Low regularity justification results for envelope approximations of nonlinear wave packets in periodic media

Karsten Matthies*,
Hannes Uecker†

June 18, 2016

Abstract

We consider a cubic nonlinear wave equation with highly oscillating cubic coefficient and wave packet initial data. Using a regularization step of the initial data, we give a low regularity justification of the Nonlinear Schrödinger equation as the envelope equation.

1 Introduction

Amplitude equations are an ubiquitous tool to describe complicated physical systems modeled by partial differential equations. A phenomenon of particular interest is the propagation of wave packets, e.g. light pulses in dispersive media. This has been addressed in various physical settings, i.e. in linear and nonlinear media with or without oscillating coefficients. A non-exhaustive list of work on amplitude equations for wave packet propagation is [Kal88, KSM92, Sch98, SU01, GS01, MN02, SW04, SU03, Sch05, BSTU06, SU07]. In this amplitude formalism (AF), first by a multiple scaling ansatz the amplitude equation is derived, which typically takes the form of a (nonlinear) Schrödinger equation with constant so called effective coefficients. Then, in a second step, the amplitude equation is justified by estimates for the error between solutions of the full system and the approximation by the AF. This is also related to certain WKB methods (in geometric and diffractive optics, see, e.g., [Lan98, AP06, Car07]), but since we restrict to one space dimension and to almost monochromatic waves (no explicit leading order phase modulation in the initial conditions) we prefer the more specific name amplitude formalism.

Our purpose here is to improve the AF to regularity assumptions on the data that are lower than in previous results. We consider the cubic Klein–Gordon equation

$$\varepsilon^2\partial_t^2 u_\varepsilon - \partial_x^2 u_\varepsilon + \varepsilon^{-2}u_\varepsilon + c_\varepsilon u_\varepsilon^3 = 0, \tag{1}$$

where $u_\varepsilon = u_\varepsilon(t,x) \in \mathbb{R}, t \geq 0, x \in \mathbb{R}$, and where c_ε is highly oscillating with period ε, i.e., $c_\varepsilon(x) := c(x/\varepsilon)$ with small ε and $c \in L^2((0,1)$ extended via $c(y+1) = c(y)$ for $y \in \mathbb{R}$. This can

*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom, k.matthies@bath.ac.uk

†Institut für Mathematik, Universität Oldenburg, D-26111 Oldenburg, hannes.uecker@uni-oldenburg.de

be generalized in various ways, for instance to oscillating coefficients for the linear terms, but for simplicity we stick to (1). On the other hand, the oscillating nonlinear coefficient c_ε gives some averaging effects which go beyond the case of constant c; see Remark 1.1 for further comments.

The conserved energy for (1) is

$$E(t) = E(u_\varepsilon(t)) := \int \frac{1}{2}\Big[(\varepsilon^2\partial_t u_\varepsilon(t,x))^2 + (\varepsilon\partial_x u_\varepsilon(t,x))^2 + u_\varepsilon^2(t,x)\Big] + \frac{\varepsilon^2}{4}c_\varepsilon u_\varepsilon^4(t,x)\,\mathrm{d}x. \tag{2}$$

If for simplicity we restrict to the case $c(y) \geq 0$ for all $y \in [0,1)$, then we immediately obtain the uniform a priori–estimate

$$\|u_\varepsilon(t)\|_E^2 := \varepsilon^4\|\partial_t u_\varepsilon(t)\|^2 + \varepsilon^2\|\partial_x u_\varepsilon(t)\|^2 + \|u_\varepsilon(t)\|^2 \leq 2E(t) = 2E(0) \text{ for all } t \tag{3}$$

in the energy norm $\|u_\varepsilon\|_E$, where here and in the following always $\|\cdot\| = \|\cdot\|_{L^2(\mathbb{R})}$, i.e.,

$$\|u_\varepsilon(t)\|^2 = \int u_\varepsilon^2(t,x)\mathrm{d}x.$$

From (3) we obtain global existence of solutions of (1). However, without the sign condition on $c(y)$, which we do not assume in this paper, (3) is wrong, and good estimates for $\|u_\varepsilon(t)\|_E$ depend on the class of initial conditions (IC).

We consider IC in the form of wave–packets

$$\begin{aligned} u_\varepsilon(0,x) &= v_0(x)e_1(0,x/\varepsilon) + \varepsilon g_1(x) + \bar{v}_0(x)e_{-1}(0,x/\varepsilon) + \varepsilon\bar{g}_1(x), \\ \partial_t u_\varepsilon(0,x) &= \varepsilon^{-2}v_1(x)e_1(0,x/\varepsilon) + \varepsilon^{-1}g_2(x,x/\varepsilon) + \varepsilon^{-2}\bar{v}_1(x)e_{-1}(0,x/\varepsilon) + \varepsilon^{-1}\bar{g}_2(x,x/\varepsilon), \end{aligned} \tag{4}$$

with $e_{\pm 1}(\tau,y) = \exp(\pm\mathrm{i}(k_0 y - \omega_0\tau)$, $y = x/\varepsilon, \tau = t/\varepsilon^2$, $v_0, g_1 \in H^1(\mathbb{R})$, $v_1, g_2 \in L^2(\mathbb{R})$, and where v_1 is related to v_0 in a certain way, specified in Theorems 2.3 and 2.5 below. Here $^-$ denotes the complex conjugate, from now on we will use the shorthand notation c.c. to denote the complex conjugate of the preceding terms. The spatial wave number $k_0 \in \mathbb{R}$ and the temporal wavenumber ω_0 are related by the dispersion relation, which for (1) takes the simple form

