the ansatz

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

where, due to (28), $\tilde{u}_j(\ell, t) = \tilde{u}_j(\ell + 1, t)$. Since $\tilde{L}(\ell, \partial_x)$ is self-adjoint in $L^2_{\chi_1}$, the eigenfunctions $(f_j(\ell, \cdot))_{j \in \mathbb{N}}$ can be chosen to form an orthonormal basis of $L^2_{\chi_1}$ for each fixed ℓ. Hence $\tilde{u}_j(\ell, t) = \langle f_j(\ell, \cdot), \tilde{u}(\ell, \cdot, t) \rangle_{\chi_1}$, and therefore

$$\partial_t^2 \tilde{u}_j(\ell, t) = -\lambda_j(\ell) \tilde{u}_j(\ell, t) + \langle f_j(\ell, \cdot), \chi_3(\cdot) \tilde{u}^{\star \vartheta}(\ell, \cdot, t) \rangle_{\chi_1}. \tag{32}$$

To write these second order equations as diagonal first order systems, for $j \in \mathbb{N}$ we introduce

$$\omega_j(\ell) = \sqrt{\lambda_j(\ell)}, \quad \omega_{-j}(\ell) = -\omega_j(\ell), \quad \tilde{v}_j(\ell) = \partial_t \tilde{u}_j(\ell)/\omega_j(\ell), \quad \tilde{v}(\ell) = (\tilde{v}_j(\ell))_{j \in \mathbb{N}}, \quad \text{and}$$

$$\tilde{Z}_j = \hat{U}^* \begin{pmatrix} \tilde{u}_j \\ \tilde{v}_j \end{pmatrix}, \quad \text{where} \quad \hat{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \quad \hat{U}^* = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Here $\tilde{Z}_j(\ell, t) = \tilde{Z}_j(\ell + 1, t)$, due to $\tilde{u}_j(\ell, t) = \tilde{u}_j(\ell + 1, t)$ and $\omega_j(\ell) = \omega_j(\ell + 1)$. We denote the mapping of $(\tilde{u}, \tilde{v})$ to $\tilde{Z}(\ell, t) = (\tilde{Z}_j(\ell, t))_{j \in \mathbb{N}}$ by $\mathcal{D}$, i.e., for $\tilde{u}(\ell, x) = \sum_{j \in \mathbb{N}} \tilde{u}_j(\ell) f_j(\ell, x)$ and $\tilde{v}(\ell, x) = \sum_{j \in \mathbb{N}} \tilde{v}_j(\ell) f_j(\ell, x)$ we set

$$\mathcal{D}(\ell) \begin{pmatrix} \tilde{u}(\ell) \\ \tilde{v}(\ell) \end{pmatrix} = (\tilde{Z}_j(\ell))_{j \in \mathbb{N}} \quad \text{with} \quad \tilde{Z}_j(\ell) = \frac{1}{\sqrt{2}} \begin{pmatrix} \tilde{u}_j(\ell) - i\tilde{v}_j(\ell) \\ \tilde{u}_j(\ell) + i\tilde{v}_j(\ell) \end{pmatrix}. \tag{33}$$

Here $\mathcal{D}$ depends on ℓ via the orthonormal system $f_n(\ell)$.

It is also convenient to write

$$\tilde{Z}_j = \begin{pmatrix} \tilde{z}_j \\ \tilde{z}_{-j} \end{pmatrix}, \quad j \in \mathbb{N},$$

which yields

$$\mathcal{D}(\ell)^{-1}((\tilde{Z}_j(\ell))_{j \in \mathbb{N}})(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sum\limits_{j \in \mathbb{N}} (\tilde{z}_j + \tilde{z}_{-j})(\ell) f_j(\ell, x) \\ \sum\limits_{j \in \mathbb{N}} i(\tilde{z}_j - \tilde{z}_{-j})(\ell) f_j(\ell, x) \end{pmatrix}. \tag{34}$$

Now (32) yields the diagonal first order system

$$\partial_t \tilde{Z}_j(\ell, t) = \Lambda_j(\ell) \tilde{Z}_j(\ell, t) - \hat{U}^* \begin{pmatrix} 0 \\ \tilde{s}_j(\ell, t) \end{pmatrix}, \tag{35}$$

with $\tilde{\Lambda}_j(\ell) = \mathrm{diag}(i\omega_j(\ell), i\omega_{-j}(\ell))$ and

$$\tilde{s}_j(\ell, t) = \omega_j^{-1}(\ell) \langle f_j(\ell, \cdot), \chi_3(\cdot) \tilde{u}^{\star \vartheta}(\ell, \cdot, t) \rangle_{\chi_1},$$

where $\tilde{u}$ is given by the first component of the right hand side of (34). In the quadratic case $\vartheta = 2$ we have

$$\tilde{s}_j(\ell, t) = \int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j_1, j_2 \in \mathbb{N}} \beta_{j_1, j_2}^j(\ell, \ell - \ell_1, \ell_1)(\tilde{z}_{j_1} + \tilde{z}_{-j_1})(\ell - \ell_1, t)(\tilde{z}_{j_2} + \tilde{z}_{-j_2})(\ell_1, t) \, d\ell_1,$$

$$\beta_{j_1, j_2}^j(\ell, \ell - \ell_1, \ell_1) = \frac{1}{2} \omega_j^{-1}(\ell) \langle f_j(\ell, \cdot), \chi_3(\cdot) f_{j_1}(\ell - \ell_1, \cdot) f_{j_2}(\ell_1, \cdot) \rangle_{\chi_1},$$

15

which we abbreviate as

$$\tilde{s}_j(\ell, t) = \int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j_1, j_2 \in \mathbb{Z}\setminus\{0\}} \tilde{b}^j_{j_1, j_2}(\ell, \ell-\ell_1, \ell_1)\tilde{z}_{j_1}(\ell-\ell_1, t)\tilde{z}_{j_2}(\ell_1, t)\, \mathrm{d}\ell_1, \tag{36}$$

$$\tilde{b}^j_{j_1, j_2}(\ell_1, \ell_2, \ell_3) = \tilde{\beta}^{|j|}_{|j_1|, |j_2|}(\ell_1, \ell_2, \ell_3)\mathrm{sgn}(jj_1j_2). \tag{37}$$

The cubic case $\vartheta = 3$ yields

$$\tilde{s}_j(\ell, t) = \int\limits_{-\frac{1}{2}}^{\frac{1}{2}}\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j_1, j_2, j_3 \in \mathbb{N}} \tilde{\beta}^j_{j_1, j_2, j_3}(\ell, \ell-\ell_1, \ell_1-\ell_2, \ell_2)$$
$$(\tilde{z}_{j_1}+\tilde{z}_{-j_1})(\ell-\ell_1, t)(\tilde{z}_{j_2}+\tilde{z}_{-j_2})(\ell_1-\ell_2, t)(\tilde{z}_{j_3}+\tilde{z}_{-j_3})(\ell_2, t)\, \mathrm{d}\ell_2\, \mathrm{d}\ell_1,$$
$$\tilde{\beta}^j_{j_1, j_2, j_3}(\ell, \ell-\ell_1, \ell_1-\ell_2, \ell_2)=\frac{1}{\sqrt{8}}\omega_j^{-1}(\ell)\langle f_j(\ell, \cdot), \chi_3(\cdot)f_{j_1}(\ell-\ell_1, \cdot)f_{j_2}(\ell_1-\ell_2, \cdot)f_{j_3}(\ell_1, \cdot)\rangle_{\chi_1},$$

i.e.

$$\tilde{s}_j(\ell, t) = \int\limits_{-\frac{1}{2}}^{\frac{1}{2}}\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j_1, j_2, j_3 \in \mathbb{Z}\setminus\{0\}} \tilde{b}^j_{j_1, j_2, j_3}(\ell, \ell-\ell_1, \ell_1-\ell_2, \ell_2)$$
$$\tilde{z}_{j_1}(\ell-\ell_1, t)\tilde{z}_{j_2}(\ell_1-\ell_2, t)\tilde{z}_{j_3}(\ell_2, t)\, \mathrm{d}\ell_2\, \mathrm{d}\ell_1, \tag{38}$$

$$\tilde{b}^j_{j_1, j_2, j_3}(\ell_1, \ell_2, \ell_3, \ell_4) = \tilde{\beta}^{|j|}_{|j_1|, |j_2|, |j_3|}(\ell_1, \ell_2, \ell_3, \ell_4)\mathrm{sgn}(jj_1j_2j_3). \tag{39}$$

Equation (35) is abbreviated in the following with

$$\partial_t \tilde{Z} = \tilde{\Lambda}\tilde{Z} + \tilde{\mathcal{N}}(\tilde{Z}). \tag{40}$$

It is obvious that we can also regard the nonlinearity $\tilde{\mathcal{N}}(\tilde{Z})$ as a symmetric bilinear or trilinear mapping as in $f(u) = u^2 = b(u, u)$ with $b(u, v) = \frac{1}{2}(uv + vu)$.

