Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Interaction of modulated pulses in the nonlinear Schrödinger equation with periodic potential

Johannes Giannoulis[1], Alexander Mielke[2,3], Christof Sparber[4]

submitted: 8th May 2007

[1] Zentrum Mathematik
Technische Universität München
Boltzmannstraße 3
85747 Garching b. München
Germany
E-Mail: giannoulis@ma.tum.de

[2] Weierstraß-Institut
für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin
Germany
E-Mail: mielke@wias-berlin.de

[3] Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof
Germany

[4] Wolfgang Pauli Institut Wien
& Fakultät für Mathematik
Universität Wien
Nordbergstraße 15
1090 Wien
Österreich
E-Mail: christof.sparber@univie.ac.at

No. 1221
Berlin 2007

2000 *Mathematics Subject Classification.* 81Q20, 34E13, 34E20, 35Q55.

Key words and phrases. Nonlinear Schrödinger equation, Bloch eigenvalue problem, two scale asymptotics, modulation equations, four-wave interaction.

This research has been supported by the DFG Priority Program 1095 *"Analysis, Modeling and Simulation of Multiscale Problems"* under contract number Mi 459/3-3. CS has been supported by the APART research grant of the Austrian academy of sciences.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We consider a cubic nonlinear Schrödinger equation with periodic potential. In a semiclassical scaling the nonlinear interaction of modulated pulses concentrated in one or several Bloch bands is studied. The notion of closed mode systems is introduced which allows for the rigorous derivation of a finite system of amplitude equations describing the macroscopic dynamics of these pulses.

1 Introduction and main result

In this work we study the asymptotic behavior for $0 < \varepsilon \ll 1$ of the following nonlinear Schrödinger equation (NLS)

$$\mathrm{i}\varepsilon\partial_t u^\varepsilon = -\frac{\varepsilon^2}{2}\Delta u^\varepsilon + V_\Gamma\left(\frac{x}{\varepsilon}\right)u^\varepsilon + \varepsilon\kappa\,|u^\varepsilon|^2 u^\varepsilon, \quad x\in\mathbb{R}^d, t\in\mathbb{R}, \tag{1.1}$$

governing the dynamics of a wave field $u^\varepsilon(t,\cdot)\in L^2(\mathbb{R}^d)$. Here, $\kappa\in\mathbb{R}$ and the potential $V_\Gamma = V_\Gamma(y)\in\mathbb{R}$ is assumed to be smooth and *periodic* with respect to some *regular lattice* $\Gamma\simeq\mathbb{Z}^d$, generated by a given basis $\{\zeta_1,\ldots,\zeta_d\}$, $\zeta_l\in\mathbb{R}^d$, *i.e.*

$$V_\Gamma(y+\gamma) = V_\Gamma(y), \quad \forall\, y\in\mathbb{R}^d, \gamma\in\Gamma\equiv\left\{\gamma = \textstyle\sum_{l=1}^d \gamma_l\zeta_l\in\mathbb{R}^d :\ \gamma_l\in\mathbb{Z}\right\}. \tag{1.2}$$

It is well known that if $\kappa<0$ the solution of (1.1) in general does not exist for all times, *i.e.* finite-time *blow-ups* may occur, *cf.* [25].

The equation (1.1) can be seen as a simplified model of the one considered in [7]. There the main motivation was to study, from a *semiclassical* point of view, the dynamics of a Bose-Einstein condensate in an optical lattice, described by V_Γ, *cf.* [8, 9, 19] for more details. To this end a rescaling of the appearing physical parameters yields an equation similarly to (1.1), but with an additional non-periodic confining potential, which we shall neglect in the following. The parameter $\varepsilon \ll 1$ then describes the *microscopic/macroscopic scale ratio*. The main assumption for the analysis presented in [7] has been that the initial data $u^\varepsilon(0,\cdot)$ is supposed to be of *WKB type* and in particular it has to be concentrated in a single (isolated) *Bloch band* $E_\ell(k)\in\mathbb{R}$. These energy bands describe the spectral subspaces corresponding to the *periodic Hamiltonian operator*

$$H^\varepsilon_{\mathrm{per}} := -\frac{\varepsilon^2}{2}\Delta + V_\Gamma\left(\frac{x}{\varepsilon}\right), \tag{1.3}$$

cf. Section 2.1 below for more details. In the linear case similar WKB approximations have been established earlier in [3, 14], yielding an approximate macroscopic description (*i.e.* on time– and length–scales of order one) of the highly oscillatory solution to (1.1). However, the question concerning a generalization of the results in [7], in particular to the case of multiple bands, has been open so far. Here we will answer this question for initial data which correspond to a sum of modulated *plane waves*.

In order to derive an approximate macroscopic description we shall proceed by a *two scale expansion method* similar to that in [7]. To this end a detailed understanding of the influence of the nonlinearity is crucial. Indeed we will show that the solution to (1.1) can be approximated (in a suitably scaled Sobolev space) via

$$u^\varepsilon(t,x) \sim \sum_{m=1}^{M} a_m(t,x)\chi_{\ell_m}\left(\frac{x}{\varepsilon};k_m\right) \mathrm{e}^{\mathrm{i}(k_m\cdot x - tE_{\ell_m}(k_m))/\varepsilon)} + \mathcal{O}(\varepsilon), \tag{1.4}$$

for $M \in \mathbb{N}$, where the set $\{(k_m,\ell_m) : m = 1, ..., M\}$ is assumed to form a *closed mode system*, see Definition 2.2,t́ of sufficient high order Λ. As we shall see Λ will depend on the spatial dimension d. The amplitudes a_m are the (local-in-time) solutions to the nonlinear system of amplitude equations

$$\mathrm{i}\partial_t a_m + \mathrm{i}\vartheta_m \cdot \nabla_x a_m = \sum_{\substack{p,q,r=1:\\ \Sigma(\mu_p,\mu_q,\mu_r)=\Sigma(\mu_m)}}^{M} \kappa_{(p,q,r,m)}\, a_p\, \overline{a}_q\, a_r, \quad m = 1,\dots,M. \tag{1.5}$$

The above system describes a so-called *four-wave interaction*, also known (most prominently in laser physics and nonlinear optics) as *four-wave mixing*, *cf.* [1, 6, 15]. By (1.4) we allow for nonlinear interactions within the same band but also consider interactions of different bands. In particular, energy or mass transfer between different bands is expected due to the presence of the nonlinearity on the right hand side of (1.5). To our knowledge this phenomenon has never been studied rigorously in the context of Schrödinger type equations.

One should note that the concept of wave mixing is strongly linked to plane waves as considered above. Indeed, if one allows for more general phases (which is possible in the case of a single pulse [7]), a rigorous understanding, even in much simpler cases, is lacking and in particular one can not expect the nonlinear interaction to be maintained on macroscopic time-scales in general. On the other hand we could, without any problems, allow for (smooth) *higher order nonlinearities* as considered in [7]. This however would result into a much more involved resonance-structure for the nonlinear interactions and in order to keep our presentation simple we restrict ourselves to the cubic case.

Before going into more details, let us briefly mention the following mathematically rigorous works which, besides [7], are most closely related to ours: In [24] the same equation as (1.1) is considered but in a slightly different scaling. Similarly, a nonlinear Schrödinger type model is derived in [5] from a semilinear wave equation with periodic coefficients and in [13] from an underlying oscillator chain model. Concerning nonlinear wave interactions, there exist several results (mostly three-wave mixing) in the context of strictly hyperbolic systems, see, *e.g.*, [17, 18, 20, 22], and in the case of microscopically discrete dynamical systems, *cf.* [10], [11] and the references given therein.

The paper is organized as follows: In Section 2 we introduce the basic notions needed in the following, *i.e.* Bloch bands, mode systems and the *closure condition.* We also discuss several illustrative examples of wave mixing. Then in Section 3 the formal derivation of the approximate solution is given in detail. In Section 4 the obtained formal asymptotic description is rigorously established and our main theorem is stated for a closed mode system of order $\Lambda = 2N + 1$, with $N > d/2$. In Section 5 it is shown that this condition can be relaxed though by introducing the concept of *weak closure*. Finally, in Appendix A, we discuss in more detail the underlying Hamiltonian structure of the amplitude equations (1.5).

2 Mode systems and resonances

For our work it is essential to study the spectral properties of

$$H_{\rm per} = -\frac{1}{2}\Delta_y + V_\Gamma(y),$$

since they will basically determine the fast degrees of freedom in our model.

2.1 Bloch's spectral problem

In what follows we denote by Y the centered *fundamental domain* of the lattice Γ, *i.e.*

$$Y := \left\{ \gamma \in \mathbb{R}^d : \ \gamma = \sum_{l=1}^{d} \gamma_l \zeta_l, \ \gamma_l \in \left[-\frac{1}{2}, \frac{1}{2} \right] \right\}. \tag{2.1}$$

By $\mathcal{Y}$ we denote the d-dimensional torus $\mathbb{R}^d_{/\Gamma}$, which is obtained also from Y by identifying opposite faces. Note that writing $H^s(\mathcal{Y})$ then automatically includes periodicity conditions. Moreover, Y^* denotes the corresponding basic cell of the dual lattice Γ^*. Equipped with periodic boundary conditions $\mathcal{Y}^*$ is usually called the *Brillouin zone* and hence we shall denote it by $\mathcal{B} \equiv \mathcal{Y}^*$.

Next, consider the so-called *Bloch eigenvalue problem* [4], *i.e.* the following spectral problem on $\mathcal{Y}$:

$$H_\Gamma(k)\chi_\ell(y;k) = E_\ell(k)\chi_\ell(y;k), \quad k \in \mathcal{B}, \, \ell \in \mathbb{N}, \tag{2.2}$$

where $E_\ell(k) \in \mathbb{R}$ and $\chi_\ell(y) \equiv \chi_\ell(y;k)$ denote the ℓ-th eigenvalue and eigenvector of the *shifted Hamiltonian operator*

$$H_\Gamma(k) := {\rm e}^{-{\rm i}k\cdot y} H_{\rm per}\, {\rm e}^{{\rm i}k\cdot y} = \frac{1}{2}\left(-{\rm i}\nabla_y + k\right)^2 + V_\Gamma\left(y\right).$$

Let us recall some well known facts for this eigenvalue problem, *cf.* [2, 3, 23, 26]. Since V_Γ is smooth and periodic, we get that, for every fixed $k \in \mathcal{B}$, $H_\Gamma(k)$ is self-adjoint on $L^2(\mathcal{Y})$ with domain $H^2(\mathcal{Y})$ and compact resolvent. Hence the spectrum of $H_\Gamma(k)$ is given by

$${\rm spec}(H_\Gamma(k)) = \{E_\ell(k) \ ; \ \ell \in \mathbb{N}\} \subset \mathbb{R}.$$

One can order the eigenvalues $E_\ell(k)$ according to their magnitude and multiplicity such that

$$E_1(k) \leq \ldots \leq E_\ell(k) \leq E_{\ell+1}(k) \leq \ldots$$

Moreover every $E_\ell(k)$ is periodic w.r.t. Γ^* and it holds that $E_\ell(k) = E_\ell(-k)$. The set $\{E_\ell(k);\, k \in \mathcal{B}\}$ is called the ℓth *energy band*. The associated eigenfunctions, the *Bloch functions*, $\chi_\ell(y;k)$ form (for every fixed $k \in \mathcal{B}$) an orthonormal basis in $L^2(\mathcal{Y})$. We choose the usual normalization such that

$$\langle \chi_{\ell_1}(k), \chi_{\ell_2}(k)\rangle_{L^2(\mathcal{Y})} \equiv \int_{\mathcal{Y}} \overline{\chi_{\ell_1}}(y;k)\, \chi_{\ell_2}(y;k)\,\mathrm{d}y = \delta_{\ell_1,\ell_2}, \quad \ell_1,\, \ell_2 \in \mathbb{N}.$$

Concerning the dependence on $k \in \mathcal{B}$, it has been shown, *cf.* [23], that for any $\ell \in \mathbb{N}$ there exists a closed subset $\mathcal{U} \subset \mathcal{B}$ such that $E_\ell(k)$ is analytic in $\mathcal{O} := \mathcal{B}\backslash\mathcal{U}$. Similarly, the eigenfunctions χ_ℓ are found to be analytic and periodic in k, for all $k \in \mathcal{O}$. Moreover it holds that

$$E_{\ell-1}(k) < E_\ell(k) < E_{\ell+1}(k), \quad \forall\, k \in \mathcal{O}.$$

If this condition is satisfied for all $k \in \mathcal{B}$ then $E_\ell(k)$ is said to be an *isolated Bloch band*. Finally we remark that

$$\operatorname{meas}\mathcal{U} = \operatorname{meas}\{k \in \mathcal{B} \mid E_{\ell_1}(k) = E_{\ell_2}(k),\ \ell_1 \neq \ell_2\} = 0.$$

In this set of measure zero one encounters so-called *band crossings*. The elements of this set are characterized by the fact that $E_\ell(k)$ is only Lipschitz continuous and hence the *group velocity* $\vartheta := \nabla_k E_\ell(k)$ does *not* exist.