$$\omega^2 = k^2 + 1. \tag{5}$$

On a the level of formal asymptotic expansions it is well known, e.g., [dSS88], that (1) has approximate solutions of the form

$$u_\varepsilon(t,x) = \psi_v(t,x) = v(t, x - \nu t/\varepsilon)e_1(t/\varepsilon^2, x/\varepsilon) + \text{c.c.}, \tag{6}$$

where $\nu = \partial_k\omega(k_0)$ (the group velocity) and $v(t,x)$ fulfills the Nonlinear Schrödinger equation (NLS)

$$\partial_t v = \frac{\mathrm{i}}{2}\partial_k^2\omega_0\partial_x^2 v - \frac{3\mathrm{i}}{2\omega_0}c^*|v|^2 v, \quad c^* = \int_0^1 c(y)\mathrm{d}y. \tag{7}$$

As our main results we give justifications of (7) with low regularity requirements on v_0. For $v_0 \in H^2(\mathbb{R})$, letting $u_\varepsilon(x,t) = \psi_v(x,t) + \varepsilon r(x,t)$ we prove, on an $\mathcal{O}(1)$ time scale, the error estimate

$$\|r(t)\|_E \leq C. \tag{8}$$

For $v_0 \in H^s(\mathbb{R}), 1 < s < 2$, letting $u_\varepsilon(x,t) = \psi_v(x,t) + \varepsilon^{s/2} r(x,t)$ we show that

$$\|r(t)\|_{H^1_\varepsilon} \leq C \tag{9}$$

with the scaled H^1 norm $\|u\|_{H^1_\varepsilon(\mathbb{R})} := \sqrt{\int_{\mathbb{R}} u^2(x) + \varepsilon^2(\partial_x u(x))^2 \mathrm{d}x}$. See Theorems 2.3 and 2.5 for the precise results. In particular, these also yield long time existence for (1) for the specific IC without assumptions on the sign of c. The reduced regularity assumptions are achieved by smooth approximations of the initial conditions via cut–off of Fourier modes of wave number k with $|k| > n$, and balancing of errors: the approximation error is small for large n, while the error estimates of the amplitude formalism will grow in n. We balance the growth in n with some power of ε, track the errors and choose an appropriate $n(\varepsilon)$ to achieve the desired estimates.

Remark 1.1 a) As already mentioned, many related results are known in the literature, see, e.g., [Kal88, SU01, BSTU06, SU07] and the references therein. Our model (1),(4) is a simple extension of the constant coefficient cubic Klein–Gordon equation, and the use of c_ε instead of a constant $c \in \mathbb{R}$ allows to illustrate some averaging effects in the effective coefficients. Moreover, it allows to illustrate a slight technical improvement (see Remark 2.1) which avoids the diagonalization of linear operators for instance used in [BSTU06] for the definition of improved approximations. On the other hand, contrary to, e.g., [BSTU06, APR09, APR11], we do not consider periodic coefficients in the linear part because we want to avoid the Bloch wave machinery needed in this case; the results given here can be extended to this case, but this requires some technical effort.

b) A more complicated version of (1) contains quadratic terms, e.g.,

$$\varepsilon^2 \partial_t^2 u_\varepsilon - \partial_x^2 u_\varepsilon + \varepsilon^{-2} u_\varepsilon + \varepsilon^{-1} b_\varepsilon u_\varepsilon^2 + c_\varepsilon u_\varepsilon^3 = 0, \tag{10}$$

where $b_\varepsilon(x) := b(x/\varepsilon)$, b sufficiently smooth with $b(y+1) = b(y)$. A priori estimates and justification results for this quadratic case are typically obtained via normal form transforms [Sha85, Sch98, Sch05, BSTU06], and require certain non–resonance conditions and a careful handling of the regularity loss in the normal form transforms. We believe that the methods from this paper can be transferred to the quadratic case, but for now restrict to the simpler cubic case.

c) Other standard versions of (1), (4) are obtained from setting

$$\tilde{u}_\varepsilon(\tau, \xi) = \varepsilon u_\varepsilon(\varepsilon^2 \tau, \varepsilon \xi), \tag{11}$$

which yields

$$\partial_\tau^2 \tilde{u}_\varepsilon = \partial_\xi^2 \tilde{u}_\varepsilon - \tilde{u}_\varepsilon - b(\xi)\tilde{u}_\varepsilon^2 - c(\xi)\tilde{u}_\varepsilon^3, \tag{12}$$

$$\tilde{u}_\varepsilon(0,\xi) = \varepsilon v_0(\varepsilon\xi)\tilde{e}_1(0,\xi) + \text{c.c.}, \quad \partial_\tau \tilde{u}_\varepsilon(0,\xi) = v_1(\varepsilon\xi)\tilde{e}_1(0,\xi) + \text{c.c.} \tag{13}$$

with $\tilde{e}_1(\tau,\xi) = \mathrm{e}^{\mathrm{i}(k_0\xi - \omega_0\tau)}$. Note that (12) does not explicitly depend on ε, which makes this scaling somewhat more natural, but the ε–dependence of the initial conditions justifies the subscript ε in $\tilde{u}_\varepsilon$. The advantage of (1), (4) is that it is somewhat closer to the underlying physics: x, t and x/ε, t/ε^2 are called the macroscopic and microscopic scales, respectively, and consequently (7) is called the macroscopic equation.