3.3 The function spaces

In this section we summarize the analytic properties of the mappings $\mathcal{D}(\ell)$ and give an estimate for bilinear mappings of the form (36). Therefore we introduce

$$\mathcal{X}^s = L^2((-1/2, 1/2], \ell^2(s))$$

equipped with the norm

$$\|\tilde{z}\|_{\mathcal{X}^s} = \left(\int_{-1/2}^{1/2} \sum_{j\in\mathbb{Z}} |\tilde{z}_j(\ell)|^2 |j|^{2s}\, \mathrm{d}\ell\right)^{1/2}.$$

16

In the space $\mathcal{X}^s$ we have the trivial estimate

$$\|\tilde{M}(\cdot)\tilde{z}(\cdot)\|_{\mathcal{X}^s} \leq \sup\{|m_j(\ell)| : \ell \in (-\frac{1}{2}, \frac{1}{2}], \ j \in \mathbb{Z}\} \ \|\tilde{z}(\cdot)\|_{\mathcal{X}^s}$$

where $\tilde{M}(\ell)$ is a diagonal operator with entries $\tilde{m}_j(\ell)$.

Lemma 3.3 *For all $\ell \in (-1/2, 1/2]$, the diagonalization operator $\mathcal{D}(\ell)$ defined in (33), is an isomorphism between $H^s([0, 2\pi)) \times H^s([0, 2\pi))$ and $\ell^2(s) \times \ell^2(s)$ and there exists a $\delta(\ell) > 0$ with*

$$\|D(\ell)\|, \|D^{-1}(\ell)\| \leq \delta(\ell). \tag{41}$$

Moreover, $\sup_{\ell \in (-1/2, 1/2]} \delta(\ell) \leq C < \infty$.

Proof. We use the equivalence of the spaces of fractional powers belonging to the sectorial operators

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

and

$$\tilde{L}_0(\ell, \partial_x)\tilde{u}(\cdot)(x) = -(\partial_x + i\ell)^2\tilde{u}(x) + \tilde{u}(x),$$

in the space $Y = L^2([0, 2\pi), \mathbb{C})$. These spaces are denoted in the following with $Y^\alpha(\ell)$ and $Y_0^\alpha(\ell)$. They are equipped with the norms $\|\tilde{u}\|_{Y^\alpha(\ell)} = \|\tilde{L}^\alpha(\ell, \partial_x)\tilde{u}(\ell, \cdot))\|_Y$ and $\|\tilde{u}\|_{Y_0^\alpha(\ell)} = \|\tilde{L}_0^\alpha(\ell, \partial_x)\tilde{u}(\ell, \cdot))\|_Y$. The equivalence then follows from the estimate for the resolvents

$$\begin{aligned} C_1\|(\tilde{L}_0(\ell, \partial_x) - \lambda I)^{-1}\|_{Y \to Y} &\leq& \|\tilde{L}(\ell, \partial_x) - \lambda I)^{-1}\|_{Y \to Y} \\ &\leq& C_2\|(\tilde{L}_0(\ell, \partial; x) - \lambda I)^{-1}\|_{Y \to Y} \end{aligned}$$

with constants $C_1, C_2 > 0$ independent of ℓ and λ in a suitable chosen sector and from the alternative definition of the spaces of fractional powers given in [He81, sec. 1.4.].

It is well known that the space $Y_0^\alpha(\ell)$ is given by the Sobolev space $H^{2\alpha}([0, 2\pi))$. Therefore, also the space $Y^\alpha(\ell)$ is equivalent to $H^{2\alpha}([0, 2\pi))$. It remains to relate the space $Y^\alpha(\ell)$ with $\ell^2(s)$ where $s = 2\alpha$. From the asymptotics of the eigenvalues it follows that there exist C_3, C_4 such that

$$C_3 j \leq \omega_j \leq C_4 j, \quad j \in \mathbb{N} \tag{42}$$

which holds uniformly in $\ell \in (-1/2, 1/2]$. Then the assertion follows directly from the spectral theorem for self adjoint operators, i.e.,

$$\|\tilde{u}\|_{Y^\alpha(\ell)} = \left(\sum_{j \in \mathbb{N}} |\tilde{u}_j|^2 |\omega_j(\ell)|^{2s}\right)^{\frac{1}{2}}$$

for $\tilde{u}(\ell, x) = \sum_{j \in \mathbb{N}} \tilde{u}_j(\ell)f_j(\ell, x)$. ∎

Clearly, Lemma 3.3 implies that $(\tilde{u}, \tilde{v}) \mapsto \tilde{Z}$ with $\tilde{Z}(\ell) = \mathcal{D}(\ell)(\tilde{u}(\ell), \tilde{v}(\ell))$ is an isomorphism between $L^2((-1/2, 1/2), H^s([0, 2\pi)) \times H^s([0, 2\pi)))$ and $\mathcal{X}^s$.

Lemma 3.4 *For each $s > 1/2$ the nonlinearity $\tilde{\mathcal{N}}(\tilde{Z})$ in (40) is a smooth mapping from $\mathcal{X}^s$ into $\mathcal{X}^{s+1}$, i.e. we have $\|\tilde{\mathcal{N}}(\tilde{Z})\|_{\mathcal{X}^{s+1}} \leq C\|\tilde{Z}\|_{\mathcal{X}^s}^\vartheta$.*

Proof. This follows from Sobolev's embedding theorem, i.e. $\|u^\vartheta\|_{H^s} \leq C\|u\|_{H^s}^\vartheta$, by going back into x–space using $\mathcal{T}$ and Lemma 3.3. In detail, writing

$$\mathcal{D}^{-1}(\tilde{Z}) = \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \in L^2((-1/2, 1/2], H^s) \times L^2((-1/2, 1/2], H^s)$$

we have

$$\|\tilde{\mathcal{N}}(\tilde{Z})\|_{\mathcal{X}^{s+1}} = \left\| \mathcal{D}\mathcal{T} \begin{pmatrix} 0 \\ u^\vartheta \end{pmatrix} \right\|_{\mathcal{X}^{s+1}} \leq C\|\mathcal{T}(u^\vartheta)\|_{L^2((-1/2,1/2], H^s)} \leq C\|u^\vartheta\|_{H^s} \leq C\|u\|_{H^s}^\vartheta$$
$$\leq C\|\tilde{u}\|_{L^2((-1/2,1/2], H^s)}^\vartheta \leq C\|\tilde{Z}\|_{\mathcal{X}^s}^\vartheta.$$

The gain in regularity is due to the the fact that the nonlinearity only depends on u and our choice of $\tilde{v}$, which yields the $\omega_j^{-1}(\ell)$ in the representation

$$\tilde{s}_j(\ell, t) = \omega_j^{-1}(\ell)\big\langle f_j(\ell, \cdot), \chi_3(\cdot)\tilde{u}^{\star\vartheta}(\ell, \cdot, t)\big\rangle_{\chi_1}. \qquad \blacksquare$$

3.4 Estimates for the normal form transform

This following estimate is used in Section 5.2 to estimate the normal form transform in case of quadratic nonlinearities

Lemma 3.5 *Nonlinear terms of the form $s(x) = \mathcal{T}^{-1}(\sum_{j\in\mathbb{Z}} \tilde{s}_j f_j)(x)$ with $\tilde{s}_j(\ell)$ given by (36) are bilinear mappings from $H^q \times H^q$ to H^s if*

$$\sup_{\ell,\ell_1} \sum_{j,j_1,j_2\in\mathbb{Z}\backslash\{0\}} |\tilde{b}_{j_1,j_2}^j(\ell, \ell - \ell_1, \ell_1)|j|^s|j_1|^{-q}|j_2|^{-q}|^2 < \infty. \qquad (43)$$

Proof. Using Young's inequality and the embedding $L^2((-1/2, 1, 2]) \subset L^1((-1/2, 1, 2])$ this follows by applying Schwarz's inequality twice from

$$\left(\sum_j \big(|j|^s \sum_{mn} \tilde{b}_{mn}^j \tilde{z}_m \tilde{z}_n\big)^2\right)^{1/2} = \left(\sum_j \big(|j|^s \sum_{mn} \tilde{b}_{mn}^j |m|^q|m|^{-q}|n|^q|n|^{-q}\tilde{z}_m\tilde{z}_n\big)^2\right)^{1/2}$$
$$\leq \quad \big(\sum_{jmn} ||j|^s\tilde{b}_{mn}^j|m|^{-q}|n|^{-q}|^2\big)^{1/2}\big(\sum_m (|\tilde{z}_m||m|^q)^2\big)^{1/2}\big(\sum_n (|\tilde{z}_n||n|^q)^2\big)^{1/2}$$
$$\leq \quad \left(\sum_{jmn} ||j|^s\tilde{b}_{mn}^j|m|^{-q}|n|^{-q}|^2\right)^{1/2}\|\tilde{z}\|_{\ell^2(q)}^2. \qquad \blacksquare$$

Lemma 3.6 *Assumption (7) on the nonlinear interaction of the Bloch modes implies that the nonlinearity from Lemma 3.5 is smooth from $H^s \times H^s$ into H^s for every $s \in (1/2, \alpha + 1/2)$.*

Proof: Under the validity of the assumption (7) we find

$$\sum_{j,j_1,j_2\in\mathbb{Z}\backslash\{0\}} |\tilde{b}^j_{j_1 j_2}|j|^s|j_1|^{-s}|j_2|^{-s}|^2$$

$$\leq \sum_{j,j_1,j_2\in\mathbb{Z}\backslash\{0\}} \left| \frac{1}{|j|}\left(\frac{C}{1+|j-j_1-j_2|}\right)^\alpha |j|^s|j_1|^{-s}|j_2|^{-s} \right|^2 = s_2 \, .$$

For the finiteness of the sums over j_1 and j_2 we need $s > 1/2$ and $\alpha > 1/2$. Consider

$$s_1 = \sum_{j_1} \left| \frac{1}{1+|j-j_1|^\alpha} |j_1|^{-s} \right|^2 \, .$$

We have two cases:

1) Let $0 < j_1 < j/2$ or $3j/2 < j_1$, then, for $s > 1/2$,

$$s_1 \leq \sum_{j_1} \left| \frac{1}{1+|\frac{j}{2}|^\alpha}|j_1|^{-s} \right|^2 \leq \frac{C}{|j|^{2\alpha}}.$$

2) Let $j/2 < j_1 < 3j/2$, then, for $\alpha > 1/2$,

$$s_1 \leq \sum_{j_1} \left| \frac{1}{1+|j-j_1|^\alpha} |\frac{j}{2}|^{-s} \right|^2 \leq \frac{C}{|j|^{2s}}$$

Thus we have

$$s_2 \leq C \sum_{j\in\mathbb{Z}\backslash\{0\}} \left| \frac{1}{|j|}\left(\frac{1}{1+|j|}\right)^{\min(s,\alpha)} |j|^s \right|^2$$

which is finite for $1 + \min(s,\alpha) - s > 1/2$, i.e. $\alpha \geq s - 1/2$. ∎

4 The cubic case

In the following we use the abbreviation $\mathcal{X} = \mathcal{X}^s$. As in the spatially homogeneous case for cubic nonlinearities the proof of the approximation result is based on a simple application of Gronwall's inequality. However, the additional non-resonance conditions (4) have to be satisfied to make the so called residual

$$\text{Res}(\tilde{Z}) = -\partial_t \tilde{Z} + \tilde{\Lambda}\tilde{Z} + \tilde{\mathcal{N}}(\tilde{Z}) \tag{44}$$

sufficiently small in the prescribed norm in which we intend to measure the magnitude of the error, i.e. here $\mathcal{O}(\varepsilon^4)$ in $\mathcal{X}$. This smallness of the residual is necessary in the proof of the approximation result given in this section.

4.1 Derivation and residual

The simple approximation (14) produces a number of $\mathcal{O}(\varepsilon^3)$ terms which do not cancel. Thus $\|\mathrm{Res}(\varepsilon\tilde{\Psi}^{\mathrm{app}})\|_{\mathcal{X}} = \mathcal{O}(\varepsilon^3)$, where $\varepsilon\tilde{\Psi}^{\mathrm{app}}$ is the approximation for Z defined via (14). As in [KSM92] these $\mathcal{O}(\varepsilon^3)$ terms will be canceled by adding higher order terms to the approximation. For notational simplicity and w.l.o.g. we choose $n_0 = 1$, i.e. we derive the NLS equation for the variable $\tilde{z}_1$. With $\mathbf{E}^j = \mathrm{e}^{j\mathrm{i}\omega_1(\ell_0)t}\mathrm{e}^{\mathrm{i}\omega_1'(\ell_0)(\ell-j\ell_0)t}$, the extension $\varepsilon\tilde{\Psi}$ of $\varepsilon\tilde{\Psi}^{\mathrm{app}}$ is given by

$$
\left.
\begin{aligned}
\tilde{z}_1 =& \varepsilon\varepsilon^{-1}\tilde{A}_1\left(\frac{\ell-\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E} + \varepsilon^3\varepsilon^{-1}\tilde{A}_{13}\left(\frac{\ell-3\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^3 \\
&+ \varepsilon^3\varepsilon^{-1}\tilde{A}_{1-1}\left(\frac{\ell+\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-1} + \varepsilon^3\varepsilon^{-1}\tilde{A}_{1-3}\left(\frac{\ell+3\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-3}, \\
\tilde{z}_{-1} =& \varepsilon\varepsilon^{-1}\tilde{A}_{-1}\left(\frac{\ell+\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-1} + \varepsilon^3\varepsilon^{-1}\tilde{A}_{-1-3}\left(\frac{\ell+3\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-3} \\
&+ \varepsilon^3\varepsilon^{-1}\tilde{A}_{-11}\left(\frac{\ell-\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^1 + \varepsilon^3\varepsilon^{-1}\tilde{A}_{-13}\left(\frac{\ell-3\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^3, \\
\tilde{z}_n =& \sum_{j=-3,-1,1,3}\varepsilon^3\varepsilon^{-1}\tilde{A}_{nj}\left(\frac{\ell-j\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^j, \qquad n \in \mathbb{Z}\setminus\{0,\pm1\}.
\end{aligned}
\right\}
\tag{45}
$$

Here, $\tilde{A}_1 = \chi_{\varepsilon,\ell_0}\hat{A}$ and $\tilde{A}_{-1} = \chi_{\varepsilon,-\ell_0}\overline{\hat{A}}$, where $\chi_{\varepsilon,\ell_0}$ is the cut off function

$$
\chi_{\varepsilon,\ell_0}(\ell) = \begin{cases} 1 & , \quad \varepsilon^{-1}(\ell-\ell_0) \in (-1/2, 1/2], \\ 0 & , \quad \text{else.} \end{cases}
$$

The $\tilde{A}_{ij}$ are variables in Bloch space, i.e., they satisfy the periodicity condition

$$
\tilde{A}_{ij}\left(\frac{\ell+1-j\ell_0}{\varepsilon}\right) = \tilde{A}_{ij}\left(\frac{\ell-j\ell_0}{\varepsilon}\right),
$$

whereas $\hat{A}$ is a variable in Fourier space. Inserting (45) into (40) and equating the coefficients of $\varepsilon^3\mathbf{E}^1$ in the equation for $\tilde{z}_1$ to zero yields

$$
\partial_T\tilde{A}_1 = \frac{\mathrm{i}\omega''(\ell_0)}{2}\kappa^2\tilde{A}_1 + \gamma\tilde{A}_1 \star_\varepsilon \tilde{A}_1 \star_\varepsilon \tilde{A}_{-1},
\tag{46}
$$

where $\ell = \ell_0 + \varepsilon\kappa$ and

$$
\gamma = 3\mathrm{i}\tilde{b}_{111}^1(\ell_0, \ell_0, \ell_0, -\ell_0) \in \mathrm{i}\mathbb{R},
\tag{47}
$$

while A_{-1} satisfies the complex conjugate equation. By (39) we have $\gamma = \mathrm{i}\nu_2$ with ν_2 from (16). In (46), the convolution $\star_\varepsilon$ means

$$
(\tilde{A}_1 \star_\varepsilon \tilde{A}_1 \star_\varepsilon \tilde{A}_{-1})(\kappa) = \int_{-1/(2\varepsilon)}^{1/(2\varepsilon)}\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)}\tilde{A}_1(\kappa-\kappa_1)\tilde{A}_1(\kappa_1-\kappa_2)\tilde{A}_{-1}(\kappa_2)\,\mathrm{d}\kappa_2\,\mathrm{d}\kappa_1.
$$

From (46) we find formally the NLS equation

$$\partial_T \hat{A} = \frac{i\omega''(\ell_0)}{2}\kappa^2 \hat{A} + \gamma \hat{A} * \hat{A} * \overline{\hat{A}}, \tag{48}$$

by letting $\varepsilon \to 0$, in particular $\int_{-1/(2\varepsilon)}^{1/(2\varepsilon)} \to \int_{\mathbb{R}}$.

Similarly, equating the coefficients of $\varepsilon^2 \mathbf{E}^j$ in the equation for $\tilde{z}_n$ to zero, $n \in \mathbb{Z} \setminus \{0\}$ and $j = \pm 3, \pm 1$, yields

$$\begin{aligned}
j\omega_1(\ell_0)\tilde{A}_{nj}(\kappa) &= \omega_n(j\ell_0)\tilde{A}_{nj}(\kappa) \\
&+ \sum_{j_1+j_2+j_3=j,\ j_i \in \{-1,1\}} \tilde{b}_{j_1,j_2,j_3}^n(j\ell_0, j_1\ell_0, j_2\ell_0, j_3\ell_0)\tilde{A}_{j_1} \star_\varepsilon \tilde{A}_{j_2} \star_\varepsilon \tilde{A}_{j_3}(\kappa).
\end{aligned} \tag{49}$$

Thus the $\tilde{A}_{nj}$ can be determined if the countably many non resonance conditions

$$|\omega_n(j\ell_0) - j\omega_1(\ell_0)| > 0, \tag{50}$$

for $j = \pm 3, \pm 1$ and $n \in \mathbb{Z} \setminus \{0\}$ except for $(n,j) = \pm(1,1)$, hold, where again the argument of ω_n is to be taken modulo 1. More precisely, we need the existence of a $C > 0$ such that

$$\inf_{n \in \mathbb{Z}, j \in \{\pm 3, \pm 1\}, (n,j) \neq \pm(1,1)} |\omega_n(j\ell_0) - j\omega_1(\ell_0)| \geq C . \tag{51}$$

Remark 4.1 (51) automatically holds for $|n|$ sufficiently large since $\omega_n(\ell) \sim |n|$. Hence (51) only needs to be checked for $n \leq N_0$ for some $N_0 > 0$ which depends on χ_j. This usually has to be done numerically. Moreover, for our choice u^3 of the nonlinearity (without derivatives) and for given $\varepsilon > 0$, $N_0 = \mathcal{O}(1/\varepsilon)$ will always be sufficient since again due to $\omega_n \sim |n|$ we have $\tilde{b}_{j_1,j_2,j_3}^n \leq C\varepsilon$ for $n > 1/\varepsilon$, hence (49) holds up to $\mathcal{O}(\varepsilon)$ even if we set $\tilde{A}_{nj} = 0$. However this remark is only useful for finite ε and not in the limit $\varepsilon \to 0$. A similar remark holds in the quadratic case, too. ⌋