2.2 Resonances and closed mode systems

Our goal is to derive an approximate description of our model for $\varepsilon \ll 1$. To this end we shall first introduce several definitions needed to do so.

Definition 2.1. For $k \in \mathcal{B}$ and $\ell \in \mathbb{N}$ we call $\mu = (k, \ell)$ a *mode* and $\mathcal{M} := \mathcal{B} \times \mathbb{N}$ the *set of all modes.*
(i) The *graph of all modes* is given by

$$\mathcal{G} = \big\{(k, E_\ell(k)) : (k,\ell) \in \mathcal{M}\big\} \subset \mathcal{B} \times \mathbb{R}.$$

(ii) Given a finite set $\mathcal{S} = \{\mu_1, \ldots, \mu_M : M \in \mathbb{N}\} \subset \mathcal{M}$ of modes, we call $\mathcal{S}^\Lambda$ the (ordered) *mode system of size* $\Lambda \in \mathbb{N}$ *generated by* $\mathcal{S}$.
(iii) We further introduce $\Sigma : \mathcal{S}^\Lambda \to \mathcal{B} \times \mathbb{R}$ by

$$\Sigma(\mu_{m_1}, \ldots, \mu_{m_\Lambda}) := \left(\sum_{\lambda=1}^{\Lambda}{}^{*} (-1)^{\lambda+1} k_{m_\lambda}\,,\ \sum_{\lambda=1}^{\Lambda} (-1)^{\lambda+1} E_{\ell_{m_\lambda}}(k_{m_\lambda}) \right),$$

where $\sum^*$ denotes summation modulo Γ^*, and we write $\mathcal{G}_{\mathcal{S}}^{(\Lambda)} := \Sigma(\mathcal{S}^\Lambda)$ for the corresponding graphs.

In the following the mapping Σ will describe the possible nonlinear interaction of modes (in every order of ε). Note that $\mathcal{G}_{\mathcal{S}}^{(1)} \subset \mathcal{G}$ and moreover, for any $\Lambda \in \mathbb{N}$, it holds $\mathcal{G}_{\mathcal{S}}^{(\Lambda)} \subset \mathcal{G}_{\mathcal{S}}^{(\Lambda+2)}$. Since we are dealing with a cubic nonlinearity, we shall see that indeed $\Lambda \in \mathbb{N}$ takes only *odd* values.

Definition 2.2. Given a finite set of modes $\mathcal{S}$ and a subset $\mathcal{T}$ of $\mathcal{M}$.
(i) An element $(\mu_{m_1}, \dots, \mu_{m_\Lambda}) \in \mathcal{S}^\Lambda$ is called *resonant of order* Λ *to* $\mathcal{T}$, if

$$\Sigma(\mu_{m_1}, \dots, \mu_{m_\Lambda}) \in \mathcal{G}_{\mathcal{T}}^{(1)}.$$

(ii) We say that $\mathcal{S}$ *is closed of order* Λ, if the group velocity $\nabla_k E_\ell(k)$ exists for all $\mu = (k, \ell) \in \mathcal{S}$ and

$$\mathcal{G}_{\mathcal{S}}^{(\Lambda)} \cap \left(\mathcal{G} \setminus \mathcal{G}_{\mathcal{S}}^{(1)}\right) = \emptyset. \tag{2.3}$$

We infer from the above definition that a single mode μ is always resonant of order $\Lambda = 1$ to itself. Note that for any $\Lambda \geq 2$ however, we obtain

$$\Sigma : (\mu_{m_1}, \dots, \mu_{m_\Lambda}) \mapsto \sigma := (k, E) \in \mathcal{B} \times \mathbb{R},$$

which does not necessarily imply $\sigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$. As for the condition (2.3), it is equivalent to saying that for all $(\mu_{m_1}, \dots, \mu_{m_\Lambda}) \in \mathcal{S}^\Lambda$ it holds: Either $\Sigma(\mu_{m_1}, \dots, \mu_{m_\Lambda}) \in (\mathcal{B} \times \mathbb{R}) \setminus \mathcal{G}$, or else there exists a $\mu \in \mathcal{S}$ such that $\Sigma(\mu_{m_1}, \dots, \mu_{m_\Lambda}) = \Sigma(\mu) \in \mathcal{G}_{\mathcal{S}}^{(1)}$. The latter obviously means that $(\mu_{m_1}, \dots, \mu_{m_\Lambda})$ is resonant of order Λ to $\mathcal{S}$. We illustrate these concepts in the examples in Section 2.4.

The condition that $\nabla_k E_\ell(k)$ has to exist implies that we can not deal with mode systems including band crossings, *i.e.* which include a $\mu = (k, \ell)$ such that $k \in \mathcal{U}$. It is known that the higher the dimension d and the higher the band index ℓ, the more likely one encounters such crossings. Thus, in terms of practical use one can expect our analysis to be restricted to cases where only a few bands with low energies are taken into account.

Remark 2.3. Also note that condition (2.3) is not equivalent to $\mathcal{G}_{\mathcal{S}}^{(\Lambda)} \cap \mathcal{G} \subset \mathcal{G}_{\mathcal{S}}^{(1)}$, since due to multiple eigenvalues $E_\ell(k)$, we may have modes $\mu \in \mathcal{S}$ and $\tilde{\mu} \notin \mathcal{S}$ with $\Sigma(\mu) = \Sigma(\tilde{\mu})$.

2.3 The system of amplitude equations

In Section 3 we shall derive a system of amplitude equations describing the macroscopic dynamics of our nonlinear wave interactions. Before going into the details of the derivation let us first discuss the general structure of the obtained system (see also Appendix A). In order to keep our presentation simple, we shall from now on consider only the case of *non-degenerated eigenvalues* $E_\ell(k)$.

Let $\mathcal{S}$ be a given finite set of modes $\{\mu_1, \dots, \mu_M\}$ which is closed of sufficient high order Λ. Then the equations obtained for the modulating amplitudes $a_m(t, x) \in \mathbb{C}$, for $m \in \{1, \dots, M\}$, are as in (1.5), *i.e.*

$$\mathrm{i}\partial_t a_m + \mathrm{i}\vartheta_m \cdot \nabla_x a_m = \sum_{\substack{p,q,r=1: \\ \Sigma(\mu_p,\mu_q,\mu_r)=\Sigma(\mu_m)}}^{M} \kappa_{(p,q,r,m)}\, a_p\, \overline{a}_q\, a_r.$$

Here $\vartheta_m \in \mathbb{R}^d$ is the group velocity corresponding to a given mode $\mu_m = (k_m, \ell_m) \in \mathcal{S}$, *i.e.* $\vartheta_m := \nabla_k E_{\ell_m}(k_m)$, and we further denote by

$$\kappa_{(p,q,r,m)} := \kappa \int_{\mathcal{Y}} \chi_{\ell_p}(y)\overline{\chi}_{\ell_q}(y)\chi_{\ell_r}(y)\overline{\chi}_{\ell_m}(y)\,\mathrm{d}y$$

the *effective coupling constant* $\kappa_{(p,q,r,m)}$, which in general is complex valued.

As in (1.1), the nonlinearity in the amplitude equations is again cubic. It takes into account the sum over all $p, q, r = 1, \dots, M$, for which the *resonance condition of order* $\Lambda = 3$ holds. Explicitly, this is equivalent to the following two conditions

$$k_p - k_q + k_r = k_m, \quad E_{\ell_p}(k_p) - E_{\ell_q}(k_q) + E_{\ell_r}(k_r) = E_{\ell_m}(k_m), \tag{2.4}$$

where the first equation has to be understood as a summation modulo Γ^*. The equations (2.4) describe a so-called *four-wave interaction*, *i.e.* three modes μ_p, μ_q, μ_r being in resonance with a fourth one $\mu_m \in \mathcal{S}$. Clearly, the nonlinear structure on the right-hand side of (1.5) is such that energy transfer between different modes can occur. In other words, even if initially some of the amplitudes $a_m(0, \cdot)$ are zero, they will not remain so in general for times $|t| \neq 0$ (see also Subsection 2.4.1 below).

Concerning existence and uniqueness of solutions to the above given amplitude equations, we only need a local-in-time result.

Lemma 2.4. *For any initial data* $(a_1(0, \cdot), \dots, a_M(0, \cdot)) \in H^S(\mathbb{R}^d)^M$*, with* $S > d/2$*, the system* (1.5) *admits a unique solution*

$$(a_1, \dots, a_M) \in C^0([0, T); H^S(\mathbb{R}^d))^M \cap C^1((0, T); H^{S-1}(\mathbb{R}^d))^M,$$

up to some (finite) time $T > 0$.

Proof. Since the left hand side of (1.5) only generates translations by a constant velocity $\vartheta_m \in \mathbb{R}^d$ and thus conserves the $L^2(\mathbb{R}^d)$ norm of each $a_m(t, x)$, the assertion of the lemma follows by a standard fixed point argument. □

Remark 2.5. Discarding for a moment the nonlinear term in (1.5) we want to point out that the remaining linear transport part is much simpler than in the case of the full WKB type approximation, as discussed in [7] (for a single mode only). In particular we do not run into any problems due to *caustics*, since the phase functions

$$\varphi_m(t, x) = k_m \cdot x - tE_{\ell_m}(k_m)$$

are globally defined. Also note that in the present work the so-called *Berry term* vanishes, since we do not take into account additional non-periodic potentials, *cf.* [7, 21].

2.4 Examples

In order to illustrate the abstract concepts defined in the former subsections we shall in the following consider several particular examples of nonlinear wave interactions appearing in our model.