d) A related class of problem, see, e.g., [MNO02, MN02] and the references therein, is given by the case of spatially non–oscillatory initial data, corresponding to $k_0 = 0$ in (4), for (typically) constant coefficient and gauge invariant nonlinear Klein–Gordon equations with $x \in \mathbb{R}^d$. In this case, the solutions formally decompose into fast oscillations in time modulated by slow envelopes in time and space, which again can be described via solutions of NLS equations, and [MNO02, MN02] give low regularity approximation results for this situation, which would essentially correspond to $v_0 \in H^1(\mathbb{R})$ (and $k_0 = 0$) in (4). While our proofs below use pointwise in time a priori estimates on the difference between the formal approximation and the solution u_ε in some energy norms, the proofs in [MNO02, MN02] are based on Strichartz estimates, which are not easily available in our case of highly oscillatory initial data ($k_0 \neq 0$) and non gauge invariant nonlinearity. $\rfloor$

Remark 1.2 Another approach to derive and justify effective equations for problems with rapidly varying data is the method of two-scale convergence (2SC). The 2SC method for wave equations is based on the same formal calculation as the amplitude formalism, and on an a–priori estimate like (3). On the linear level, terms involving the solution u_ε then have good convergence properties within the 2SC, and this allows to *derive and justify* a limiting equation in one step, for initial data $v_0 \in H^1(\mathbb{R})$. See, e.g., [APR09], where this has been carried out for a linear wave equation with $x \in \mathbb{R}^d$, which moreover on the linear level is much more general than (1).

A justification of (7) for (1) for $v_0 \in H^1(\mathbb{R})$ using methods from 2SC would be an important achievement, and an improvement of, or at least an alternative to, our results (8) and (9). However, 2SC is considerably harder for nonlinear equations, as so called strong 2SC is needed to obtain some information on nonlinear expressions involving u_ε. In [Spa06], 2SC methods have been combined with multiple scales expansions to prove justification results for the homogenization of nonlinear Schrödinger equations with a large rapidly varying potential and highly regular initial data. The linear Schrödinger case with $v_0 \in H^1$ is considered in [AP05, AP06], and in [All08, Remark 7.5] it is claimed that the results of [AP05, All08] for the linear case in some special cases generalize to the nonlinear case. However, it appears that sufficient information to treat nonlinear terms in the framework of 2SC is not easily available for solutions of (1), and we failed to make 2SC methods work for (1),(4) without directly using the results (8) or (9), which in particular means that again we cannot deal with $v_0 \in H^1 \setminus H^s, s > 1$. $\rfloor$

2 The justification results

Our derivation and justification of (7) proceeds in two steps. First we plug the ansatz

$$\begin{aligned} u_\varepsilon(t,x) &= \psi_v(t,x) + \text{h.o.t.} := v(t,\xi)e_1(t/\varepsilon^2, x/\varepsilon) + \text{c.c.} + \text{h.o.t.}, \\ \xi &= x - \nu t/\varepsilon, \quad e_j = \mathrm{e}^{\mathrm{i}j(k_0 y - \omega_0 \tau)}, \end{aligned} \tag{14}$$

into (1), where $y = x/\varepsilon$, $\tau = t/\varepsilon^2$, and h.o.t. denotes at this point unspecified higher order (in ε) terms. Sorting with respect to ε then yields a hierarchy of equations, to be successively solved. All $\mathcal{O}(\varepsilon^{-2})$ terms vanish due to the dispersion relation $\omega_0^2 = k_0^2 + 1$, all $\mathcal{O}(\varepsilon^{-1})$ terms vanish by the choice $\nu = \partial_k \omega(k_0)$, and at $\mathcal{O}(\varepsilon^0)$ we obtain

$$[-2\mathrm{i}\omega_0\partial_t v + (\nu^2-1)\partial_\xi^2 v + 3c(x/\varepsilon)|v|^2 v]e_1 + c(x/\varepsilon)v^3 e_3 + \text{c.c.} = 0. \tag{15}$$

Clearly, the splitting into harmonics e_1 and e_3 in (15) is not self–evident, unless c is constant. However, based on the separation of scales between the arguments $\xi = x - \nu t/\varepsilon$ of v and $y = x/\varepsilon$ of c, the standard procedure is to average the coefficients of e_1 in y over the periodicity cell $(0,1)$ of c, i.e., to require

$$\int_0^1 (-2\mathrm{i}\omega_0\partial_t v + (\nu^2-1)\partial_\xi^2 v + 3c(y)|v|^2 v)\mathrm{d}y = -2\mathrm{i}\omega_0\partial_t v + (\nu^2-1)\partial_\xi^2 v + 3c^*|v|^2 v \stackrel{!}{=} 0, \qquad (16)$$

which is the NLS (7) for v. Thus, the $\mathcal{O}(\varepsilon^0 e_1)$ terms vanish in an averaged sense, i.e., at $\mathcal{O}(\varepsilon^0)$ the so called residual has the form

$$\mathrm{Res}_0(\psi_v) := (c(x/\varepsilon) - c^*)|v|^2 v e_1 + c(x/\varepsilon)v^3 e_3. \qquad (17)$$

The complete residual is

$$\mathrm{Res}(\psi_v) = \varepsilon^2\partial_t^2\psi_v - \partial_x^2\psi_v + \varepsilon^{-2}\psi_v + c_\varepsilon\psi_v^3 = \mathrm{Res}_0 + \varepsilon\mathrm{Res}_1,$$

and it is easy to see that under mild conditions on v the terms Res_1 are $\mathcal{O}(1)$ bounded in natural norms (see below).