Remark 4.2 In Remark 2.1 we explained that in the spatially homogeneous case, cf. (10), we only have one non-resonance condition, namely $|\pm \mu(3k_0) - 3\mu(k_0)| > 0$, to achieve $\mathcal{O}(\varepsilon^4)$ for the residual. If we consider (10) artificially in Bloch space, cf. Example 3.1, then we find that similarly (51) reduces to

$$|\omega_1(3\ell_0) - 3\omega_1(\ell_0)| > 0. \tag{52}$$

This holds due to $f_n(\ell, x) = e^{inx}$, (with $n \in \mathbb{Z}$ as in Example 3.1), hence

$$\begin{aligned}
\tilde{\beta}_{j_1,j_2,j_3}^j(\ell, \ell_1, \ell_2, \ell_3) &= \frac{1}{\sqrt{8}}\omega_j^{-1}(\ell)\frac{1}{2\pi}\int_0^{2\pi} e^{i(j-j_1-j_2-j_3)x}\, dx \\
&= \begin{cases} \frac{1}{\sqrt{8}}\omega_j^{-1}(\ell) & \text{if } j - j_1 - j_2 - j_3 = 0 \\ 0 & \text{else} \end{cases},
\end{aligned}$$

independent of $\ell, \ldots, \ell_3$, and consequently $\tilde{b}_{j_1,j_2,j_3}^n = 0$ in (49) except for $n = j = 3$. Thus we can set $\tilde{A}_{nj} = 0$ for $n \neq 3$. ⌋

From a formal point of view we now have that the residual is of order $\mathcal{O}(\varepsilon^4)$. This immediately can be made rigorous by

Lemma 4.3 *Let $m, s \geq 0$ and let $g(\ell)$ satisfy $|g(\ell)| \leq C|\ell - \ell_0|^s$. Then*

$$\|g(\cdot)\hat{A}(\varepsilon^{-1}(\cdot - \ell_0))\|_{L^2(m)} \leq C\varepsilon^{s+1/2}\|\hat{A}\|_{L^2(m+s)}.$$

Proof. This follows immediately from the fact that the left hand side can be estimated by $\sup_{\ell \in \mathbb{R}} |g(\ell)(1 + \varepsilon^{-1}|\ell - \ell_0|)^{-s}|\|\hat{A}\|_{L^1(m+s)}$ and by the scaling properties of L^2. ∎

Lemma 4.4 *Let $\hat{A} \in C([0, T_0], L^2(s_A))$ be a solution of (3). Then*

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|\mathrm{Res}(\varepsilon\tilde{\Psi})\|_{\mathcal{X}} \leq C_{\mathrm{Res}}\varepsilon^{7/2}, \qquad \sup_{t \in [0, T_0/\varepsilon^2]} \|\varepsilon\tilde{\Psi} - \varepsilon\tilde{\Psi}^{app}\|_{\mathcal{X}} \leq C\varepsilon^{3/2},$$

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|\tilde{\Psi}\|_{L^1((-1/2,1/2], H^s)} \leq C.$$

Proof. Applying Lemma 4.3 to, for instance,

$$|\omega_n(\ell) - \omega_n(\ell_0) - \omega_n'(\ell_0)(\ell - \ell_0) - \omega_n''(\ell_0)(\ell - \ell_0)^2/2| \leq C(\ell - \ell_0)^3,$$

and to

$$\|\varepsilon^{-1}S_{1/\varepsilon}(\tilde{A}_1 - \hat{A})\|_{L^2} = \|\varepsilon^{-1}(1-\chi)S_{1/\varepsilon}\hat{A}\|_{L^2} \leq \varepsilon^{-1/2}\sup_{\ell \in \mathbb{R}}|(1-\chi(\epsilon\ell))(1+|\ell|)^{-3}|\|\hat{A}\|_{L^2(s_A)},$$

the formal arguments from above can be made rigorous. This also explains $s_A \geq 3$. ∎

The last estimate in Lemma 4.4 is used to control convolutions that come from the nonlinearity, for instance,

$$\|\tilde{\Psi} \star \tilde{R}\|_{L^2((-1/2,1/2),\ell^2)} \leq C\|\tilde{\Psi}\|_{L^1((-1/2,1/2),\ell^2)}\|\tilde{R}\|_{L^2((-1/2,1/2),\ell^2)},$$

which arises from interactions of the approximation with the error defined below. Note that $\|\tilde{\Psi}\|_{L^2((-1/2,1/2),\ell^2)} \leq C\varepsilon^{-1/2}$.

4.2   Error estimates

In order to justify the NLS equation we write a solution $\tilde{Z}$ of (40) as a sum of the approximation $\varepsilon\tilde{\Psi}$ and the error $\varepsilon^{3/2}\tilde{R}$, i.e. $\tilde{Z} = \varepsilon\tilde{\Psi} + \varepsilon^{3/2}\tilde{R}$. Inserting this into (40) gives

$$\partial_t \tilde{R} = \tilde{\Lambda}\tilde{R} + \tilde{G}(\tilde{\psi}, \tilde{R}) \tag{53}$$

with

$$\|\tilde{G}(\tilde{\Psi}, \tilde{R})\|_{\mathcal{X}} \leq C_1\varepsilon^2\|\tilde{R}\|_{\mathcal{X}} + C_2\varepsilon^{5/2}\|\tilde{R}\|_{\mathcal{X}}^2 + C_{\mathrm{Res}}\varepsilon^2 \tag{54}$$

where C_1 is independent of $\|\tilde{R}\|_{\mathcal{X}}$, and $C_2 = C_2(M)$ is independent of $\|\tilde{R}\|_{\mathcal{X}}$ as long as $\|\tilde{R}\|_{\mathcal{X}} \leq M$ where M is a constant defined below independent of $0 < \varepsilon \ll 1$. Obviously $\tilde{\Lambda}$

generates a strongly continuous group $\tilde{S}(t) = e^{t\tilde{\Lambda}} : \mathcal{X} \to \mathcal{X}$ where $\|\tilde{S}(t)\|_{\mathcal{X} \to \mathcal{X}} \leq C_s$ with C independent of t. For simplicity we assume $R(0) = 0$. Then the variation of constant formula yields

$$\tilde{R}(t) = \int_0^t \tilde{S}(t-\tau)\tilde{G}(\tilde{\Psi}, \tilde{R})(\tau)d\tau \ .$$

By (54) we obtain local existence and uniqueness for (53), and, by Gronwall's inequality,

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\tilde{R}(t)\|_{\mathcal{X}} \leq C_s T_0 (C_{\mathrm{Res}} + 1)e^{C_s C_1 T_0} =: M \tag{55}$$

for all $\varepsilon \in (0, \varepsilon_0)$, if $\varepsilon_0 > 0$ is chosen so small that

$$\varepsilon_0^{1/2} C_2 M^2 \leq 1. \tag{56}$$

In detail, to M defined in (55) we have a constant C_2 in (54) and then an $\varepsilon_0 > 0$ defined in (56). Thus we have shown the following theorem.

Theorem 4.5 *Assume the non–resonance condition (4). For all C_1 and $T_0 > 0$ there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all solutions $\hat{A} \in C([0, T_0], L^2(s_A))$ of (48) with*

$$\sup_{T\in[0,T_0]} \|\hat{A}(T)\|_{L^2(s_A)} \leq C_1$$

the following holds. For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions $\tilde{Z} \in C([0, T_0/\varepsilon^2], \mathcal{X})$ of (40) such that

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\tilde{Z}(t) - \varepsilon\tilde{\Psi}(t)\|_{\mathcal{X}} \leq C_2\varepsilon^{3/2} \ .$$

In x–space Lemma 4.4 and Theorem 4.5 yield Theorem 1.1 via

$$\|u(x,t) - \varepsilon\psi(x,t)\|_{H^s} = \|\mathcal{T}^{-1}[\mathcal{D}^{-1}(\tilde{Z} - \varepsilon\tilde{\Psi}^{\mathrm{app}})]_1\|_{H^s} = \|\mathcal{T}^{-1}[\mathcal{D}^{-1}(\tilde{Z} - \varepsilon\tilde{\Psi}) + \mathcal{O}(\varepsilon^{3/2})]_1\|_{H^s}$$
$$\leq C\|[\mathcal{D}^{-1}\varepsilon^{3/2}\tilde{R}]_1\|_{L^2((-1/2,1/2),H^s(0,2\pi))} + C\varepsilon^{3/2}$$
$$\leq C\varepsilon^{3/2}\|\tilde{R}\|_{\mathcal{X}^s} + C\varepsilon^{3/2} \leq C\varepsilon^{3/2}$$

5   The quadratic case

This section contains the proof of the approximation result for (1) in case of a quadratic non-linearity. In order to transfer the proof from cubic nonlinearities to quadratic nonlinearities, and especially to prove estimates on the long time scale of order $\mathcal{O}(1/\varepsilon^2)$ the quadratic terms are eliminated by a normal form transform. This leads to the infinitely many non-resonance conditions (5),(6), and to assumption (7) on the quadratic interaction of Bloch modes.