2.4.1 Four-wave interaction for three pulses

In this example we shall restrict ourselves to a set of modes $\{\mu_m = (k_m, \ell_m) : m = 1, 2, 3\}$ which is closed of order $\Lambda = 3$, at least, and which satisfies the following resonance conditions

$$k_1 - k_2 + k_1 = k_3, \quad E_{\ell_1}(k_1) - E_{\ell_2}(k_2) + E_{\ell_1}(k_1) = E_{\ell_3}(k_3). \tag{2.5}$$

(Again the first summation has to be understood modulo Γ^*.) This yields the following set of amplitude equations

$$\begin{cases} \mathrm{i}\partial_t a_1 + \mathrm{i}\vartheta_1 \cdot \nabla_x a_1 = W_1(a) a_1 + 2\overline{\kappa}_{(1,2,1,3)}\, \overline{a}_1\, a_2\, a_3, \\ \mathrm{i}\partial_t a_2 + \mathrm{i}\vartheta_2 \cdot \nabla_x a_2 = W_2(a) a_2 + \kappa_{(1,2,1,3)}\, a_1^2\, \overline{a}_3, \\ \mathrm{i}\partial_t a_3 + \mathrm{i}\vartheta_3 \cdot \nabla_x a_3 = W_3(a) a_3 + \kappa_{(1,2,1,3)}\, a_1^2\, \overline{a}_2, \end{cases} \tag{2.6}$$

where we shortly denote

$$W_m(a) := \kappa_{(m,m,m,m)} |a_m|^2 + 2 \sum_{\substack{j=1,2,3: \\ j\neq m}} \kappa_{(m,j,j,m)} |a_j|^2 \ \in \mathbb{R}.$$

We consequently expect the solution of (1.1) to be asymptotically described by

$$u^\varepsilon(t,x) \sim \sum_{m=1}^{3} a_m(t,x) \chi_{\ell_m}\left(\frac{x}{\varepsilon}; k_m\right) \mathrm{e}^{\mathrm{i}(k_m \cdot x - E_{\ell_m}(k_m)t)/\varepsilon} + \mathcal{O}(\varepsilon), \quad \ell_m \in \mathbb{N}.$$

It is easily seen that the nonlinearities $W_m(a)a_m$ can not transfer energy from one band to the other as they are homogeneous in the respective $a_m(t,x)$ and that all of the appearing $\kappa_{(m,j,j,m)}$, $m, j = 1, 2, 3$ are indeed *real valued*. What is more important though is the fact that the above given amplitude system (2.6) includes an *invariant sub-system*. Namely, if initially $a_1(0,\cdot) = 0$, it remains so for all $t \in \mathbb{R}$ and thus the above given system simplifies to

$$\begin{cases} \mathrm{i}\partial_t a_2 + \mathrm{i}\vartheta_2 \cdot \nabla_x a_2 = \left(\kappa_{(2,2,2,2)} |a_2|^2 + 2\kappa_{(2,3,3,2)} |a_3|^2\right) a_2, \\ \mathrm{i}\partial_t a_3 + \mathrm{i}\vartheta_3 \cdot \nabla_x a_3 = \left(\kappa_{(3,3,3,3)} |a_3|^2 + 2\kappa_{(2,3,3,2)} |a_2|^2\right) a_3. \end{cases}$$

However, such a decoupling does not exist if initially $a_2(0,\cdot) = 0$, since this amplitude will be generated during the course of time by the remaining two others. Analogously a_3 is generated by interaction of a_1 and a_2.

More formally we can also infer this fact from the closure condition, since any possible combination of k_2 and k_3 via the mapping Σ (with $\Lambda = 3$) yields either k_2, k_3, or any other value (like $2k_2 - k_3$) which is not in $\mathcal{G}_S^{(1)}$ by assumption (recall that the system of modes μ_1, μ_2, μ_3 is assumed to be closed of order $\Lambda = 3$). However, if one aims to follow the same argument for, *e.g.*, k_1 and k_2, the first equation in (2.5) shows that this sub-system is not closed.

2.4.2 The case of several pulses within one Bloch band

This is a particular situation where we keep the band index $\ell \in \mathbb{N}$ fixed and thus only consider the interaction of several pulses within a single Bloch band. Hence the solution to (1.1) takes the asymptotic form

$$u^\varepsilon(t,x) \sim \sum_{m=1}^{M} a_m(t,x)\chi_\ell\left(\frac{x}{\varepsilon};k_m\right)\mathrm{e}^{\mathrm{i}(k_m\cdot x - E_\ell(k_m)t)/\varepsilon} + \mathcal{O}(\varepsilon), \quad \ell \in \mathbb{N}.$$

In particular we can expect this description to be correct in cases where the Bloch band $E_\ell(k)$ admits only a moderate variation $\Delta_\ell := \mathrm{var}_{k\in\mathcal{B}}\, E_\ell(k)$ and is well separated from the rest of the spectrum of $H_\Gamma(k)$ by a sufficiently large gap, *i.e.*

$$\min\{|E_\ell(k) - E_n(k)| \,:\, n \in \mathbb{N},\ n \neq \ell\} \gg \Delta_\ell > 0.$$

It is easy to show then that a four-wave interaction can *always* be realized within such a band. Choose $k_{\max}$ and $k_{\min}$ such that $E_\ell(k_{\min}) \leq E_\ell(k) \leq E_\ell(k_{\max})$ for all $k \in \mathcal{B}$. We are looking for a triple (k_1, k_2, k_3) satisfying (2.5) with $\ell_j = \ell$, $j = 1,2,3$. To this end we note that $k_3 = 2k_1 - k_2$, modulo Γ^*, and define

$$e(k_1,k_2) \equiv 2E_\ell(k_1) - E_\ell(k_2) - E_\ell(2k_1 - k_2).$$

Then we have $e(k_{\max}, k_{\min}) > 0 > e(k_{\min}, k_{\max})$ and by a simple application of the intermediate-value theorem, we easily find k_1, k_2 with $k_1 \neq k_2$ and $e(k_1,k_2) = 0$.

2.4.3 The case of a single pulse decomposed into several bands

This is a second particular case, where we expect an asymptotic description for solutions to (1.1) given by

$$u^\varepsilon(t,x) \sim \sum_{\ell=1}^{L} a_\ell(t,x)\chi_\ell\left(\frac{x}{\varepsilon};k_0\right)\mathrm{e}^{\mathrm{i}(k_0\cdot x - E_\ell(k_0)t)/\varepsilon} + \mathcal{O}(\varepsilon), \quad k_0 \in \mathbb{R},$$

where k_0 is some given wave vector. One should have the following intuition: Given an initial plane wave of the form

$$u^\varepsilon_{\mathrm{in}}(x) = f(x)\mathrm{e}^{\mathrm{i}k_0\cdot x/\varepsilon},$$

one decomposes the slowly varying *macroscopic* amplitude $f(x)$ into a sum of terms, each of which is concentrated on a single Bloch band. This is possible due to Bloch's theorem, which ensures that $L^2(\mathbb{R}^d) = \bigoplus_{\ell=1}^\infty \mathcal{H}_\ell$, where $\mathcal{H}_\ell$ are the so-called band spaces. Strictly speaking though, one would require countably many terms in the decomposition, which we can take into account here. However for any practical purposes (and if $f(x)$ is sufficiently smooth and rapidly decaying) only the first few Bloch bands need to be taken into account as all higher bands give negligible contributions, *cf.* [16]. We shall not go into more details here on the precise definition of $\mathcal{H}_\ell$ etc. but refer to [2, 3, 23, 26] for more details on these classical results (see also [16] for a numerical approach).

Concerning the possible generation of resonant modes, it is clear that in this case the first condition in (2.4) is trivially fulfilled, since we are only dealing with a single wave vector k_0. Thus, it is merely a question on the precise structure of the bands $E_\ell(k_0)$ whether one can indeed expect resonances.

3 Formal derivation of the approximate solution

In the following we consider a finite set $\mathcal{S}$ of modes which is closed of order $\Lambda = 2N+1$, for some $N \in \mathbb{N}$, to be determined later.

For solutions of (1.1) we seek an asymptotic two-scale expansion of the following form

$$u_N^\varepsilon(t,x) := \sum_{n=0}^{N} \varepsilon^n \, v_n \left(t, x, \frac{t}{\varepsilon}, \frac{x}{\varepsilon}\right), \tag{3.1}$$

where

$$v_n(t,x,\tau,y) := \sum_{\sigma \in \mathcal{G}_\mathcal{S}^{(2n+1)}} A_{n,\sigma}(t,x,y) \mathbf{E}_\sigma(\tau, y), \tag{3.2}$$

with $\mathbf{E}_\sigma(\tau,y) := \mathrm{e}^{\mathrm{i}\sigma\cdot(y,-\tau)}$, for $\sigma \in \mathcal{G}_\mathcal{S}^{(2n+1)} \equiv \Sigma(\mathcal{S}^{2n+1})$. Note that in the most simple case, where $n=0$, we get that $\sigma = (k, E_\ell(k))$, with $(k,\ell) \in \mathcal{S}$, which yields $\mathbf{E}_\sigma(\tau,y) = \mathrm{e}^{\mathrm{i}(k\cdot y - E_\ell(k)\tau)}$, a simple plane wave. Moreover, if $\sigma_j \in \mathcal{G}_\mathcal{S}^{(2n_j+1)}$, for $j=1,2,3$, this yields

$$\mathbf{E}_{\sigma_1}\overline{\mathbf{E}}_{\sigma_2}\mathbf{E}_{\sigma_3} = \mathbf{E}_{\sigma_1-\sigma_2+\sigma_3}, \tag{3.3}$$

with $\sigma_1 - \sigma_2 + \sigma_3 \in \mathcal{G}_\mathcal{S}^{(2(n_1+n_2+n_3)+3)}$. Of course we also impose that

$$A_{n,\sigma}(\cdot,\cdot,y+\gamma) = A_{n,\sigma}(\cdot,\cdot,y), \quad \forall\, y \in \mathbb{R}^d,\ \gamma \in \Gamma.$$

3.1 The general strategy

As already said before, we shall only consider the case of simple eigenvalues $E_\ell(k)$, for simplicity. Plugging the ansatz (3.1), (3.2) into (1.1) and expanding in powers of ε, we formally obtain

$$\mathrm{i}\varepsilon\partial_t u_N^\varepsilon - H_{\mathrm{per}}^\varepsilon u_N^\varepsilon - \varepsilon\kappa |u_N^\varepsilon|^2 u_N^\varepsilon = \sum_{n=0}^{N} \varepsilon^n X_n + \mathrm{res}(u_N^\varepsilon),$$

with the residual

$$\mathrm{res}(u_N^\varepsilon) = \sum_{n=N+1}^{3N+1} \varepsilon^n X_n. \tag{3.4}$$