Given a solution v of the NLS, the question is whether the ansatz (14) gives an approximation of a solution of (1) on an $\mathcal{O}(1)$ time–scale, which is the natural time–scale for the NLS evolution. In the context of the AF this problem of *justification* is usually dealt with as follows. Defining the (scaled) error $\rho(x,t)$ via

$$u_\varepsilon(x,t) = \psi_v(x,t) + \varepsilon\rho(x,t),$$

we want to use a–priori estimates on the error equation

$$\varepsilon^2\partial_t^2\rho = \partial_x^2\rho - \varepsilon^{-2}\rho - f, \quad f = f(t,x) = c_\varepsilon(3\psi_v^2\rho + 3\varepsilon\psi_v\rho^2 + \varepsilon^2\rho^3) - \varepsilon^{-1}\mathrm{Res}(\psi_v), \qquad (18)$$

to show that ρ stays $\mathcal{O}(1)$ bounded in a suitable norm, on an $\mathcal{O}(1)$ time scale. In the energy–norm we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\rho\|_E^2 = 2\varepsilon^2\int(\partial_t\rho)f\mathrm{d}x \le 2\varepsilon^2\|\partial_t\rho\|_{L^2}\|f\|_{L^2} \le 2\|\rho\|_E\|f\|_{L^2} \qquad (19)$$

and using a Gronwall argument it would be sufficient to have

$$\|\varepsilon^{-1}\mathrm{Res}(\psi_v)\|_{L^2} = \mathcal{O}(1) \qquad (20)$$

to have $\|\rho\|_E$ bounded. However, already from (17) we see that (20) does not hold in general if v is $\mathcal{O}(1)$ in $L^2 \cap L^\infty$, say.

2.1 The improved residual

Our first main idea is to use an improved approximation in the form

$$\phi_v(t,x) := ve_1(t/\varepsilon^2, x/\varepsilon) + \varepsilon^2\left[a_1(x/\varepsilon)e_1(t/\varepsilon^2, x/\varepsilon)|v|^2 v + a_3(x/\varepsilon)e_3(t/\varepsilon^2, x/\varepsilon)v^3\right] + \mathrm{c.c.}, \quad (21)$$

with a_j to be determined. First we note that if a_j is bounded in $H^1((0,1))$ and $\sup_{t\in[0,t_0]}\|v(t,\cdot)\|_{H^2} \le C_1$, then we have

$$\|\phi_v - \psi_v\|_E \le C\varepsilon^2, \qquad (22)$$

such that by the triangle inequality it is sufficient to estimate $r = \varepsilon^{-1}(u_\varepsilon - \phi_v)$.

Plugging ϕ_v into (1) we require, at $\mathcal{O}(\varepsilon^0)$, using (16),

$$[L_1 a_1(y) + c(y) - c^*] e_1 |v|^2 v + [L_3 a_3(y) + c(y)] e_3 v^3 + \text{c.c.} = 0, \tag{23}$$

where $L_1 a = [-\partial_y^2 - 2\mathrm{i}k_0 \partial_y] a$ and $L_3 a = [-(3\omega_0)^2 + \omega^2(3k_0) - 6\mathrm{i}k_0\partial_y - \partial_y^2] a$ are elliptic operators on $L^2((0,1))$. Thus, we need to solve

$$L_1 a_1 = -c + c^* \quad \text{and} \quad L_3 a_3 = -c, \tag{24}$$

where $c \in L^2_{\text{per}}((0,1))$. Even though L_1 has the one dimensional kernel spanned by constants, here we obtain $a_1 \in H^2_{\text{per}}((0,1))$ with $\|a_1\|_{H^2((0,1))} \leq C\|c(\cdot) - c^*\|_{L^2((0,1))}$ since

$$\langle c(\cdot) - c^* \rangle := \int_0^1 c(y) - c^* \mathrm{d}y = 0.$$

Similarly using a Fourier representation, we see that $L_3 : H^2((0,1)) \to L^2((0,1))$ is invertible due to the non-vanishing imaginary part of the symbol for non-constant modes and due to the nonresonance condition $\omega(3k_0)^2 - (3\omega_0)^2 \neq 0$ for the constant term, hence we have $\|a_3\|_{H^2((0,1))} \leq C\|c\|_{L^2((0,1))}$.