5.1 Derivation and residual

In order to derive the NLS equation we now make the ansatz

$$\tilde{z}_1 = \varepsilon\varepsilon^{-1}\tilde{A}_1\left(\frac{\ell - \ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E} + \varepsilon^2\varepsilon^{-1}\tilde{A}_{12}\left(\frac{\ell - 2\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^2$$

$$+ \varepsilon^2\varepsilon^{-1}\tilde{A}_{10}\left(\frac{\ell}{\varepsilon}, \varepsilon^2 t\right) + \varepsilon^2\varepsilon^{-1}\tilde{A}_{1-2}\left(\frac{\ell + 2\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-2}$$

$$\tilde{z}_{-1} = \varepsilon\varepsilon^{-1}\tilde{A}_{-1}\left(\frac{\ell + \ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-1} + \varepsilon^2\varepsilon^{-1}\tilde{A}_{-10}\left(\frac{\ell}{\varepsilon}, \varepsilon^2 t\right)$$

$$+ \varepsilon^2\varepsilon^{-1}\tilde{A}_{-12}\left(\frac{\ell - 2\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^2 + \varepsilon^2\varepsilon^{-1}A_{-1-2}\left(\frac{\ell + 2\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^{-2}$$

$$\tilde{z}_n = \sum_{j=0,-2,2}\varepsilon^2\varepsilon^{-1}\tilde{A}_{nj}\left(\frac{\ell - j\ell_0}{\varepsilon}, \varepsilon^2 t\right)\mathbf{E}^j, \qquad n \in \mathbb{Z} \setminus \{0, \pm 1\},$$

where $\tilde{A}_1 = \chi_{\varepsilon,\ell_0}\hat{A}$ and $\tilde{A}_{-1} = \chi_{\varepsilon,-\ell_0}\hat{\bar{A}}$. The $\tilde{A}_{ij}$ are variables in Bloch space, i.e., they satisfy the periodicity condition

$$\tilde{A}_{ij}\left(\frac{\ell + 1 - j\ell_0}{\varepsilon}\right) = \tilde{A}_{ij}\left(\frac{\ell - j\ell_0}{\varepsilon}\right),$$

whereas $\hat{A}$ is a variable in Fourier space. Plugging this into (40) and equating the coefficients of $\varepsilon^3\mathbf{E}$ in the equation for $\tilde{z}_1$ at ℓ_0 to zero we find

$$\partial_T\tilde{A}_1(\kappa, T) = \frac{\omega_1''(\ell_0)}{2}\kappa^2\tilde{A}_1(\kappa, T) + \sum_{n\in\mathbb{Z}}2\tilde{b}_{n1}^1(\ell_0, \ell_0, 0)(\tilde{A}_1 \star_\varepsilon \tilde{A}_{n0})(\kappa, T)$$

$$+ \sum_{n\in\mathbb{Z}}2\tilde{b}_{n-1}^1(\ell_0, -\ell_0, 2\ell_0)(\tilde{A}_{-1} \star_\varepsilon \tilde{A}_{n2})(\kappa, T). \tag{57}$$

Similarly, setting $\varepsilon^2 E^j$ in the equation for $\tilde{z}_n$ at ℓ_0 to zero with $n \in \mathbb{Z} \setminus \{0\}$ and $j = 0, -2, 2$ we find

$$\begin{aligned}2\omega_1(\ell_0)\tilde{A}_{n2} &= \omega_n(2\ell_0)\tilde{A}_{n2} + \tilde{b}_{11}^n(2\ell_0, \ell_0, \ell_0)\tilde{A}_1\star_\varepsilon\tilde{A}_1, \\ 0\tilde{A}_{n0} &= \omega_n(0)\tilde{A}_{n0} + 2\tilde{b}_{1-1}^n(0, \ell_0, -\ell_0)\tilde{A}_1\star_\varepsilon\tilde{A}_{-1}.\end{aligned} \tag{58}$$

Eliminating the $\tilde{A}_{nj}$ in the equation for $\tilde{A}_1$ by these algebraic relations gives

$$\partial_T\tilde{A}_1 = \frac{\mathrm{i}\omega_1''(\ell_0)}{2}\kappa^2\tilde{A}_1 + \gamma\tilde{A}_1 \star_\varepsilon \tilde{A}_1 \star_\varepsilon \tilde{A}_{-1}$$

with

$$\mathrm{i}\mathbb{R} \ni \gamma = 2\mathrm{i}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{-2}{\omega_n(0)}\tilde{b}_{n1}^1(\ell_0, \ell_0, 0)\tilde{b}_{1-1}^n(0, \ell_0, -\ell_0)$$

$$+ 2\mathrm{i}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{2}{2\omega_1(0) - \omega_n(2\ell_0)}\tilde{b}_{n-1}^1(\ell_0, -\ell_0, 2\ell_0)\tilde{b}_{11}^n(2\ell_0, \ell_0, \ell_0). \tag{59}$$

From this we find formally the NLS equation

$$\partial_T \hat{A} = \frac{\mathrm{i}\omega''(\ell_0)}{2}\kappa^2 \hat{A} + \gamma \hat{A} * \hat{A} * \overline{\hat{A}}, \tag{60}$$

by letting $\varepsilon \to 0$. For the elimination of $\tilde{A}_{n0}$ and $\tilde{A}_{n2}$ we used the non-resonance conditions

$$\omega_n(2\ell_0) - 2\omega_1(\ell_0) \neq 0, \quad \omega_n(0) \neq 0 \tag{61}$$

for all $n \in \mathbb{Z} \setminus \{0\}$ uniformly in n.

For the validity proof we again need that the residual for the approximation is sufficiently small, here of order $\mathcal{O}(\varepsilon^{9/2})$ in $\mathcal{X}$. Exactly as before this can be achieved by adding additional terms of higher order to the approximation which is possible if the non resonance conditions

$$|\omega_n(j\ell_0) - j\omega_1(\ell_0)| > 0$$

holds uniformly for $|j| \leq 4$ and $n \in \mathbb{Z} \setminus \{0\}$ except for $(n, j) = \pm(1, 1)$.

In order to justify the NLS equation we again write the solution $\tilde{Z}$ as sum of such an approximation $\varepsilon\tilde{\Psi}$ satisfying

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|\mathrm{Res}(\varepsilon\tilde{\Psi})\|_{\mathcal{X}} \leq C_{\mathrm{Res}}\varepsilon^{9/2} \tag{62}$$

and an error $\varepsilon^{5/2}R$. The precise construction of the approximation is as in Section 4. The smaller residual ($\mathcal{O}(\varepsilon^{9/2})$ in (62) in contrast to $\mathcal{O}(\varepsilon^{7/2})$ in Lemma 4.4) explains the more restrictive lower bound $s_A \geq 4$ in Theorem 1.2.

5.2 Error estimates

Writing the nonlinear terms $\tilde{\mathcal{N}}(\tilde{Z})$ in (40) as a symmetric bilinear mapping $\tilde{B}(\tilde{Z}, \tilde{Z})$ we find for the error

$$\partial_t \tilde{R} = \tilde{\Lambda}\tilde{R} + 2\varepsilon\tilde{B}(\tilde{\Psi}, \tilde{R}) + \varepsilon^{5/2}\tilde{B}(\tilde{R}, \tilde{R}) + \varepsilon^{-5/2}\mathrm{Res}(\varepsilon\tilde{\Psi}).$$

The main difficulty in proving a bound of order $\mathcal{O}(1)$ for $\tilde{R}$ on the long time scale $\mathcal{O}(1/\varepsilon^2)$ obviously stems from the term

$$2\varepsilon\tilde{B}(\tilde{\Psi}, \tilde{R}).$$

As in the spatially homogeneous case this term will be eliminated by a normal form transformation. In order to do so additional non-resonance condition have to be valid. We make a near identity transformation

$$\tilde{R} = \tilde{W} + 2\varepsilon\tilde{Q}(\tilde{\Psi}, \tilde{W}) \tag{63}$$

with $\tilde{Q}$ also a symmetric bilinear mapping. This yields

$$\partial_t \tilde{W} = \tilde{\Lambda}\tilde{W} + 2\varepsilon[\tilde{B}(\tilde{\Psi}, \tilde{W}) + \tilde{\Lambda}\tilde{Q}(\tilde{\Psi}, \tilde{W}) - \tilde{Q}(\tilde{\Lambda}\tilde{\Psi}, \tilde{W}) - \tilde{Q}(\tilde{\Psi}, \tilde{\Lambda}\tilde{W})] + \mathcal{O}(\varepsilon^2).$$

Hence, in order to eliminate the dangerous $\mathcal{O}(\varepsilon)$-terms in this equation we have to find a $\tilde{Q}$ such that the terms in the parenthesis $[\cdots]$ vanish, i.e. find $\tilde{Q}$ such that

$$-\tilde{\Lambda}\tilde{Q}(\tilde{\Psi},\tilde{W}) + \tilde{Q}(\tilde{\Lambda}\tilde{\Psi},\tilde{W}) + \tilde{Q}(\tilde{\Psi},\tilde{\Lambda}\tilde{W}) = \tilde{B}(\tilde{\Psi},\tilde{W}).$$

The j-th component of $\tilde{Q}(\tilde{\Psi},\tilde{W})$ can be written as

$$\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m,n\in\mathbb{Z}\setminus\{0\}} \tilde{q}_{m,n}^j(\ell,\ell-\ell_1,\,\ell_1)\tilde{\psi}_m(\ell-\ell_1,t)\tilde{W}_n(\ell_1,t)\,\mathrm{d}\ell_1\,,$$

similar to the j-th component of $\tilde{B}(\tilde{\Psi},\tilde{W})$ which we wrote in sec. 3 as

$$\int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{m,n\in\mathbb{Z}\setminus\{0\}} \tilde{b}_{m,n}^j(\ell,\ell-\ell_1,\,\ell_1)\tilde{\psi}_m(\ell-\ell_1,t)\tilde{W}_n(\ell_1,t)\,\mathrm{d}\ell_1,$$

cf. (36). Using the diagonal form of $\tilde{\Lambda}(\ell)$ we find that Q can be defined if the equation

$$\mathrm{i}(-\omega_j(\ell) + \omega_m(\ell-\ell_1) + \omega_n(\ell_1))\tilde{q}_{mn}^j(\ell,\ell-\ell_1,\ell_1) = \tilde{b}_{mn}^j(\ell,\ell-\ell_1,\ell_1)$$

can be resolved with respect to $\tilde{q}_{mn}^j$. Since $\varepsilon\tilde{\psi}(\ell)$ is only of order $\mathcal{O}(\varepsilon)$ in the $\tilde{z}_1$ variable close to ℓ_0 and in the $\tilde{z}_{-1}$ variable close to $-\ell_0$, the term $\tilde{b}_{mn}^j$ has only to be eliminated for these coordinates and wave numbers. Thus, applying Lemma 4.3, the order $\mathcal{O}(\varepsilon)$ terms can be eliminated if the non resonance condition (6) is satisfied.