Introducing, for any (general) $\sigma \equiv (k,E) \in \mathcal{B}\times\mathbb{R}$, the operators

$$L_0^\sigma := E - H_\Gamma(k), \quad L_1^\sigma := \mathrm{i}\partial_t + \mathrm{i}k\cdot\nabla_x + \mathrm{div}_x\,\nabla_y, \quad L_2 := \frac{1}{2}\Delta_x \tag{3.5}$$

this yields

$$X_0 := \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(1)}} (L_0^\sigma A_{0,\sigma})\mathbf{E}_\sigma, \tag{3.6}$$

and also

$$\begin{aligned} X_1 := & \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(3)}} (L_0^\sigma A_{1,\sigma})\mathbf{E}_\sigma + \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(1)}} (L_1^\sigma A_{0,\sigma})\mathbf{E}_\sigma \\ & - \kappa \sum_{\substack{\sigma_j\in G_{\mathcal{S}}^{(1)}: \\ j=1,2,3}} A_{0,\sigma_1}\overline{A}_{0,\sigma_2}A_{0,\sigma_3}\mathbf{E}_{\sigma_1-\sigma_2+\sigma_3} \end{aligned} \tag{3.7}$$

where we have used the relation (3.3). In general we get for $n = 2, \dots, 3N+1$,

$$\begin{aligned} X_n := & \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(2n+1)}} (L_0^\sigma A_{n,\sigma})\mathbf{E}_\sigma + \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(2n-1)}} (L_1^\sigma A_{n-1,\sigma})\mathbf{E}_\sigma + \sum_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(2n-3)}} L_2 A_{n-2,\sigma}\mathbf{E}_\sigma \\ & - \kappa \sum_{n_1+n_2+n_3=n-1} \sum_{\substack{\sigma_j\in G_{\mathcal{S}}^{(2n_j+1)}: \\ j=1,2,3}} A_{n_1,\sigma_1}\overline{A}_{n_2,\sigma_2}A_{n_3,\sigma_3}\mathbf{E}_{\sigma_1-\sigma_2+\sigma_3}. \end{aligned} \tag{3.8}$$

Here one should note that $A_{n,\sigma}(t,x,y) \equiv 0$, for all $n \geq N+1$, by assumption.
Now we shall subsequently construct $A_{n,\sigma}$ such that $X_n \equiv 0$ for $n = 0, \dots, N$. To this end we have to compare *equal* coefficients of $\mathbf{E}_\sigma$. We consequently obtain, from (3.6)–(3.8), equations of the form

$$L_0^\sigma A_{n,\sigma} = F_{n,\sigma}, \quad \sigma \in \mathcal{G}_{\mathcal{S}}^{(2n+1)}, \tag{3.9}$$

where the r.h.s. $F_{n,\sigma}$ can be determined from the coefficients $(A_{m,\sigma})_{\sigma\in\mathcal{G}_{\mathcal{S}}^{(2m+1)}}$ for $m = 0, 1, \dots, n-1$. More precisely

$$\begin{aligned} F_{n,\sigma} := & - L_1^\sigma A_{n-1,\sigma} - L_2 A_{n-2,\sigma} \\ & + \kappa \sum_{n_1+n_2+n_3=n-1} \sum_{\substack{\sigma_j\in\mathcal{G}_{\mathcal{S}}^{(2n_j+1)}: \\ \sigma_1-\sigma_2+\sigma_3=\sigma}} A_{n_1,\sigma_1}\overline{A}_{n_2,\sigma_2}A_{n_3,\sigma_3}. \end{aligned} \tag{3.10}$$

Note that here the summation index has changed in comparison to (3.8). To proceed further we need to distinguish two possible cases:
Case I. On the one hand, for $\sigma \in \mathcal{G}_{\mathcal{S}}^{(2n+1)} \setminus \mathcal{G}_{\mathcal{S}}^{(1)}$ the closure condition up to order $2n+1$ implies invertibility of L_0^σ, *i.e.* $(L_0^\sigma)^{-1} \in \mathrm{Lin}(L^2(\mathcal{Y}), H^2(\mathcal{Y}))$ and we obtain

$$A_{n,\sigma}(t,x,y) = (L_0^\sigma)^{-1} F_{n,\sigma}(t,x,y). \tag{3.11}$$

The corresponding modes are called *non-resonant*.
Case II. On the other hand, if $\sigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$, then L_0^σ has a nontrivial kernel. In order to distinguish this case more prominently from the one above, we shall from now

on use the notation $\varsigma \equiv \sigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$, which also characterizes the basic resonant modes $\mu \in \mathcal{S}$ via $\varsigma = \Sigma(\mu)$.

Using the orthogonal projections $\mathbb{P}_\varsigma$ onto this kernel the necessary and sufficient solvability condition for (3.9) in this case is then given by

$$\mathbb{P}_\varsigma F_{n,\varsigma}(t,x,y) = 0, \tag{3.12}$$

which yields (recall that $\dim(\operatorname{ran}\mathbb{P}_\varsigma) = 1$, by assumption)

$$\langle \chi_\varsigma,\, F_{n,\varsigma}\rangle_{L^2(\mathcal{Y})} = 0, \quad \text{for } 0 \neq \chi_\varsigma \in \operatorname{ran}\mathbb{P}_\varsigma. \tag{3.13}$$

Under this condition, we consequently obtain

$$A_{n,\varsigma}(t,x,y) = a_{n,\varsigma}(t,x)\chi_\varsigma(y) + A^\perp_{n,\varsigma}(t,x,y) \tag{3.14}$$

with

$$A^\perp_{n,\varsigma} := (L_0^\varsigma)^{-1}(1-\mathbb{P}_\varsigma)F_{n,\varsigma}. \tag{3.15}$$

Note that here $a_{n,\varsigma}$ is still undetermined. However the condition (3.13) provides a partial differential equation for $a_{n-1,\varsigma}$, the so far undetermined part of the previous step.

Remark 3.1. In the case of non-simple eigenvalues $E_\ell(k)$, *i.e.* $\dim(\operatorname{ran}\mathbb{P}_\varsigma) = R > 1$, we can simply use a smooth orthonormal basis $\{\chi_{\varsigma,r}\}_{r=1}^R$ of $\operatorname{ran}\mathbb{P}_\varsigma$ and generalize the above given formulas (3.12), (3.14) accordingly.

3.2 Explicit calculations

In the following we shall determine the approximate solution in more detail, by following the above described strategy.

$n = 0$: We need to solve $X_0 \equiv 0$ and immediately note that in this case, Case I above is obsolete. Then, in Case II, equation (3.6) implies

$$L_0^\varsigma A_{0,\varsigma} = 0, \quad \varsigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$$

and thus (3.14) simplifies to

$$A_{0,\varsigma} = a_{0,\varsigma}(t,x)\chi_\varsigma(y), \tag{3.16}$$

with $a_{0,\varsigma}$ still to be determined.

$n = 1$: In the next step we have to solve $X_1 \equiv 0$. In Case I, *i.e.* for $\sigma \in \mathcal{G}_{\mathcal{S}}^{(3)} \setminus \mathcal{G}_{\mathcal{S}}^{(1)}$, the equations (3.7) and (3.11) imply

$$\begin{aligned} A_{1,\sigma} &= \kappa\, (L_0^\sigma)^{-1}\Big(\sum_{\substack{\varsigma_1,\varsigma_2,\varsigma_3 \in \mathcal{G}_{\mathcal{S}}^{(1)} \\ \sigma = \varsigma_1 - \varsigma_2 + \varsigma_3}} A_{0,\varsigma_1}\overline{A}_{0,\varsigma_2}A_{0,\varsigma_3}\Big) \\ &= \kappa \sum_{\substack{\varsigma_1,\varsigma_2,\varsigma_3 \in \mathcal{G}_{\mathcal{S}}^{(1)}: \\ \sigma = \varsigma_1 - \varsigma_2 + \varsigma_3}} a_{0,\varsigma_1}\overline{a}_{0,\varsigma_2}a_{0,\varsigma_3}\, (L_0^\sigma)^{-1}\left(\chi_{\varsigma_1}\overline{\chi}_{\varsigma_2}\chi_{\varsigma_3}\right), \end{aligned} \tag{3.17}$$

where for the second equality we simply insert (3.16). We proceed with Case II: To this end the solvability condition (3.12) for $\varsigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$ allows us to determine the so far still unknown $a_{0,\varsigma}$, obtained before. By (3.13), this yields

$$\int_{\mathcal{Y}} \overline{\chi}_\varsigma(y)\, F_{1,\varsigma}(t,x,y)\,\mathrm{d}y = 0, \tag{3.18}$$

where

$$F_{1,\varsigma} = L_1^\varsigma A_{0,\varsigma} - \kappa \sum_{\substack{\varsigma_1,\varsigma_2,\varsigma_3 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \varsigma = \varsigma_1 - \varsigma_2 + \varsigma_3}} A_{0,\varsigma_1} \overline{A}_{0,\varsigma_2} A_{0,\varsigma_3}. \tag{3.19}$$

From the definition of L_1^ς in (3.5) and using the following basic identity

$$\langle \chi_\ell,\, (-\mathrm{i}\nabla_y + k)\chi_\ell \rangle_{L^2(\mathcal{Y})} = \nabla_k E_\ell(k),$$

a straightforward calculation shows, *cf.* the appendix of [7], that (3.18) can be written as

$$\partial_t a_{0,\varsigma} + \vartheta_\varsigma \cdot \nabla_x a_{0,\varsigma} + \sum_{\substack{\varsigma_1,\varsigma_2,\varsigma_3 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \varsigma = \varsigma_1 - \varsigma_2 + \varsigma_3}} \mathrm{i}\kappa_{(\varsigma_1,\varsigma_2,\varsigma_3,\varsigma)} a_{0,\varsigma_1} \overline{a}_{0,\varsigma_2} a_{0,\varsigma_3} = 0, \tag{3.20}$$

where we denote

$$\kappa_{(\varsigma_1,\varsigma_2,\varsigma_3,\varsigma)} := \kappa \int_{\mathcal{Y}} \chi_{\varsigma_1}(y)\overline{\chi}_{\varsigma_2}(y)\chi_{\varsigma_3}(y)\overline{\chi}_\varsigma(y)\,\mathrm{d}y.$$

Since any $\varsigma \in \mathcal{G}_{\mathcal{S}}^{(1)}$ corresponds to a unique $m = 1,\dots,M$, via $\varsigma = \Sigma(\mu_m) = (k_m, E_{\ell_m}(k_m))$, with $\mu_m = (k_m, \ell_m) \in \mathcal{S}$, we can shortly write

$$a_{0,\varsigma}(t,x) \equiv a_m(t,x), \quad \chi_\varsigma(y) \equiv \chi_{\ell_m}(y;k_m), \quad \vartheta_\varsigma \equiv \vartheta_m.$$

Hence the amplitude equations (3.20) can be equivalently written in the form (1.5) used before.

In summary we have now fully determined the expressions (3.16) and (3.17), and from (3.14) we finally get that

$$A_{1,\varsigma}(t,x,y) = a_{1,\varsigma}(t,x)\chi_\varsigma(y) + (L_0^\varsigma)^{-1}(1-\mathbb{P}_\varsigma)F_{1,\varsigma}(t,x,y),$$

where again the coefficients $a_{1,\varsigma}$ are still arbitrary and have to be determined by the solvability condition for $n = 2$. Since this yields an initial value problem for $a_{1,\varsigma}$ (see below) we are free to choose its value at time $t = 0$. For simplicity we shall put $a_{1,\varsigma}(0,\cdot) = 0$.