Remark 2.1 If c is constant, then $a_1 = 0$ and $a_3 = \dfrac{c}{\omega(3k_0)^2 - (3\omega_0)^2}$, and such relations have been used since at least [Kal88, KSM92] to achieve small residuals. On the other hand, for non-constant c, equations related to (24) have often been solved via diagonalization in Fourier (or Bloch) space, see, e.g., [BSTU06]. The ansatz (21) instead of the less specific version $\phi_v(t,x) := ve_1 + \text{c.c.} + \varepsilon^2 a(x,t,x/\varepsilon,t/\varepsilon^2)$ gives the splitting (24) and thus avoids this diagonalization. $\rfloor$

Our main idea is to carefully keep track of terms of low and high orders of derivatives in the residual, and the associated ε orders. This is aimed at approximating $v \in C([0,t_0], H^m(\mathbb{R}))$, $1 < m \leq 2$, by sequences (v_n) with $v_n \in C([0,t_0], H^4(\mathbb{R}))$, and trading some powers of ε for regularity.

Lemma 2.2 *Let $c \in L^2((0,1))$, $k_0 \in \mathbb{R}$. There exists a $C_3 > 0$ such that for all solutions $v \in C([0,t_0], H^4(\mathbb{R}))$ of (7) and all $t \in [0,t_0]$ we have*

$$\|\varepsilon^{-1}\text{Res}(\psi_{v(t)})\|_{L^2} \leq C_3(\varepsilon\|v(t)\|_{H^4} + \varepsilon^2\|v(t)\|^2_{H^2} + 1)(\|v(t)\|_{H^1} + \|v(t)\|^9_{H^1}). \tag{25}$$

Proof. In the following we drop the notation of complex conjugate terms and suppress the explicit t dependence of v in the estimates, i.e., for instance, $\|v\|_\infty = \|v(t)\|_\infty$. The term involving the highest derivatives in $\text{Res}(\phi_v(t))$ is

$$\varepsilon^2(\partial_t^2 v)e_1 = \varepsilon^2 \partial_t \left[\frac{\mathrm{i}}{2}\partial_k^2\omega_0\partial_x^2 v - \frac{3\mathrm{i}}{2\omega_0}c^*|v|^2 v\right] e_1 = \varepsilon^2\left[\frac{\mathrm{i}}{2}\partial_k^2\omega_0\partial_x^2\partial_t v - \frac{3\mathrm{i}}{2\omega_0}c^*(2|v|^2\partial_t v + v^2\overline{v}_t)\right] e_1,$$

where we replaced $\partial_t v$ by the rhs of (7). Doing this once more we obtain a fourth derivative as the highest spatial derivative of v, i.e.,

$$\begin{aligned}\varepsilon^2(\partial_t^2 v)e_1 =& \varepsilon^2\left[-\frac{1}{4}(\partial_k^2\omega_0)^2\partial_x^4 v + \varepsilon^2\frac{3}{4\omega_0}\partial_k^2\omega_0 c^*\partial_k^2\omega_0\partial_x^2(|v|^2 v)\right] e_1 \\ &- \varepsilon^2\frac{3\mathrm{i}}{2\omega_0}c^*\left[-|v|^2\mathrm{i}(\frac{\partial_k^2\omega_0}{2}\partial_x^2 v + \frac{3}{\omega_0}c^*|v|^2 v) + v^2\mathrm{i}(\frac{\mathrm{i}}{2}\partial_k^2\omega_0\partial_x^2\overline{v} + \frac{3}{2\omega_0}c^*|v|^2\overline{v})\right] e_1.\end{aligned}$$

Thus, by using $\|v\|_{L^\infty} \le C\|v\|_{H^1}$ and $\|\partial_x v\|_{L^\infty} \le C\|v\|_{H^2}$ we have

$$\varepsilon^2\|(\partial_t^2 v)e_1\|_{L^2} \le C\varepsilon^2\left[\|v\|_{H^4} + \|v\|_{H^2}\|v\|_{H^1}^2 + \|v\|_{H^1}^5\right]. \tag{26}$$

The other terms involving ∂_t^2 are $\varepsilon^4(a_1(y)e_1\partial_t^2(|v|^2v) + a_3(y)e_3\partial_t^2 v^3)$, and these can be estimated in a similar fashion to obtain

$$\varepsilon^4\|(a_1(y)e_1\partial_t^2(|v|^2v) + a_3(y)e_3\partial_t^2 v^3)\|_{L^2} \le \varepsilon^4\left[\|v\|_{H^4} + (\|v\|_{H^2}+1)^2\right]C(\|v\|_{H^1} + \|v\|_{H^1}^7). \tag{27}$$

The remaining terms involving $\partial_t v$ are of the form $\varepsilon^2\left[a_1(y)e_1\omega_0\partial_t(|v|^2v) + a_3(y)e_3 3\omega_0\partial_t(v^3)\right]$, such that their L^2-norm can be estimated by

$$\varepsilon^2(\|v\|_{H^2}+1)^2 C(\|v\|_{H^1} + \|v\|_{H^1}^5). \tag{28}$$

The lowest order terms in ε are due to residuals of the form

$$\varepsilon^2\varepsilon^{-1}\left[(a_1'(y) + \mathrm{i}k_0 a_1(y))e_1\partial_x(|v|^2v) + (a_3'(y) + 3\mathrm{i}k_0 a_3(y))e_3\partial_x v^3\right],$$

which can be estimated by

$$\varepsilon^2\varepsilon^{-1}\|a_1'(y)e_1\partial_x(|v|^2v) + a_3'(y)e_3\partial_x v^3\|_{L^2} \le \varepsilon C(\|v\|_{H^1} + \|v\|_{H^1}^3). \tag{29}$$