Due to Lemma 3.6 the assumption (7) implies

$$\sup_{\ell,\ell_1} \sum_{j,j_1,j_2\in\mathbb{Z}/\{0\}} |\tilde{b}_{j_1 j_2}^j(\ell,\ell-\ell_1,\ell_1)|j|^s|j_1|^{-s}|j_2|^{-2}|^2 < \infty.$$

Thus, by (6),

$$\sup_{\ell,\ell_1} \sum_{j,j_1,j_2\in\mathbb{Z}/\{0\}} |\tilde{q}_{j_1 j_2}^j(\ell,\ell-\ell_1,\ell_1)|j|^s|j_1|^{-s}|j_2|^{-s}|^2 < \infty.$$

Lemma 3.5 then implies the boundedness of Q as a mapping from H^s into H^s with an $\mathcal{O}(1)$-bound.

Therefore, the transformation (63) is an isomorphism in $\mathcal{X}$ for $\varepsilon > 0$ sufficiently small. After the transformation we find

$$\partial_t \tilde{W} = \tilde{\Lambda}\tilde{W} + \tilde{g}(\tilde{\psi},\tilde{W})$$

with

$$\|\tilde{g}(\tilde{\psi},\tilde{W})\|_{\mathcal{X}} \leq C_1\varepsilon^2\|\tilde{W}\|_{\mathcal{X}} + C_2\varepsilon^{5/2}\|\tilde{W}\|_{\mathcal{X}}^2 + C_{\mathrm{Res}}\varepsilon^2$$

where C is independent of $\|\tilde{W}\|_{\mathcal{X}}$ if $\|\tilde{W}\|_{\mathcal{X}} \leq M$, with M independent of $0 < \varepsilon \ll 1$. The rest of the approximation result works then line for line as in Section 4.

Thus we have the following result:

26

Theorem 5.1 *Assume (5)–(7). Then for all C_1 and $T_0 > 0$ there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that for all solutions $\hat{A} \in C([0,T_0], L^2(s_A))$ of (60) with*

$$\sup_{T \in [0,T_0]} \|\hat{A}(T)\|_{L^2(s_A)} \leq C_1$$

the following holds. For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions $\tilde{z} \in C([0, T_0/\varepsilon^2], \mathcal{X})$ of (40) such that

$$\sup_{t \in [0,T_0/\varepsilon^2]} \|\tilde{z}(t) - \varepsilon\tilde{\psi}(t)\|_{\mathcal{X}} \leq C_2\varepsilon^{3/2} \ .$$

The rest of the proof of Theorem 1.2 now follows exactly as at the end of Section 4.

6 Estimates for Maxwell's equations

In this section we prove the approximation property for the model coming from nonlinear optics considered in Section 1.3. Here, for technical reasons we add a term $-\chi_2(x)E(x,t)$ on the right hand side of (9) which is needed for the equivalence of the norms $\|\cdot\|_{H_t^s}$ and $E_{s,0}(\cdot)$ defined below. The validity question for (9) as original system remains open.

Thus, suppose that we constructed the approximation $\varepsilon\psi$ for this model as in Section 4 such that the the following holds. The residual

$$\mathrm{Res}(E) = -\partial_x^2 E + \chi_1\partial_t^2 E + \chi_3\partial_t^2(E^3) + \chi_2 E$$

satisfies

$$\sup_{t \in [0,T_0/\varepsilon^2]} \|\mathrm{Res}(\varepsilon\psi)\|_{H^s} \leq C_{\mathrm{Res}}\varepsilon^{\beta+2}$$

for chosen $\beta > 1$ and $s > 1/2$, if $A \in C([0,T_0], H^{s_A})$ satisfies the corresponding NLS, $s_A \geq \max\{3, s\}$.

In the following we use the notation

$$\|u\|_{H_t^s} = \sum_{j=0}^{s} \left(\int |\partial_t^j u(x,t)|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \ .$$

Note that integration is with respect to x, but differentiation with respect to t. We use the estimates

$$\sup_{x \in \mathbb{R}} |\partial_t^j R| \leq \|\partial_t^j R\|_{H^1} = \|\partial_t^j R\|_{L^2} + \|\partial_t^j \partial_x R\|_{L^2} \ ,$$

$$\int u_1 u_2 u_3 u_4 \, \mathrm{d}x \leq \sup_{x \in \mathbb{R}} |u_1 u_2| \, \|u_3\|_{L^2}\|u_4\|_{L^2} \ .$$

We write a solution E of

$$\partial_x^2 E = \chi_1\partial_t^2 E + \chi_3\partial_t^2(E^3) + \chi_2 E$$

27

as a sum of an approximation $\varepsilon\psi$ an an error function $\varepsilon^\beta R$. Then R satisfies

$$
\begin{aligned}
\partial_x^2 R = {} & \chi_1\partial_t^2 R + \chi_2 R + 3\varepsilon^2\chi_3\partial_t^2(\psi^2 R) + 3\varepsilon^{1+\beta}\chi_3\partial_t^2(\psi R^2) + \varepsilon^{2\beta}\chi_3\partial_t^2(R^3) \\
& + \varepsilon^{-\beta}\mathrm{Res}(\varepsilon\psi).
\end{aligned} \tag{64}
$$

We apply ∂_t^n to (64), multiply by $\partial_t^{n+1} R$ and integrate with respect to x. For $n \geq 2$ we find, using the abbreviation $\tilde{R} = \partial_t^n R$,

$$
\int \chi_2 \tilde{R}\partial_t\tilde{R}\,\mathrm{d}x = \frac{1}{2}\,\partial_t \int \chi_2\tilde{R}^2 dx, \quad \int \chi_1\partial_t^2\tilde{R}\partial_t\tilde{R}\,\mathrm{d}x = \frac{1}{2}\,\partial_t\int \chi_1(\partial_t\tilde{R})^2\,\mathrm{d}x\,,
$$

$$
\int \partial_x^2\tilde{R}\partial_t\tilde{R}\,\mathrm{d}x = -\int(\partial_x\tilde{R})\partial_t(\partial_x\tilde{R})\,\mathrm{d}x = -\frac{1}{2}\,\partial_t\int(\partial_x\tilde{R})^2\,\mathrm{d}x\,,
$$

$$
\begin{aligned}
\int \chi_3\partial_t^{n+2}(\psi^2 R)\partial_t\tilde{R}\,\mathrm{d}x &= g_1 + \int \chi_3\psi^2(\partial_t\tilde{R})(\partial_t^2\tilde{R})\,\mathrm{d}x \\
&= g_1 + \frac{1}{2}\,\partial_t\int \chi_3\psi^2(\partial_t\tilde{R})^2\,\mathrm{d}x - \frac{1}{2}\int \chi_3(\partial_t(\psi^2))(\partial_t\tilde{R})^2\,\mathrm{d}x,
\end{aligned}
$$

$$
\begin{aligned}
\int \chi_3\partial_t^{n+2}(\psi R^2)\partial_t\tilde{R}\,\mathrm{d}x &= g_2 + \int \chi_3\psi R\partial_t(\partial_t\tilde{R})^2\,\mathrm{d}x \\
&= g_2 + \partial_t\int \chi_3\psi R(\partial_t\tilde{R})^2\,\mathrm{d}x - \int \chi_3\partial_t(\psi R)(\partial_t\tilde{R})^2\,\mathrm{d}x\,,
\end{aligned}
$$

$$
\begin{aligned}
\int \chi_3\partial_t^{n+2}(R^3)\partial_t\tilde{R}\,\mathrm{d}x &= g_3 + \int(\partial_t\tilde{R})3R^2\partial_t^2\tilde{R}\,\mathrm{d}x \\
&= g_3 + \frac{3}{2}\,\partial_t\int R^2(\partial_t\tilde{R})^2\,\mathrm{d}x - \frac{3}{2}\int \partial_t(R^2)(\partial_t\tilde{R})^2\,\mathrm{d}x\,,
\end{aligned}
$$

with

$$
|g_1| \leq C\|\psi\|_{H_t^{n+2}}^2\|R\|_{H_t^{n+1}}^2\,, \quad |g_2| \leq C\|\psi\|_{H_t^{n+2}}\|R\|_{H_t^{n+1}}^3\,, \quad |g_3| \leq C\|R\|_{H_t^{n+1}}^4.
$$