$n \geq 2$: From here we proceed inductively, as described above, by solving $X_n \equiv 0$. The only difference that occurs is that in Case II the coefficients $a_{n,\varsigma}$ do not solve a nonlinear initial value problem, but rather a system of *linear, inhomogeneous* transport equations. Indeed, lengthy calculations show that

$$\begin{aligned}\partial_t a_{n,\varsigma} + \vartheta_\varsigma \cdot \nabla_x a_{n,\varsigma} + \sum_{\substack{\tilde\varsigma,\varsigma_1,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \tilde\varsigma-\varsigma_1+\varsigma_2=\varsigma}} 2\mathrm{i}\kappa_{(\tilde\varsigma,\varsigma_1,\varsigma_2,\varsigma)} a_{n,\tilde\varsigma}\overline{a}_{0,\varsigma_1} a_{0,\varsigma_2} \\ + \sum_{\substack{\varsigma_1,\tilde\varsigma,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \varsigma_1-\tilde\varsigma+\varsigma_2=\varsigma}} \mathrm{i}\kappa_{(\varsigma_1,\tilde\varsigma,\varsigma_2,\varsigma)} a_{0,\varsigma_1}\overline{a}_{n,\tilde\varsigma} a_{0,\varsigma_2} + \mathrm{i}\Theta_n = 0,\end{aligned} \tag{3.21}$$

with source term

$$\begin{aligned}\Theta_n := & -\mathbb{P}_\varsigma \left(L_1^\varsigma A_{n,\varsigma}^\perp + L_2 A_{n-1,\varsigma}\right) \\ & + \kappa\, \mathbb{P}_\varsigma \Big(\sum_{\substack{\tilde\varsigma,\varsigma_1,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \tilde\varsigma-\varsigma_1+\varsigma_2=\varsigma}} 2A_{n,\tilde\varsigma}^\perp \overline{A}_{0,\varsigma_1} A_{0,\varsigma_2} + \sum_{\substack{\varsigma_1,\tilde\varsigma,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)}:\\ \varsigma_1-\tilde\varsigma+\varsigma_2=\varsigma}} A_{0,\varsigma_1} \overline{A}_{n,\tilde\varsigma}^\perp A_{0,\varsigma_2} \Big) \\ & + \kappa\, \mathbb{P}_\varsigma \Big(\sum_{\substack{\varsigma_1,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)} \not\ni \sigma:\\ \sigma-\varsigma_1+\varsigma_2=\varsigma}} 2A_{n,\sigma} \overline{A}_{0,\varsigma_1} A_{0,\varsigma_2} + \sum_{\substack{\varsigma_1,\varsigma_2 \in \mathcal{G}_{\mathcal{S}}^{(1)} \not\ni \sigma:\\ \varsigma_1-\sigma+\varsigma_2=\varsigma}} A_{0,\varsigma_1} \overline{A}_{n,\sigma} A_{0,\varsigma_2} \Big) \\ & + \kappa\, \mathbb{P}_\varsigma \Big(\sum_{\substack{n_1+n_2+n_3=n:\\ n_j \le n-1}} \sum_{\substack{\sigma_j \in \mathcal{G}_{\mathcal{S}}^{(2n_j+1)}:\\ \sigma_1-\sigma_2+\sigma_3=\varsigma}} A_{n_1,\sigma_1} \overline{A}_{n_2,\sigma_2} A_{n_3,\sigma_3} \Big).\end{aligned}$$

Again we shall put $a_{n,\varsigma}(0,\cdot) = 0$ for simplicity, since we are free to choose the initial values for (3.21). Finally, we note that, in order to satisfy $X_N \equiv 0$ (*i.e.* in the last step), we do not need to determine the corresponding $a_{N,\varsigma}$ and thus we can impose $a_{N,\varsigma}(t,\cdot) \equiv 0$.

4 Justification of the amplitude equations

In order to obtain our main result, Theorem 4.5, we have to justify the above given formal calculations rigorously. In particular we need a nonlinear stability result on our approximation.

4.1 Estimates on the approximate solution and on the residual

In the above section we derived an approximate solution u_N^ε which admits an asymptotic expansion of the following form

$$
\begin{aligned}
u_N^\varepsilon(t,x) = &\sum_{\varsigma\in G_{\mathcal{S}}^{(1)}} a_{0,\varsigma}(t,x)\chi_\varsigma\left(\frac{x}{\varepsilon}\right)\mathbf{E}_\varsigma\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)\\
&+\sum_{n=1}^{N}\varepsilon^n\sum_{\varsigma\in G_{\mathcal{S}}^{(1)}}\left(a_{n,\varsigma}(t,x)\chi_\varsigma\left(\frac{x}{\varepsilon}\right)+A_{n,\varsigma}^{\perp}\left(t,x,\frac{x}{\varepsilon}\right)\right)\mathbf{E}_\varsigma\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)\\
&+\sum_{n=1}^{N}\varepsilon^n\sum_{\sigma\in G_{\mathcal{S}}^{(2n+1)}\setminus G_{\mathcal{S}}^{(1)}} A_{n,\sigma}\left(t,x,\frac{x}{\varepsilon}\right)\mathbf{E}_\sigma\left(\frac{t}{\varepsilon},\frac{x}{\varepsilon}\right)
\end{aligned}
\tag{4.1}
$$

where the first two lines on the r.h.s. include only resonant modes while the last one takes into account the generated non-resonant terms. To this end we required the appearing finite set of modes to be closed of order $\Lambda = 2N+1$.

To proceed further let us, for $s\in[0,\infty)$, introduce the scaled Sobolev spaces

$$H_\varepsilon^s := \left\{f^\varepsilon \in L^2(\mathbb{R}^d)\ ;\ \|f^\varepsilon\|_{H_\varepsilon^s}<\infty\right\},$$

with

$$\|f^\varepsilon\|_{H_\varepsilon^s}^2 := \int_{\mathbb{R}^d}(1+|\varepsilon p|^2)^s\,|\widehat{f}(p)|^2\,\mathrm{d}p.$$

Here $\widehat{f}$ denotes the usual Fourier transform of f in $L^2(\mathbb{R}^d)$. Note that in H_ε^s the following Gagliardo–Nirenberg inequality holds

$$\forall\, s>\frac{d}{2}\ \exists\, C_\infty>0:\quad \|\,f\,\|_{L^\infty}\le C_\infty\|\,f\,\|_{H^s}\le C_\infty\varepsilon^{-d/2}\|\,f\,\|_{H_\varepsilon^s}, \tag{4.2}$$

where the factor $\varepsilon^{-d/2}$ is easily obtained by scaling. In the following lemma we collect the a-priori estimates on u_N^ε needed in order to prove our main result.

Lemma 4.1. *For $d,N\in\mathbb{N}$ and $s\in[0,\infty)$ let $K=\max\{0,s+\frac{d}{2}-2\}$ and $S>N+s+\frac{d}{2}$, or, if $d=1$, $S\ge N+s+1$. Moreover, assume that $\partial^\alpha V_\Gamma\in L^\infty(\mathcal{Y})$, for all $|\alpha|\le K$, and let $(a_1,\dots,a_M)\in C^0([0,T),H^S(\mathbb{R}^d))^M$ be a solution of the amplitude equations* (1.5)*.*

Then, the approximate solution u_N^ε, given in (4.1)*, satisfies the following estimates. For each $T_*\in(0,T)$ there exist positive constants $C_a,C_b,C_r>0$, such that, for $\varepsilon\in(0,1)$, $|\alpha|\le s$, and $t\in[0,T_*]$, it holds*

$$\|\,(\varepsilon\partial_x)^\alpha u_N^\varepsilon(t,\cdot)\,\|_{L^\infty}\le C_a,\quad \|\,u_N^\varepsilon(t,\cdot)\,\|_{H_\varepsilon^s}\le C_b,\quad \|\,\mathrm{res}(u_N^\varepsilon)(t,\cdot)\,\|_{H_\varepsilon^s}\le C_r\varepsilon^{N+1}.$$

Proof. Since u_N^ε is given by (4.1), in order to prove the estimates of the lemma we need to establish the regularity of $A_{n,\sigma}$ for $n=0,\dots,N$, $\sigma\in\mathcal{G}_{\mathcal{S}}^{(2n+1)}$, which have been calculated formally in Section 3. Under the assumptions $(a_1,\dots,a_M)\in C^0([0,T),H^S(\mathbb{R}^d))^M$ and $\partial^\alpha V_\Gamma\in L^\infty(\mathcal{Y})$, for $|\alpha|\le K$, we shall first prove that it holds

$$A_{n,\sigma}\in C^0([0,T),H^{S-n}(\mathbb{R}^d,H^{K+2}(\mathcal{Y}))). \tag{4.3}$$

Indeed, by (3.16), $A_{0,\varsigma}(t,x,y)=a_{0,\varsigma}(t,x)\chi_\varsigma(y)$, where $a_{0,\varsigma}$ solves (3.20) and χ_ς is given by (2.2), and we get immediately (4.3) for $n=0$ (*cf.* Lemma 2.4 and Lemma 4.3 below).

From here it follows, on the one hand, that for $\sigma\in\mathcal{G}_{\mathcal{S}}^{(3)}\setminus\mathcal{G}_{\mathcal{S}}^{(1)}$, by (3.9) and (3.17), $F_{1,\sigma}=L_0^\sigma A_{1,\sigma}\in C^0([0,T),H^S(\mathbb{R}^d,H^{\tilde{s}}(\mathcal{Y})))$ with $\tilde{s}=K+2$ and hence, by (3.10), $A_{1,\sigma}\in C^0([0,T),H^{S-1}(\mathbb{R}^d,H^{\tilde{s}}(\mathcal{Y})))$. On the other hand, for $\varsigma\in\mathcal{G}_{\mathcal{S}}^{(1)}$ we obtain by (3.19) $F_{1,\varsigma}\in C^0([0,T),H^{S-1}(\mathbb{R}^d,H^{\tilde{s}}(\mathcal{Y})))$, and thus, by (3.15), $A_{1,\varsigma}^\perp\in C^0([0,T),H^{S-1}(\mathbb{R}^d,H^{\tilde{s}}(\mathcal{Y})))$. Moreover, $a_{1,\varsigma}\chi_\varsigma\in C^0([0,T),H^{S-1}(\mathbb{R}^d,H^{\tilde{s}}(\mathcal{Y})))$, since $a_{1,\varsigma}$ solves the linear inhomogeneous transport equation (3.21) (for $n=1$) with initial condition $a_{1,\varsigma}(0,\cdot)=0$, coefficients in $C^0([0,T),H^S(\mathbb{R}^d))$, and source term $\Theta_1\in C^0([0,T),H^{S-2}(\mathbb{R}^d))$. Hence, by (3.14), we obtain (4.3) for $n=1$. Proceeding inductively, we obtain (4.3) for all $n=0,\dots,N$ via the above given steps.

Having established (4.3), we aim to show the first estimate of the lemma. To this end we have to guarantee that for each $n=0,\dots,N$, $\sigma\in\mathcal{G}_{\mathcal{S}}^{(2n+1)}$, there exists a $C>0$, such that

$$\left\|(\varepsilon\partial_x)^\alpha A_{n,\sigma}\left(t,\cdot,\frac{\cdot}{\varepsilon}\right)\right\|_{L^\infty}\le C$$

holds true for all $\varepsilon\in(0,1)$, $|\alpha|\le s$, and $t\in[0,T_*]$. Using the Gagliardo–Nirenberg inequality (4.2), we therefore require

$$(\varepsilon\partial_x)^\alpha A_{n,\sigma}\left(t,\cdot,\frac{\cdot}{\varepsilon}\right)\in H^m(\mathbb{R}^d)\quad\text{with } m>\frac{d}{2}, \tag{4.4}$$

for all $|\alpha|\le s$, i.e.

$$A_{n,\sigma}\left(t,\cdot,\frac{\cdot}{\varepsilon}\right)\in H_\varepsilon^{m^*}(\mathbb{R}^d)\quad\text{with } m^*>s+\frac{d}{2}.$$

Thus, by (4.3), we need $m^*=S-n+\tilde{s}>s+\frac{d}{2}$, that is, either $S-n>|\beta|+\frac{d}{2}$ and $\tilde{s}\ge s-|\beta|$, or $S-n\ge|\beta|$ and $\tilde{s}>s-|\beta|+\frac{d}{2}$, for all $|\beta|\le s$ and all $n=0,\dots,N$. It turns out that in order to satisfy either of these conditions, the optimal conditions we have to impose concerning the regularity of $A_{n,\sigma}$, given in (4.3), are $S-n>s+\frac{d}{2}$ and $\tilde{s}=K+2>s+\frac{d}{2}$, for $n=0,\dots,N$, which directly yields the assumptions on S and K stated in the lemma.