All other terms involving $\partial_x v$ are of order ε^2 and in a similar fashion can be estimated by

$$\varepsilon^2 C(\|v\|_{H^1} + \|v\|_{H^1}^3)(1 + \|v\|_{H^2}). \tag{30}$$

All further terms do not involve any derivatives of v and are of order ε^2 or higher. Their L^2-norm can be estimated by $\varepsilon^2 C(\|v\|_{H^1} + \|v\|_{H^1}^9)$. Collecting the above estimates and estimating the linear terms in $\|v\|_{H^2}$ by $\|v\|_{H^4}$ yields (25). $\square$

2.2 The error estimates

We now give two theorems that estimate the error under weak regularity conditions on v_0. The first one, with $v_0 \in H^2(\mathbb{R})$ yields the expected scaling $u = \psi_v + \varepsilon\rho$ with $\sup_{t\in[0,t_0]}\|\rho(t)\|_E \le C$.

Theorem 2.3 *Let $c \in L^2((0,1))$. For all $C_1 > 0$ and $t_0 > 0$ there exist $\varepsilon_0, C_2 > 0$ such that for all solutions $v \in C([0,t_0], H^2(\mathbb{R}))$ of (7) with $\sup_{t\in[0,t_0]}\|v(t,\cdot)\|_{H^2} \le C_1$, and all $\varepsilon \in (0,\varepsilon_0)$ the following holds. If*

$$u_\varepsilon(0,x) = \psi_v(0,x) + \varepsilon g_1(x, x/\varepsilon) \text{ and } \partial_t u_\varepsilon(0,x) = \frac{\mathrm{d}}{\mathrm{d}t}\psi_v(0,x) + \varepsilon^{-1}g_2(x,x/\varepsilon), \tag{31}$$

where $g_1, \varepsilon^{-1}\partial_x g_1, \partial_y g_1, g_2$ are bounded by C_1 in $L^2(\mathbb{R})$, then there exist a unique mild solution $u_\varepsilon \in C([0,t_0], H^1(\mathbb{R})) \cap C^1([0,t_0], L^2(\mathbb{R}))$ of (1) with initial conditions $u_\varepsilon(0,x)$ and $\partial_t u_\varepsilon(0,x)$, which can be written as $u_\varepsilon(t,x) = \psi_v(t,x) + \varepsilon\rho(t,x)$ with

$$\sup_{0\le t\le t_0}\|\rho\|_E \le C_2. \tag{32}$$

Remark 2.4 a) In 1D, given $t_0 > 0$ and $v_0 \in H^m(\mathbb{R})$, $m \geq 1$, the associated solutions of the NLS fulfill $v \in C([-t_0, t_0], H^m(\mathbb{R}))$ with $\|v\|_{H^m} \leq C_m(t_0, \|v_0\|_{H^m})$. The constant C_m is independent of t_0 for $m = 1$: In the defocussing case $c^* > 0$ this follows directly from the conservation of the coercive Hamiltonian $H(v) = \frac{\partial_k^2 \omega_0}{4} \int_{\mathbb{R}} v_x^2 \mathrm{d}x + \frac{3c^*}{8\omega_0} \int_{\mathbb{R}} |v|^4 \mathrm{d}x$. For the focussing case we can use the mass conversation $\|v(t))\|_{L^2} = \text{const}$ and the Gagliardo-Nierenberg estimate (see (37) below) to bound the negative part from below by $-C\|v\|_{H^1}\|v\|_{L^2}^3$, which is enough to give a uniform bound for $\|v\|_{H^1}$. The general statement then follows with a result on the preservation of regularity, e.g. [Tao06, Prop. 3.11]. The result is applicable as the H^1 bound implies uniform L^∞ estimates. Note that the constant C_m will in general grow exponentially in t_0 for $m > 1$. From [Tao06, Prop. 3.8] we obtain local Lipschitz continuity with respect to initial conditions, i.e., $\|v_1(t) - v_2(t)\|_{H^m} \leq L(t_0, \|v_1(0)\|_{H^m}, \|v_2(0)\|_{H^m})\|v_1(0) - v_2(0)\|_{H^m}$ for two solutions with IC $v_1(0)$ and $v_2(0)$.

b) Given Lemma 2.2, a naive condition to close the a–priori estimate (19) would be $v_0 \in H^4$, and our main contribution is to improve this by trading powers of ε in (25) for lower regularity of v. A similar idea has been used in [BSTU06] for the justification of the NLS for a generalization of (1) (in its rescaled form (12)) to periodic coefficients also for the linear terms, using a somewhat heavy machinery of Bloch wave transform and diagonalization of the linear part. Transferring back the result from [BSTU06] for (12), (13) to (1),(4) we obtain Theorem 2.3 under the condition $v_0 \in H^3(\mathbb{R})$, and thus a somewhat weaker result, with a significantly more complicated proof.

c) By requiring (31) we consider a special case of initial data (4), i.e., IC for wave–packets ψ_v that move to the right, and which thus can be described by a single NLS equation for v. In general, e.g., for $\partial_t u_\varepsilon(0, \cdot) \equiv 0$, the solution will decompose into two wave–packets, one moving left, and described by a NLS equation for v_-, and one moving right described by a NLS equation for v_+. As this is merely a question of book keeping, here we restrict to a single NLS.