Hence we define

$$
\tilde{E}_0(R) = \int(\partial_x R)^2\,\mathrm{d}x + \int \chi_1(\partial_t R)^2\,\mathrm{d}x + \int \chi_2 R^2\,\mathrm{d}x
$$

and

$$
\begin{aligned}
\tilde{E}_\varepsilon(R,\tilde{R}) &= \int(\partial_x\tilde{R})^2\,\mathrm{d}x + \int \chi_1(\partial_t\tilde{R})^2\,\mathrm{d}x + \int \chi_2\tilde{R}^2\,\mathrm{d}x + \varepsilon^2\int \chi_3\psi^2(\partial_t\tilde{R})^2\,\mathrm{d}x \\
&\quad + 6\varepsilon^{1+\beta}\int \chi_3\psi R(\partial_t\tilde{R})^2\,\mathrm{d}x + 3\varepsilon^{2\beta}\int R^2(\partial_t\tilde{R})^2\,\mathrm{d}x.
\end{aligned}
$$

Moreover, we define the final energies

$$E_{s,0}(R) = \sum_{j=0}^{s} \tilde{E}_0(\partial_t^j R), \quad \text{and} \quad E_{s,\varepsilon}(R) = \sum_{j=0}^{s} \tilde{E}_\varepsilon(R, \partial_t^j R).$$

For a given $M > 0$ there exist $\varepsilon_0 > 0$, C_1, $C_2 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have

$$C_1 E_{s,0}(R) \leq E_{s,\varepsilon}(R) \leq C_2 E_{s,0}(R)$$

as long as $E_{s,0}(R) \leq M$. For these $\varepsilon > 0$ the above estimates together with, e.g.,

$$\left| \int (\partial_t \tilde{R})(\partial_t^n \mathrm{Res}(\varepsilon \psi)) \, \mathrm{d}x \right| \leq \|\partial_t \tilde{R}\|_{L^2} \|\mathrm{Res}(\varepsilon \psi)\|_{H_t^n},$$

immediately show

$$\begin{aligned}
\partial_t E_{s,\varepsilon}(R) &\leq C_3 \varepsilon^2 E_{s,\varepsilon}(R) + C_4 \varepsilon^{1+\beta} E_{s,\varepsilon}(R)^{3/2} + C_5 \varepsilon^{2\beta} E_{s,\varepsilon}(R)^2 + C_{Res}\varepsilon^2 \\
&\leq (C_3 + 1)\varepsilon^2 E_{s,\varepsilon}(R) + C_{Res}\varepsilon^2 ,
\end{aligned}$$

if

$$C_4 \varepsilon^{\beta-1} M^{1/2} + C_5 \varepsilon^{2\beta-2} M \leq 1 \tag{65}$$

with C_4, C_5 independent of $0 < \varepsilon \ll 1$. Then by Gronwall's inequality

$$E_{s,\varepsilon}(R) \leq (C_{\mathrm{Res}} T_0) e^{(C_3+1)T_0} =: M .$$

Finally choose $\varepsilon_0 > 0$ so small that (65) is satisfied for all $\varepsilon \in (0, \varepsilon_0)$. Since the spatial derivatives of R are related to the temporal derivates of R through the equation for the error, we have the following result:

Theorem 6.1 *Let $A \in C([0, T_0], H^{s_A})$ be a solution of the NLS equation. Then there exist $\varepsilon_0 > 0$, $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have solutions E of (9) such that*

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|E - \varepsilon \psi\|_{H^s} \leq C \varepsilon^{3/2} .$$

A   Estimates for the Fourier expansion of Bloch modes

We show that our technical assumption (7) is at least satisfied if χ_1 is constant, under mild assumptions on χ_2 and χ_3.

Lemma A.1 *Assumption (7) on the nonlinear interaction of the Bloch modes is satisfied with $\alpha = 2-\delta$ for every $\delta > 0$ in the case that χ_1 is independent of x, $\chi_2 \in C_b^0((0, 2\pi))$, and $\chi_3 \in C_b^2((0, 2\pi))$.*

Proof. The proof consists of two steps. In a first step we give estimates for the Fourier expansion of the Bloch modes f_n (Lemma A.2). In a second step Lemma A.2 is used to conclude (7).

Without loss of generality we set $\chi_1 \equiv 1$. Thus, with the abbreviation $\chi = \chi_2$ we consider the eigenvalue problem

$$\partial_x^2 u - \chi(x)u = -\lambda u \quad \text{with} \quad \chi(x) = \chi(x + 2\pi), \quad \text{hence} \quad \chi(x) = \sum_{n \in \mathbb{Z}} \chi_n e^{inx}, \quad (66)$$

with $(\chi_n) \in \ell^2$. The eigenfunctions are given by Bloch waves, and so we have to consider

$$(\partial_x + i\ell)^2 \tilde{u} - \chi(x)\tilde{u} = -\lambda \tilde{u} \quad \text{with} \quad \tilde{u}(x) = \tilde{u}(x + 2\pi) .$$

For the n_0^{th} eigenfunction $\tilde{u}_{n_0}$ we make the ansatz

$$\tilde{u}_{n_0}(x) = \sum_{n \in \mathbb{Z}} u_{n_0,n} e^{inx} ,$$

and obtain

$$(in + i\ell)^2 u_{n_0,n} - \sum_m \chi_{n-m} u_{n_0,m} = -\lambda_{n_0} u_{n_0,n} \quad \text{for all} \quad n \in \mathbb{Z} . \quad (67)$$

Lemma A.2 *There exists a $C > 0$ such that for all $n_0 \in \mathbb{N}$ the following holds. If u_{n_0} solves (67) with $\|u_{n_0}\|_{\ell^2} = 1$ then*

$$|u_{n_0,n}| \le \frac{C}{||n| + n_0|(1 + ||n| - n_0|)} \quad \text{for all} \quad n \in \mathbb{Z}.$$

Proof. For small n_0 the estimate holds due to the elliptic regularity $\|\tilde{u}_{n_0}\|_{H^2} \le C\|\tilde{u}_{n_0}\|_{L^2}$, cf. the proof of Lemma 3.3. Hence it remains to consider large n_0. We introduce

$$\lambda_{n_0} = (n_0 + \ell)^2 + \tilde{\lambda}.$$

With the abbreviation $u_n = u_{n_0,n}$ we then obtain

$$((n_0 + n + 2\ell)(n_0 - n) + \tilde{\lambda})u_n - \sum_m \chi_{n-m} u_m = 0 \quad \text{for} \quad n \ne n_0, \quad (68)$$

$$\tilde{\lambda} u_{n_0} - \sum_m \chi_{n_0-m} u_m = 0 \quad \text{for} \quad n = n_0. \quad (69)$$

We apply a Lyapunov–Schmidt reduction and resolve (68) with respect to

$$u_n = u_n^*(\chi, \ell, \tilde{\lambda})u_{n_0}$$

which is at least possible for n_0 sufficiently large. Inserting this into (69) shows

$$(\tilde{\lambda} - \sum_m \chi_{n_0-m} u_m^*(\chi, \ell, \tilde{\lambda}))u_{n_0} = 0 .$$

We choose $\tilde{\lambda}$ such that $(\tilde{\lambda} - \sum_m \chi_{n_0-m} u_m^*(\chi, \ell, \tilde{\lambda})) = 0$, i.e. for $\tilde{\lambda}$ chosen in this way we find nontrivial solutions $u_{n_0} \neq 0$.

In detail, we start with

$$
\begin{aligned}
|\tilde{\lambda}| &= |\sum_m \chi_{n_0-m} u_m^*(\chi)| \leq \|\chi * u_m^*\|_{\ell^\infty} \leq \|\chi\|_{\ell^2} \|u^*\|_{\ell^2} \\
&\leq \|\chi\|_{\ell^2} \|u^* \cdot u_{n_0}\|_{\ell^2} |u_{n_0}|^{-1} \leq \|\chi\|_{\ell^2} |u_{n_0}|^{-1}.
\end{aligned}
$$

Then from (68) we have

$$
\begin{aligned}
|u_n| &\leq \frac{1}{|n_0 + n + 2\ell||n_0 - n|}(|\sum_m \chi_{n-m} u_m| + |\tilde{\lambda} u_n|) \\
&\leq \frac{C}{|n_0 + n||n_0 - n|}(\|\chi * u\|_{\ell^\infty} + \|\chi\|_{\ell^2} |u_n| |u_{n_0}|^{-1}) \\
&\leq \frac{C}{|n_0 + n||n_0 - n|} \|\chi\|_{\ell^2} (\|u\|_{\ell^2} + 1) \leq \frac{C}{|n_0 + n||n_0 - n|}
\end{aligned}
$$

where we used $\|u\|_{\ell^2} = 1$ and $|u_n| \leq C|u_{n_0}|$. $\blacksquare$

Continuation of the Proof of Lemma A.1. For $\chi_3(x) = \sum_{p \in \mathbb{Z}} \hat{\chi}_p e^{ipx} \in C_b^2$ we have $|\hat{\chi}_p| \leq C(1 + |p|)^{-2}$. By definition we have (omitting the arguments ℓ_1, ℓ_2, ℓ_3)

$$
\begin{aligned}
|\langle f_j, \chi_3 f_m f_n \rangle_{\chi_1}| &= \left| \left\langle \sum_{s_j} f_{js_j} e^{is_j x}, \sum_{s_m} f_{ms_m} e^{is_m x} \sum_{s_n} f_{ns_n} e^{is_n x} \sum_{p \in \mathbb{Z}} \hat{\chi}_p e^{ipx} \right\rangle_{\chi_1} \right| \\
&= \left| \sum_{s_m} \sum_{s_n} \sum_p f_{j(s_m+s_n+p)} f_{ms_m} f_{ns_n} \hat{\chi}_p \right|.
\end{aligned}
$$