The second estimate of the lemma then follows immediately by (4.4). In order to obtain the third estimate, having in mind the definitions (3.4) and (3.8) with $A_{n,\sigma}\equiv 0$ for $n\ge N+1$ and $a_{N,\varsigma}\equiv 0$ (*cf.* the note after (3.21)), it is necessary and sufficient to assure additionally that $\Delta a_{N-1,\varsigma}\in H^s(\mathbb{R}^d)$, i.e., $a_{0,\varsigma}\in H^{N+s+1}(\mathbb{R}^d)$. This yields the additional condition $S\ge N+s+1$, and concludes the proof. □

Remark 4.2. Note that in order to derive the above given a-priori estimates for u_N, the required regularity imposed on χ_ς is indeed independent of N.

We shall also need the following result on the linear time-evolution.

Lemma 4.3. *For $\varepsilon \in (0,1)$ denote the unitary propagator corresponding to the linear Hamiltonian $H^\varepsilon_{\text{per}}$, defined in* (1.3), *by*

$$U^\varepsilon(t) := \mathrm{e}^{-\mathrm{i}H^\varepsilon_{\text{per}} t/\varepsilon}.$$

For $s \in [0,\infty)$ assume $\partial^\alpha V_\Gamma \in L^\infty(\mathcal{Y})$ for all $|\alpha| \leq \max\{0, s-2\}$. Then there exists a $C_l > 0$, such that

$$\| U^\varepsilon(t) f^\varepsilon \|_{H^s_\varepsilon} \leq C_l \|f^\varepsilon\|_{H^s_\varepsilon} \quad \textit{for all } t \in \mathbb{R}. \tag{4.5}$$

Proof. Recalling the definition of $H^\varepsilon_{\text{per}}$, given in (1.3), we clearly have the basic L^2 estimate,

$$\|f^\varepsilon(t)\|_{L^2} \equiv \|U^\varepsilon(t) f^\varepsilon\|_{L^2} = \|f^\varepsilon\|_{L^2}, \quad \text{for all } t \in \mathbb{R},$$

since $H^\varepsilon_{\text{per}}$ is self-adjoint. Moreover, since $U^\varepsilon(t)$ obviously commutes with (its generator) $H^\varepsilon_{\text{per}}$, so does any power of the latter and we therefore obtain

$$\|(H^\varepsilon_{\text{per}})^{s/2} f^\varepsilon(t)\|_{L^2} = \|(H^\varepsilon_{\text{per}})^{s/2} f^\varepsilon\|_{L^2}, \quad \text{for all } s, t \in \mathbb{R}.$$

Without loss of generality we assume here that $H^\varepsilon_{\text{per}} \geq 1$, otherwise we may add a constant independent of $\varepsilon \in (0,1)$.
For $s = 1$ this is nothing but energy conservation. Using the conservation of the L^2 norm and the condition $V_\Gamma \in L^\infty(\mathcal{Y})$ we immediately obtain the desired result (4.5) for $s = 1$ with $C_l = (1+4\|V_\Gamma\|_\infty)^{1/2}$.
For general $s \in \mathbb{N}\backslash\{1\}$, we use integration by parts and the assumed regularity of V_Γ to find ε-independent constants $C_1, C_2 > 0$ such that

$$C_1\|(H^\varepsilon_{\text{per}})^{s/2} f^\varepsilon(t)\|_{L^2} \leq \|f^\varepsilon(t)\|_{H^s_\varepsilon} \leq C_2\|(H^\varepsilon_{\text{per}})^{s/2} f^\varepsilon(t)\|_{L^2},$$

i.e. we have equivalence of the norms uniformly in ε. For general $s \geq 0$ the same holds true by interpolation. In summary we obtain boundedness of the unitary group $U^\varepsilon(t)$ in all H^s_ε and the assertion of the lemma is proved. ☐

4.2 Stability of the approximation

First let us recall the following Moser-type lemma, *cf.* [22, Lemma 8.1], which we shall use in the proof of Theorem 4.5 below.

Lemma 4.4. *Let $R > 0$, $s \in [0,\infty)$, and $\mathcal{N}(z) = \kappa|z|^2 z$ with $\kappa \in \mathbb{R}$. Then there exists a $C_s = C_s(R, s, d, \kappa) > 0$ such that if*

$$\|(\varepsilon\partial)^\alpha f\|_{L^\infty} \leq R, \ \forall\, |\alpha| \leq s \textit{ and } \|g\|_{L^\infty} \leq R,$$

then

$$\|\mathcal{N}(f+g) - \mathcal{N}(f)\|_{H^s_\varepsilon} \leq C_s \|g\|_{H^s_\varepsilon}.$$

Proof. The proof can be found in [22] for $s \in \mathbb{N}$ and follows by interpolation for general $s \in [0, \infty)$. $\square$

With the above results at hand, we are now able to establish our main result, which justifies rigorously the validity of the amplitude equations (1.5), describing the macroscopic dynamics of M modulated pulses for a closed mode system of order $\Lambda = 2N + 1$ (with N depending on d).

Theorem 4.5. *For $d \in \mathbb{N}$ choose $s, S \in [0, \infty)$ and $N \in \mathbb{N}$ such that $N, s > \frac{d}{2}$ and $S > N + s + \frac{d}{2}$, or, if $d = 1$, $S \geq N + s + 1$. Let $K = \max\{0, s + \frac{d}{2} - 2\}$ and assume $\partial^\alpha V_\Gamma \in L^\infty(\mathcal{Y})$ for all $|\alpha| \leq K$. Moreover, let the finite system of modes $\mathcal{S} = \{\mu_1, \dots, \mu_M : M \in \mathbb{N}\}$ be closed of order $\Lambda = 2N + 1$.*
Then for any solution $(a_1, \dots, a_M) \in C^0([0, T), H^S(\mathbb{R}^d))^M$ of the amplitude equations (1.5), and any $t_* \in (0, T)$, $\beta \in (\frac{d}{2}, N]$, $c > 0$, there exist an $\varepsilon_0 \in (0, 1)$ and a $C > 0$, such that for all $\varepsilon \in (0, \varepsilon_0)$, the approximate solution u_N^ε, constructed above, and any exact solution u^ε of (1.1) with

$$\| u^\varepsilon(0, \cdot) - u_{N-1}^\varepsilon(0, \cdot) \|_{H_\varepsilon^s} \leq c\varepsilon^\beta,$$

satisfy

$$\| u^\varepsilon(t, \cdot) - u_{N-1}^\varepsilon(t, \cdot) \|_{H_\varepsilon^s} \leq C\varepsilon^\beta \quad \textit{for all } t \in [0, t_*].$$

Remark 4.6. Note, that the approximate solution u_N^ε contains terms up to order $\mathcal{O}(\varepsilon^N)$, whereas the above given error estimates include only u_{N-1}^ε. The reason for this is that for our proof we need to work with u_N^ε but still the obtained error is only of the order $\mathcal{O}(\varepsilon^\beta)$ with $\beta \leq N$. We can therefore eventually neglect the last term in u_N^ε, or, loosely speaking, we can move it to the right hand side of the above given estimate.

Proof. We write the exact solution of (1.1) in the form

$$u^\varepsilon(t, x) = u_N^\varepsilon(t, x) + \varepsilon^\beta w^\varepsilon(t, x)$$

and denote $\varrho^\varepsilon(t) := \| w^\varepsilon(t) \|_{H_\varepsilon^s}$. Then, clearly, $\varrho^\varepsilon(0) \leq c$, and we will show that there exist $C > 0$, $\varepsilon_0 \in (0, 1)$, such that, for all $\varepsilon \in (0, \varepsilon_0]$, it holds $\varrho^\varepsilon(t) \leq C$ for $t \in [0, t_*]$. Inserting this ansatz into (1.1), written as

$$\mathrm{i}\partial_t u^\varepsilon = -\frac{\varepsilon}{2}\Delta u^\varepsilon + \frac{1}{\varepsilon} V_\Gamma\left(\frac{x}{\varepsilon}\right) u^\varepsilon + \mathcal{N}(u^\varepsilon), \quad \text{where } \mathcal{N}(z) = \kappa\, |z|^2 z,$$

and applying Duhamel's formula we get

$$\begin{aligned} w^\varepsilon(t) = U^\varepsilon(t) w^\varepsilon(0) + \varepsilon^{-\beta} \int_0^t U^\varepsilon(t - \tau) \left(\mathcal{N}(u_N^\varepsilon(\tau) + \varepsilon^\beta w^\varepsilon(\tau)) - \mathcal{N}(u_N^\varepsilon(\tau))\right) \mathrm{d}\tau \\ - \varepsilon^{-(\beta+1)} \int_0^t U^\varepsilon(t - \tau) \operatorname{res}(u_N^\varepsilon(\tau)) \,\mathrm{d}\tau, \end{aligned}$$

where $\operatorname{res}(u_N^\varepsilon)$ is defined in equation (3.4). Hence, using Lemma 4.1 and Lemma 4.3 to estimate the residual and the linear semi-group, respectively, we obtain

$$\varrho^\varepsilon(t) \leq \; C_l c + C_l C_r \varepsilon^{N-\beta} t + \varepsilon^{-\beta} \int_0^t C_l \big\| \mathcal{N}(u_N^\varepsilon(\tau) + \varepsilon^\beta w^\varepsilon(\tau)) - \mathcal{N}(u_N^\varepsilon(\tau)) \big\|_{H_\varepsilon^s} \, \mathrm{d}\tau,$$

since $\varrho^\varepsilon(0) \leq c$, by assumption. Using $N \geq \beta$ and $\varepsilon \in (0,1)$, we consequently obtain

$$\varrho^\varepsilon(t) \leq \; C_l(c + C_r t_*) + \varepsilon^{-\beta} \int_0^t C_l \big\| \mathcal{N}(u_N^\varepsilon(\tau) + \varepsilon^\beta w^\varepsilon(\tau)) - \mathcal{N}(u_N^\varepsilon(\tau)) \big\|_{H_\varepsilon^s} \, \mathrm{d}\tau,$$

for $t \leq t_*$.

Now, we set $C := C_l(c + C_r t_*) \mathrm{e}^{C_l C_s t_*}$ and choose a $D > \max\{c, C\}$. Then, since $D > c \geq \varrho^\varepsilon(0)$ and $\varrho^\varepsilon(t)$ is continuous, there exists, for every $\varepsilon \in (0,1)$, a positive time $t_D^\varepsilon > 0$, such that $\varrho^\varepsilon(t) \leq D$ for $t \leq t_D^\varepsilon$.