d) Our choice of perturbations g_1, g_2 ensures that $(u_\varepsilon(0, .), \partial_t u_\varepsilon(0, .)) \in H^1(\mathbb{R}) \times L^2(\mathbb{R})$, where local existence and uniqueness holds by, e.g., semigroup methods, or Duhamel's formula. However, our theorem also provides the existence of a long-term solution. $\rfloor$

Proof. Using (22) we work with the improved ansatz (21). First we want to estimate $\varepsilon^{-1}\mathrm{Res}(\phi_v)$, but for $v(t) \notin H^4(\mathbb{R})$ we cannot use Lemma 2.2 directly. Instead we approximate $v_0 \in H^2(\mathbb{R})$ by smooth $v_{n,0}$ using a cutoff in Fourier space at wave-number n, i.e.,

$$v_{n,0}(x) = \mathcal{F}^{-1}(\chi_n \hat{v}_0)(x). \tag{33}$$

Here χ_n is the characteristic function of $[-n, n]$, and $\hat{v} = \mathcal{F}v$ and $v = \mathcal{F}^{-1}\hat{v}$ denote the Fourier transform and its inverse, respectively, which is an isomorphism of $H^m(\mathbb{R})$ and

$$L^2(m) := \left\{ \hat{u} : \mathbb{R} \to \mathbb{C} : \|\hat{u}\|_{L^2(m)} := \sqrt{\int (1 + k^{2m})|\hat{u}(k)|^2 \mathrm{d}k} < \infty \right\}.$$

Then $v_{n,0} \to v_0$ in $H^2(\mathbb{R})$ by Lebesgue dominated convergence, $v_{n,0} \in H^m(\mathbb{R})$ for all m, and for all $m \geq 2$ there exists a C_m such that $\|v_{n,0}\|_{H^m} \leq C_m\|\hat{v}_{n,0}\|_{L^2(m)} \leq CC_m n^{m-2}\|v_{n.0}\|_{H^2}$.

We set $\phi_n(t) = \phi_{v_n(t)}$ where $v_n(t)$ is the solution of the NLS (7) with IC $v_{n,0}$, and start with the triangle estimate

$$\|u - \phi_v\|_E \leq \|u - \phi_n\|_E + \|\phi_n - \phi_v\|_E.$$

From Remark 2.4a we have $\|v_n(t)-v(t)\|_{H^2} \le C\|v_n(0)-v(0)\|_{H^2}$ and hence

$$\|\phi_n-\phi_v\|_E^2 \le C\left(\varepsilon^4\int|\partial_t v_n-\partial_t v|^2\mathrm{d}x+\varepsilon^2\int|\partial_x v_n-\partial_x v|^2\mathrm{d}x+\int|v_n-v|^2\mathrm{d}x\right)\le C\varepsilon^2, \quad (34)$$

where the second estimate is obtained by choosing $n=\varepsilon^{-1/2}$, since

$$\int|v_n(0)-v(0)|^2\mathrm{d}x=\int_{|k|\ge n}\frac{k^4}{k^4}|\hat{v}_n(0)-\hat{v}(0)|^2\mathrm{d}x\le n^{-4}\|v_n(0)-v(0)\|_{H^2}^2\le C\varepsilon^2.$$

It remains to estimate $\|u-\phi_n\|_E$. Setting $r_n=\varepsilon^{-1}(u-\phi_n)$, the equation for r_n reads

$$\begin{aligned}\varepsilon^2\partial_t^2 r_n &= \partial_x^2 r_n-\varepsilon^{-2}r_n-f_n, \quad f_n=c_\varepsilon(3\phi_n^2 r_n+3\varepsilon\phi_n r^2+\varepsilon^2 r_n^3)-\varepsilon^{-1}\mathrm{Res}(\phi_n), \qquad (35)\\ r_n(0)&=\frac{1}{\varepsilon}(v(0)-v_n(0))e_1+g_1,\\ \partial_t r_n(0)&=-\frac{\omega_0}{\varepsilon^3}(v(0)-v_n(0))e_1+(\partial_t v(0)-\partial_t v_n(0))e_1+\frac{1}{\varepsilon^2}g_2,\end{aligned}$$

such that $\|r_n(0)\|_E=\mathcal{O}(1)$. Next,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|r_n\|_E^2=2\varepsilon^2\int(\partial_t r_n)f_n\mathrm{d}x\le 2\varepsilon^2\|\partial_t r_n\|_{L^2}\|f_n\|_{L^2}\le 2\|r_n\|_E\|f_n\|_{L^2}, \qquad (36)$$

and we need to estimate $\|f_n\|_{L^2}$. By the (1D) Gagliardo–Nirenberg inequality

$$\|u\|_{L^p}\le C\|\partial_x u\|_{L^2}^{\alpha}\|u\|_{L^2}^{1-\alpha} \text{ for all } u\in H^1(\mathbb{R}), \text{ where } \alpha=\frac{1}{2}-\frac{1}{p}, \qquad (37)$$

see, e.g., [Bre11, §8.6.1], we obtain, with $C_{\mathrm{Res},n}=\|\varepsilon^{-1}\mathrm{Res}_n\|_{L^2}$, and recalling that $\|u(t)\|_E^2=\varepsilon^4\|\partial_t u(t)\|^2+\varepsilon^2\|\partial_x u(t)\|^2+\|u(t)\|^2$,