Thus, using Lemma A.2 we obtain

$$
\begin{aligned}
|\langle f_j, \chi_3 f_m f_n \rangle_{\chi_1}| \leq C \sum_{s_m} \sum_{s_n} \sum_p & \frac{1}{(|j+s_m+s_n+p| + 1)(|j-s_m-s_n-p| + 1)(1+|p|)^2)} \\
\times & \frac{1}{(|m+s_m| + 1)(|m-s_m| + 1)(|n-s_n| + 1)|(|n+s_n| + 1)}.
\end{aligned}
$$

Next we use that

$$
\sum_p \frac{1}{|j + p| + 1} \frac{1}{|j - p| + 1} \frac{1}{(1 + |p|)^2}
$$

can be estimated as follows:

1) for $|p| < j/2$ or $|p| > 3j/2$ this is

$$
\leq \sum_p \frac{2}{|j|} \frac{2}{|j|} \frac{1}{|p|^2} \leq \frac{C}{|j|^2} \,,
$$

2) for $j/2 < |p| < 3j/2$ this is

$$
\leq \sum_p \frac{1}{(|j - |p|| + 1)^2} \frac{4}{|j|^2} \leq \frac{C}{|j|^2}
$$

31

Finally, for given $\beta \in (0, 1]$ we provide an estimate for

$$\sum_s \frac{1}{(|k-s|+1)^\beta} \frac{1}{(|k+s|+1)^\beta} \frac{1}{(|m-s|+1)^\beta} \frac{1}{(|m+s|+1)^\beta}. \tag{70}$$

Due to the symmetry of the sum we can restrict w.l.o.g. to $k > m > 0$, $s > 0$. Then discarding the $+1$s and using

$$\frac{1}{(k-s)(m-s)} \leq \frac{1}{(k-m)}$$

and

$$\sum_s \frac{1}{(k+s)(m+s)} \leq \sum_s \frac{1}{k^\alpha s^{1-\alpha} s} \leq \frac{C}{k^\alpha} \leq \frac{2C}{(k+m)^\alpha}$$

for a $\alpha \in (0, 1)$ we find that the (70) can be estimated by $C(k-m)^{-\beta}(k+m)^{-\alpha\beta}$.

Using these inequalities $|\langle f_j, \chi_3 f_m f_n \rangle_{\chi_1}|$ can be estimated by

$$\leq \sum_{s_m} \sum_{s_n} \frac{1}{|j+x_m+s_n|+1} \frac{1}{|j-s_m-s_n|+1} \frac{1}{|m+s_m|+1} \frac{1}{|m-s_m|+1}$$

$$\times \frac{1}{|n-s_n|+1} \frac{1}{|n+s_n|+1}$$

$$\leq C \sum_{s_m} \Big(\frac{1}{|m+s_m|+1} \frac{1}{|m-s_m|+1} \frac{1}{|j-n-s_m|+1} \frac{1}{|n-j-s_m|+1} \Big)^{\beta_1}$$

$$\leq C \Big(\frac{1}{|j-n-m|+1} \Big)^{2\beta_2}$$

for all $0 < \beta_1 < \beta_2 < 1$ with different constants C depending on β_1 and β_2. ∎

References

[Ac00]   A. B. Aceves. Optical gap solitons: Past, present, and future; theory and experiments. *Chaos* 10:584-589, 2000.

[Ag01]   G. P. Agrawal. *Nonlinear Fiber Optics*. Academic Press 2001.

[BKT80] A. R. Bishop, J. A. Krumhansl, and S. E. Trullinger. Solitons in condensed matter: A paradigm. *Physica D* 1:1-44, 1980.

[BS01]   N. A. R. Bhat and J. E. Sipe. Optical pulse propagation in nonlinear photonic crystals. *Phys. Rev. E* 64:056604, 2001.

[Bu02]   K. Busch. Photonic band structure theory: assessment and perspective. *C.R. Physique* 3:53-66, 2002.

[CMS04] R. Carles, P. A. Markowich, C. Sparber. Semiclassical asymptotics for weakly nonlinear Bloch waves. Preprint 2004.

[Co02]   T. Colin.   Rigorous derivation of the nonlinear Schr¨odinger equation and Davey-Stewartson systems from quadratic hyperbolic systems. *Asymptot. Anal.* 31(1):69-91, 2002.

[DGPS99] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari.   Theory of Bose-Einstein condensation in trapped gases. *Rev. Mod. Phys.* 71(3):463-512, 1999.

[DEGM82] R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris. *Solitons and Nonlinear Wave Equations*. Academic Press 1982.

[DHS01] S. O. Demokritov, B. Hillebrands, and A. N. Slavin.   Brillouin light scattering studies of confined spin waves: linear and nonlinear confinement. *Phys. Rep.* 348(6):441-489, 2001.

[dSS88] C. M. de Sterke and J. E. Sipe.   Envelope-function approach for the electrodynamics of nonlinear periodic structures. *Phys. Rev. A*, 38(10):5149-5165, 1988.

[FW98]   S. Flach and C. R. Willis. Discrete breathers. *Phys. Rep.*, 295(5):181-264, 1998.

[GS01]   Th. Gallay and G. Schneider.   KP-description of unidirectional dispersive waves - The model case. *Proc. Royal Soc. Edinb. A*, 131(4):885-898, 2001.

[GWH01] R. H. Goodman, M. I. Weinstein, and P. J. Holmes.   Nonlinear propagation of light in one-dimensional periodic structures. *J. Nonlinear Sci.*, 11(2):123-168, 2001.

[He81]   D. Henry. *Geometric theory of semilinear parabolic equations*. Lecture Notes in Mathematics 840. Springer, 1981.

[HFBW01] D. Hermann, M. Frank, K. Busch, and P. W¨olfle. Photonic band structure computations. *Optics Express* 8(3):167-172, 2001.

[Kal88] L. A. Kalyakin. Asymptotic decay of a one-dimensional wavepacket in a nonlinear dispersive medium. *Math. USSR Sbornik*, 60(2):457-483, 1988.

[Kat66] T. Kato. *Perturbation theory for linear operators.* Die Grundlehren der mathematischen Wissenschaften, Band 132 Springer-Verlag New York, Inc., New York, 1966.

[KP03]   Th. Kappeler and J. P¨oschel. *KdV & KAM*, Springer-Verlag, 2003.

[KSM92] P. Kirrmann, G. Schneider, and A. Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. of the Royal Society of Edinburgh*, 122A:85-91, 1992.

[La98]   D. Lannes. Dispersive effects for nonlinear geometrical optics with rectification. *Asymptot. Anal*. 18:111-146, 1998.

[Lv94]   V. S. L'vov. *Wave Turbulence Under Parametric Ecitation: Applications to Magnetics*. Springer-Verlag 1994.

[RS78]   M. Reed and B. Simon. *Methods of Modern Mathematical Physics IV*. Academic Press, 1978.

[Sca99]   B. Scarpellini. *Stability, Instability, and Direct Integrals*. Chapman & Hall/CRC, 1999.

[Sch95]   G. Schneider. Validity and Limitation of the Newell-Whitehead equation. *Math. Nachr.*, 176:249-263, 1995.

[Sch98]   G. Schneider. Justification of modulation equations for hyperbolic systems via normal forms. *NoDEA*, 5:69-82, 1998.

[SU01]   G. Schneider and H. Uecker. Nonlinear coupled mode dynamics in hyperbolic and parabolic periodically structured spatially extended systems. *Asymptot. Anal.*, 28(2):163-180, 2001.

[SU03]   G. Schneider and H. Uecker. Existence and stability of exact pulse solutions for Maxwell's equations from nonlinear optics. *ZAMP*, 54:677-712, 2003.

[Schn04] G. Schneider. Justification and failure of the Nonlinear Schr̈odinger approximation in case nontrivial quadratic resonances. Preprint 2004

[Tke03]   L. Tkeshelashvili. *Interaction of Nonlinear Waves in Photonic Crystals*. Dissertation, Fakulẗat f̈ur Physik, Universiẗat Karlsruhe 2003.

[TPB04]   L. Tkeshelashvili, S. Pereira, and K. Busch. General theory of nonresonant wave interaction: Giant soliton shift in photonic band gap materials. *Europhys. Lett.* 68(2):205-211, 2004.

[Te03]   S. Teufel Adiabatic perturbation theory in quantum dynamics. *Lecture Notes in Mathematics* 1821, Springer 2003.

[Uec01]   H. Uecker. Stable modulating multi-pulse solutions for dissipative systems with resonant spatially periodic forcing. *J. of Nonlin. Sci.*, 11(2):89-121, 2001.

[ZK97]   V. E. Zakharov and E. A. Kuznetsov. Hamiltonian formalism for nonlinear waves. *Sov. Phys. Uspekhi* 40(11):1087-1116, 1997.

[ZS72]   V. E. Zakharov and A. B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Soviet Physics JETP* 34(1): 62-69, 1972.