The Gagliardo–Nirenberg inequality (4.2) yields, for $s > d/2$, that

$$\| \, \varepsilon^\beta w^\varepsilon(t) \, \|_{L^\infty} \leq \varepsilon^{\beta - d/2} C_\infty D \qquad \text{for } t \leq t_D^\varepsilon.$$

Hence, using $\beta - d/2 > 0$ there exists an $\varepsilon_0 \in (0,1)$, such that

$$\| \, \varepsilon^\beta w^\varepsilon(t) \, \|_{L^\infty} \leq C_a, \quad \text{for } \varepsilon \in (0, \varepsilon_0] \text{ and } t \leq t_D^\varepsilon.$$

Moreover, by Lemma 4.1 we have $\|(\varepsilon\partial)^\alpha u_N^\varepsilon(t) \, \|_{L^\infty} \leq C_a$ for $|\alpha| \leq s$, $\varepsilon \in (0,1)$, and $t < T$. Thus, we can apply Lemma 4.4 (with $R = C_a$) in order to estimate the nonlinear term and obtain

$$\varrho^\varepsilon(t) \leq \; C_l(c + C_r t_*) + C_l C_s \int_0^t \varrho^\varepsilon(\tau) \, \mathrm{d}\tau, \quad \text{for } \varepsilon \in (0, \varepsilon_0] \text{ and } t \leq t_D^\varepsilon.$$

Gronwall's lemma then yields

$$\varrho^\varepsilon(t) \leq \; C_l(c + C_r t_*) \mathrm{e}^{C_l C_s t} \leq C, \quad \text{for } \varepsilon \in (0, \varepsilon_0] \text{ and } t \leq t_*.$$

Since $C < D$, we conclude that the assumptions needed in order to apply Lemma 4.4 are fulfilled for all $\varepsilon \in (0, \varepsilon_0]$ and all $t \leq t_*$, that is $t_D^\varepsilon \geq t_*$. Thus the above given estimate proves that

$$\| \, u^\varepsilon(t, \cdot) - u_N^\varepsilon(t, \cdot) \, \|_{H_\varepsilon^s} = \mathcal{O}(\varepsilon^\beta), \quad \text{for } t \in [0, t_*].$$

However, since $N \geq \beta$ and

$$\| \, u_N^\varepsilon(t, \cdot) - u_{N-1}^\varepsilon(t, \cdot) \, \|_{H_\varepsilon^s} = \mathcal{O}(\varepsilon^N), \quad \text{for } t \in [0, t_*],$$

we can finally replace $u_N^\varepsilon(t, \cdot)$ by $u_{N-1}^\varepsilon(t, \cdot)$ in our stability result, which consequently proves the assertion of the theorem. $\square$

We want to stress that our stability result is an advancement when compared to the result of [7], in the sense that our asymptotic estimates no longer suffer from a loss of accuracy in powers of ε (as has been the case in the cited work). We infer in particular, that our nonlinear setting allows for the same kind of stability result as one would expect in the linear case.

Theorem 4.5 also shows that no blow-up can occur in the exact solution u^ε within the interval $t \in [0, T)$, determined by the existence time of (1.5). Hence, if the system (1.5) indeed admits global-in-time solutions we deduce that the solutions u^ε, starting close to such modulated pulses, exist for arbitrary long times (but blow-up may occur after the modulational structure is lost on longer time scales.)

Finally note that for $d = 1$ the assumption $N > d/2$ and the required closure of order $\Lambda = 2N + 1$ imply that, at least, $N = 1$ and thus $\Lambda = 3$, whereas for $d = 3$ spatial dimensions we need at least $N = 2$ and hence $\Lambda = 5$. This increase of the required Λ for higher spatial dimensions d can be relaxed though, as we shall show in the following section.

5 The case of higher-order resonances

In Section 3, we derived the approximate solution u_N^ε under the assumption that $\mathcal{S}$ satisfies a closure condition of sufficient high order Λ, *i.e.* we required $\Lambda = 2N + 1$ and $N > d/2$. We shall show now that this closure condition can be significantly relaxed if we generalize the ansatz (3.1), (3.2) slightly by allowing for a larger set of modes $\mathcal{G}_{2N+1}$. To this end we consider the new ansatz

$$u_N^\varepsilon(t, x) = \sum_{n=0}^{N} \sum_{\sigma \in \mathcal{G}_{2n+1}} \varepsilon^n A_{n,\sigma}(t, x, y) \mathbf{E}_\sigma(\tau, y), \tag{5.1}$$

where the set of modes $\mathcal{G}_{2N+1}$ is then defined inductively as follows: Again we start from $\mathcal{S} = \{(k_1, \ell_1), ..., (k_M, \ell_M)\}$, which is now assumed to be *only* closed of order $\Lambda = 3$ and we set $\mathcal{G}_1 = \mathcal{G}_\mathcal{S}^{(1)}$. However, for $n \geq 2$ we will allow the sets $\mathcal{G}_{2n+1}$ to be larger than $\mathcal{G}_\mathcal{S}^{(2n+1)}$. The reason for this enlargement is motivated by the fact, that

$$\widetilde{\mathcal{G}}_\mathcal{S}^{(2n+1)} = \mathcal{G}_\mathcal{S}^{(2n+1)} \cap \left(\mathcal{G} \setminus \mathcal{G}_\mathcal{S}^{(1)}\right)$$

may be nonempty for some $n \geq 2$. Then, the corresponding equation (3.9) might not be solvable. To circumvent this problem, we set

$$\mathcal{G}_{2n-1} = \mathcal{G}_\mathcal{S}^{(2n-1)} \cup \widetilde{\mathcal{G}}_\mathcal{S}^{(2n+1)}.$$

At stage $n-1$ the corresponding $A_{n-1,\sigma}$ with $\sigma \in \widetilde{\mathcal{G}}_\mathcal{S}^{(2n+1)}$ take the usual product form $a_{n-1,\sigma}(t, x)\chi_\sigma(y)$ with $\chi_\sigma \in \ker L_0^\sigma \subset H^2(\mathcal{Y})$ and free coefficients $a_{n-1,\sigma}$. Then, in step n the corresponding equations for $\sigma \in \widetilde{\mathcal{G}}_\mathcal{S}^{(2n+1)}$ can be solved as in Case II of Subsection 3.1, *i.e.* we obtain a linear transport equation for the free coefficients from the solvability condition for the previous step.

To make this more precise, we associate with the sequence $\mathcal{G}_1, \mathcal{G}_3, \ldots, \mathcal{G}_{2N+1}$ another sequence $\check{\mathcal{G}}_3, \check{\mathcal{G}}_5, \ldots, \check{\mathcal{G}}_{2N+1}$ as follows

$$\check{\mathcal{G}}_{2n+1} = \bigcup_{\substack{n_1,n_2,n_3=0,\ldots,n-1\\ n_1+n_2+n_3=n-1}} \Big(\mathcal{G}_{2n_1+1} - \mathcal{G}_{2n_2+1} + \mathcal{G}_{2n_3+1}\Big). \tag{5.2}$$

Hence, $\check{\mathcal{G}}_{2n+1}$ contains all modes of order $\mathcal{O}(\varepsilon^n)$ that can be generated by the cubic nonlinearity $|u|^2u$ from the already given terms of lower order.

Definition 5.1. A mode set $\mathcal{S} = \{(k_1, \ell_1), ..., (k_M, \ell_M)\}$ is called *weakly closed of order* $2N+1$, if there exists a sequence $\mathcal{G}_1 \subset \mathcal{G}_3 \subset \cdots \subset \mathcal{G}_{2N+1} \subset \mathcal{B} \times \mathbb{R}$ such that the following conditions hold:

(i) $\mathcal{G}_{2n+1}$ is finite for $n = 0, 1, ..., N$;

(ii) $\mathcal{G}_1 = \mathcal{G}_{\mathcal{S}}^{(1)} = \Sigma(\mathcal{S})$;

(iii) $\check{\mathcal{G}}_{2n+1} \subset \mathcal{G}_{2n+1}$ for $n = 1, ..., N$;

(iv) if $\sigma \in \mathcal{G}_{2n+1} \cap \mathcal{G}$ with $n \in \mathbb{N}$, then either $\sigma \in \mathcal{G}_{2n-1}$ or $\sigma \notin \check{\mathcal{G}}_{2n+1}$;

(v) for each $\sigma \in \mathcal{G}_{2N+1} \cap \mathcal{G}$ the group velocity $\vartheta_\sigma = \nabla_k E_\ell(k)$ exists,

where $\check{\mathcal{G}}_3 \subset \cdots \subset \check{\mathcal{G}}_{2N+1}$ is the sequence associated with $\mathcal{G}_1, ..., \mathcal{G}_{2N+1}$ according to (5.2).

Condition (iv) means that any occurring resonant mode must either occur already in an earlier step and hence has a corresponding free coefficient or it appears for the first time but it is not yet generated by the nonlinear interaction. This will become clearer in the following examples.

Example 5.2. We first show that a system which is closed of order $2N+1$ is also weakly closed of order $2N+1$. For this we simply set $\mathcal{G}_{2n+1} = \mathcal{G}_{\mathcal{S}}^{(2n+1)}$. By construction, the associated sequence is $\check{\mathcal{G}}_{2n+1} = \mathcal{G}_{2n+1}$ and the conditions (i) to (iii) hold immediately. Moreover, (iv) follows since $\mathcal{G}_{2n+1} \cap \mathcal{G} = \mathcal{G}_1 = \mathcal{G}_{\mathcal{S}}^{(1)}$.

Example 5.3. We now show that a weak closure of order Λ is in general a weaker condition than closure of order Λ is. For this consider the case $\mathcal{S} = \{(k_1, \ell_1), (k_2, \ell_2)\}$ such that $\mathcal{G}_1 = \mathcal{G}_{\mathcal{S}}^{(1)} = \{\sigma_1, \sigma_2\} \subset \mathcal{G}$ with $\sigma_j = (k_j, E_{\ell_j}(k_j))$. An easy induction argument gives

$$\mathcal{G}_{\mathcal{S}}^{(2n+1)} = \{\, (n{+}1{-}j)\sigma_1 + (j{-}n)\sigma_2 \;;\; j = 0, 1, \ldots, 2n+1 \,\}$$

Now assume that $\mathcal{G} \cap \mathcal{G}_{\mathcal{S}}^{(2N^*+1)} \subset \{\sigma_1, \sigma_2, 3\sigma_1{-}2\sigma_2\}$ for some $N^* \geq 2$, *i.e.* we have a closure of order 3 but not of order 5.

We claim that a weak closure of order $2N+1$ still holds for any N such that $3N \leq 2N^*$. For $n = 0, \ldots, N$ we let

$$\mathcal{G}_{2n+1} = \{\, j\sigma_1 + (1{-}j)\sigma_2 \;;\; j = -\beta_n, \ldots, \alpha_n \,\},$$

with $\alpha_n = [3(n+1)/2]$ and $\beta_n = [3n/2]$, where $[\cdot]$ denotes the integer part. By construction of the sequences $(\alpha_n, \beta_n)_n$ we thus have $\mathcal{G}_3 = \check{\mathcal{G}}_3 \cup \{3\sigma_1 - 2\sigma_2\}$ and $\mathcal{G}_{2n+1} = \check{\mathcal{G}}_{2n+1}$ for $n = 2, \ldots, N$. Hence conditions (i) to (iii) of Definition 5.1 are fulfilled. Condition (iv) also holds since $\mathcal{G}_{2n+1} \cap \mathcal{G} \subset \mathcal{G}_1 \cup \{3\sigma_1 - 2\sigma_2\}$.

Repeating the construction of Section 3.2 analogously we obtain the following result.

Lemma 5.4. *Let $\mathcal{S} = \{(k_1, \ell_1), \ldots, (k_M, \ell_M)\}$ be a set of modes that is closed of order $\Lambda = 3$ and weakly closed of order $\Lambda = 2N+1$ for some $N \in \mathbb{N}$. Then an approximation*

$$u_N^\varepsilon(t,x) = \sum_{n=0}^{N} \varepsilon^n \sum_{\sigma \in \mathcal{G}_{2n+1}} A_{n,\sigma}(t,x,y)\mathbf{E}_\sigma \tag{5.3}$$

can be constructed as in Sections 3.1 and 3.2, where $\mathcal{G}_1 \subset \cdots \mathcal{G}_{2N+1}$ is the sequence guaranteed by Definition 5.1.

Hence, via the same steps as in Section 4 one can easily obtain a justification of the amplitude equations (1.5) under these relaxed resonance conditions.

Corollary 5.5. *Let $d, N \in \mathbb{N}$ and $\mathcal{S}$ be a mode system that is closed of order $\Lambda = 3$ and weakly closed of order $\Lambda = 2N+1$. Then, under the same assumptions as before the statement of Theorem 4.5 holds analogously with the approximate solution given by* (5.3).

A Hamiltonian structure and conservation laws of the amplitude equations

Recalling the amplitude equations in the general form (1.5), *i.e.*

$$\mathrm{i}\partial_t a_m = -\mathrm{i}\vartheta_m \cdot \nabla_x a_m + \sum_{\substack{p,q,r=1: \\ \Sigma(\mu_p,\mu_q,\mu_r)=\Sigma(\mu_m)}}^{M} \kappa_{(p,q,r,m)} a_p \overline{a}_q a_r, \tag{A.1}$$

for $m = 1, \ldots, M$, we want to highlight the Hamiltonian structure of this equation as well as to deduce the conserved quantities which follow from the specific structure of the resonances. For a general method on how to derive the *reduced Hamiltonian structure* and the first integrals for (A.1) we refer to [11, 12]. Here we just collect the results that are obtained with the general theory developed there.

We first note that (A.1) is a Hamiltonian system generated by

$$\mathcal{H}^{\mathrm{red}}(a) = \int_{\mathbb{R}^d} \sum_{m=1}^{M} \mathrm{Im}\,(\overline{a}_m\, \vartheta_m \cdot \nabla a_m) + \sum_{\substack{p,q,r=1: \\ \Sigma(\mu_p,\mu_q,\mu_r)=\Sigma(\mu_m)}}^{M} \frac{\kappa_{(p,q,r,m)}}{2}\, a_p \overline{a}_q a_r \overline{a}_m \,\mathrm{d}x,$$

and the symplectic two-form $\mathrm{i} = \sqrt{-1}$. Hence, (A.1) takes the form

$$\mathrm{i}\partial_t a_m = \delta_{\overline{a}_m} \mathcal{H}^{\mathrm{red}}(a).$$

Note that the reduced Hamiltonian $\mathcal{H}^{\mathrm{red}}(a)$ is not obtained as the lowest order expansion of the original Hamiltonian $\mathcal{H}^\varepsilon(u^\varepsilon)$ of the full system, which reads

$$\mathcal{H}^\varepsilon(u^\varepsilon) = \int_{\mathbb{R}^d} \frac{\varepsilon^2}{2}|\nabla u|^2 + V_\Gamma\left(\frac{x}{\varepsilon}\right)|u^\varepsilon|^2 + \frac{\varepsilon\kappa}{2}|u|^4\,\mathrm{d}x.$$

Indeed, inserting the ansatz $u^\varepsilon = u^\varepsilon_N$, as given in (4.1), into $\mathcal{H}^\varepsilon(u^\varepsilon)$ we find, as $\varepsilon \to 0$, that

$$\mathcal{H}^\varepsilon(u^\varepsilon_N) = \mathcal{I}(a) + \mathcal{O}(\varepsilon), \quad \text{with } \mathcal{I}(a) = \int_{\mathbb{R}^d} \sum_{m=1}^M E_{\ell_m}(k_m)|a_m|^2\,\mathrm{d}x.$$

Even though not a Hamiltonian, $\mathcal{I}(a)$ clearly is a conserved quantity for the amplitude system (A.1).

However, the energy levels $E_{\ell_m}(k_m) \equiv \omega_m$ do not occur explicitly in the amplitude system (A.1). Hence, any choice of $\widetilde{\omega}_m \in \mathbb{R}$ that is compatible with the resonance conditions (2.4) leads to additional conserved quantities. More precisely, if $\widetilde{\omega}_1, \dots, \widetilde{\omega}_M$ are chosen such that for any $p, q, r, m \in \{1, \dots, M\}$ the resonance conditions (2.4) imply the identity

$$\widetilde{\omega}_p - \widetilde{\omega}_q + \widetilde{\omega}_r = \widetilde{\omega}_m,$$

then

$$\widetilde{\mathcal{I}}(a) = \int_{\mathbb{R}^d} \sum_{m=1}^M \widetilde{\omega}_m |a_m|^2\,\mathrm{d}x$$

defines a first integral.

A particular choice is $\widetilde{\omega}_1 = \dots = \widetilde{\omega}_M = 1$, which leads to the trivial fact that the L^2 norm is preserved. The latter can of course also be deduced from the fact that the L^2 norm was preserved in the original problem for u^ε or from the fact that the mode system is invariant under the phase shifts $(a_1, \dots, a_M) \mapsto (\mathrm{e}^{\mathrm{i}\alpha}a_1, \dots, \mathrm{e}^{\mathrm{i}\alpha}a_M)$, with $\alpha \in \mathbb{R}$. Similarly, $\widetilde{\mathcal{I}}(a)$ can be understood as a first integral with respect to suitably chosen phase shifts that are compatible with the resonance structure, *i.e.*

$$\widetilde{T}_\alpha : (a_1, \dots, a_M) \mapsto (\mathrm{e}^{\mathrm{i}\widetilde{\omega}_1\alpha}a_1, \dots, \mathrm{e}^{\mathrm{i}\widetilde{\omega}_M\alpha}a_M),.$$

Finally, it should be mentioned that the mode system (A.1) is translation invariant. This provides first integrals associated with the translation operators in the coordinate directions

$$\mathcal{I}^{\mathrm{trans}}_\theta(a) = \int_{\mathbb{R}^d} \sum_{m=1}^M \mathrm{Im}\left(\overline{a}_m\,\theta \cdot \nabla a_m\right)\mathrm{d}x, \quad \theta \in \mathbb{R}^d.$$

So far, we are not able to show that these conserved quantities are enough to provide a global existence result. Note in particular, that for the original problem the case $\kappa > 0$ leads to an energy that is definite and allows us to conclude global existence for the nonlinear Schrödinger equation (1.1). For the mode system (A.1) the sign of κ is no longer helpful, since the term involving derivatives is indefinite. Note also that global existence for the mode system cannot be inferred from global existence of (1.1), since we can not expect the solutions u^ε to remain in the form of modulated pulses for all times.

References

[1] G. Bartal, O. Manela, and M. Segev, *Spatial Four Wave Mixing in Nonlinear Periodic Structures*, Phys. Rev. Lett. **97** (2006), 073906–073910.

[2] E. I. Blount, *Formalisms of band theory*, Solid State Physics **13**, Academic Press, New York, 305–373 (1962).

[3] A. Bensoussan, J. L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland Pub. Co. (1978).

[4] F. Bloch, *Über die Quantenmechanik der Elektronen in Kristallgittern*, Z. Phys. **52** (1928), 555–600.

[5] K. Busch, G. Schneider, L. Tkeshelashvili, and H. Uecker, *Justification of the Nonlinear Schrödinger equation in spatially periodic media*, ZAMP to appear.

[6] G. K. Campbell, J. Mun, M. Boyd, E. W. Streed, W. Ketterle, and D. E. Pritchard, *Parametric amplification of atoms* Phys. Rev. Lett. **96** (2006), 020406–020410.

[7] R. Carles, P. Markowich, and C. Sparber, *Semi-classical asymptotics for weakly nonlinear Bloch waves*, J. Stat. Phys. **117** (2004), no.1/2, 369–400.

[8] D. Choi and Q. Niu, *Bose-Einstein condensates in an optical lattice*, Phys. Rev. Lett. **82** (1999), 2022–2025.

[9] B. Deconinck, B. Frigyik, and J. Nathan Kutz, *Dynamics and stability of Bose-Einstein condensates: the nonlinear Schrödinger equation with periodic potential*, J. Nonlinear Sci. **12** (2002), 169–205.

[10] J. Giannoulis, *Three-wave interaction of modulated pulses in discrete lattices*, in preparation.

[11] J. Giannoulis, M. Herrmann, and A. Mielke, *Continuum descriptions for the dynamics in discrete lattices: derivation and justification*, in: Analysis, Modeling and Simulation of Multiscale Problems, A. Mielke, ed., Springer, Heidelberg, 2006.

[12] J. Giannoulis, M. Herrmann, and A. Mielke, *Lagrangian and Hamiltonian two-scale reduction*, in preparation.

[13] J. Giannoulis and A. Mielke, *Dispersive evolution of pulses in oscillator chains with general interaction potentials*, Discr. Cont. Dynam. Systems Ser. B **6** (2006), 493–523.

[14] J. C. Guillot, J. Ralston, and E. Trubowitz, *Semiclassical asymptotics in solid-state physics*, Comm. Math. Phys. **116** (1998), 401–415.

[15] K. M. Hilligsoe and K. Molmer, *Phase-matched four wave mixing and quantum beam splitting of matter waves in a periodic potential*, Phys. Rev. A **71** (2005), 041602–041606.

[16] Z. Huang, S. Jin, P. Markowich, and C. Sparber, *A Bloch decomposition based split-step pseudo spectral method for quantum dynamics with periodic potentials*, SIAM J. Sci. Comput. to appear.

[17] J. Hunter, A. Majda, and R. Rosales, *Resonantly interacting weakly nonlinear hyperbolic waves II: several space variables*, Stud. Appl. Math **75** (1986), 187–226.

[18] J. Joly, G. Metivier, and J. Rauch, *Recent results in nonlinear geometric optics*, Internat. Series Num. Math. 130, Birkhäuser (1999).

[19] M. Kraemer, C. Menotti, L. Pitaevskii, and S. Stringari, *Bose-Einstein condensates in 1D optical lattices: compressibility, Bloch bands and elementary excitations*, Eur. Phys. J. D **27** (2003), 247–263.

[20] A. Majda and R. Rosales, *Resonantly interacting weakly nonlinear hyperbolic waves I: a single space variable*, Stud. Appl. Math **71** (1984), 149–179.

[21] G. Panati, H. Spohn, and S. Teufel, *Effective dynamics for Bloch electrons: Peierls substitution and beyond*, Comm. Math. Phys. **242** (2003), 547–578.

[22] J. Rauch, *Lectures on Nonlinear Geometrical Optics*, IAS/Park City Math. Series 5, AMS (1999).

[23] M. Reed, B. Simon, *Methods of modern mathematical physics IV. Analysis of operators*, Academic Press (1978).

[24] C. Sparber, *Effective mass theorems for nonlinear Schrödinger equations*, SIAM J. Appl. Math. **66** (2006), issue 3, 820–842.

[25] C. Sulem and P. L. Sulem, *The nonlinear Schrödinger equation*, Applied Math. Sciences 139, Springer (1999).

[26] C. H. Wilcox, *Theory of Bloch waves*, J. d'Anal. Math. **33** (1978), 146–167.