$$\begin{aligned}\|f_n\|_{L^2} \le& C_v^2\|r_n\|_{L^2}+C_v\varepsilon\|r_n^2\|_{L^2}+\varepsilon^2\|r_n^3\|_{L^2}+C_{\mathrm{Res},n}\\ =&C_v^2\|r_n\|_{L^2}+C_v\varepsilon\|r_n\|_{L^4}^2+\varepsilon^2\|r_n\|_{L^6}^3+C_{\mathrm{Res},n}\\ \le& C_v^2\|r_n\|_{L^2}+C_vC\varepsilon\|r_n\|_{L^2}^{3/2}\|\partial_x r_n\|_{L^2}^{1/2}+C\varepsilon^2\|r_n\|_{L^2}^2\|\partial_x r_n\|_{L^2}^1+C_{\mathrm{Res},n}\\ \le& C_v^2\|r_n\|_E+C_vC\varepsilon^{1/2}\|r_n\|_E^2+C\varepsilon\|r_n\|_E^3+C_{\mathrm{Res},n}\le(C_v^2+C_R\varepsilon^{1/2})\|r_n\|_E+C_{\mathrm{Res},n}, \quad (38)\end{aligned}$$

as long as $\|r_n\|_E\le R$ with a constant C_R which depends on R, determined below, but not on ε.

As $\|v_n\|_{H^2}$ is bounded, from Lemma 2.2 we obtain $C_{\mathrm{Res},n}\le C(1+\varepsilon^2 C+\varepsilon\|v_n\|_{H^4})\le C_{\mathrm{Res}}$, where again we used $\|v_n\|_{H^4}\le n^2\|v_n\|_{H^2}\le\varepsilon^{-1}\|v_n\|_{H^2}$ for $n=\varepsilon^{-1/2}$. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|r_n\|_E^2\le 2\varepsilon^2\|\partial_t r_n\|_{L^2}\|f_n\|_{L^2}\le 2\|r_n\|_E\|f_n\|_{L^2}\le 2(C_v^2+1/2+C_R\varepsilon^{1/2})\|r_n\|_E^2+C_{\mathrm{Res}}^2. \quad (39)$$

Now we use that for any $C_R>0$ there exists an $\varepsilon_0>0$ such that

$$C_R\varepsilon^{1/2}\le 1/2 \text{ for all } 0<\varepsilon\le\varepsilon_0. \qquad (40)$$

Then, by Gronwall's inequality, for $0\le t\le t_0$,

$$\|r_n\|_E^2\le(\|r(0)\|_E^2+C_{\mathrm{Res}}^2)\mathrm{e}^{2(C_v^2+1)t_0}=:R^2, \qquad (41)$$

and we are done, i.e., for this R we find C_R in (38) and then $\varepsilon_0>0$ from (40). Combining (34) and (41) yields (32). □

We now state and prove a justification theorem when the envelope v is only in $H^s(\mathbb{R})$ with $1 < s < 2$. For solutions of (7) with $v \in C([0,t_0], H^s(\mathbb{R}))$ we cannot ensure $\|\psi_v(t)\|_E < \infty$ because one time derivative of v corresponds to two space derivatives and hence $\partial_t v \in L^2(\mathbb{R})$ is equivalent to $v \in H^2(\mathbb{R})$. Thus, we will bound the error in the scaled H^1 norm

$$\|\rho\|_{H^1_\varepsilon} = \sqrt{\int_{\mathbb{R}} \rho^2(x) + \varepsilon^2(\partial_x \rho(x))^2 \mathrm{d}x}. \tag{42}$$

The error $\varepsilon^{s/2}\rho$ is smallest for s close to 2 and the proof shows that a larger s also yields larger ε_0. The case $s = 1$ cannot be treated in the same way due to the lack of a bound of the nonlinearity as in (47).

Theorem 2.5 *Let $c \in L^2((0,1))$ and $s > 1$. For all $C_1 > 0$ and $t_0 > 0$, there exist $\varepsilon_0, C_2 > 0$ such that for all solutions $v \in C([0,t_0], H^s(\mathbb{R}))$ of (7) with $\sup_{t\in[0,t_0]} \|v(t,\cdot)\|_{H^s} \le C_1$, and all $\varepsilon \in (0,\varepsilon_0)$ the following holds. If*

$$u_\varepsilon(0,x) = \psi_v(0,x) + \varepsilon g_1(x, x/\varepsilon), \quad \partial_t u_\varepsilon(0,x) = -\frac{\omega_0}{\varepsilon^2}\psi_v(0,x) + \varepsilon^{-1} g_2(x, x/\varepsilon), \tag{43}$$

where $g_1, \varepsilon^{-1}\partial_x g_1, \partial_y g_1, g_2$ are bounded by C_1 in $L^2(\mathbb{R})$, then there exist a unique mild solution $u_\varepsilon \in C([0,t_0], H^1(\mathbb{R})) \cap C^1([0,t_0], L^2(\mathbb{R}))$ of (1) with initial conditions $u_\varepsilon(0,x)$ and $\partial_t u_\varepsilon(0,x)$, which can be written as $u_\varepsilon(t,x) = \psi_v(t,x) + \varepsilon^{s/2}\rho(t,x)$ with

$$\sup_{0\le t\le t_0} \|\rho\|_{H^1_\varepsilon} \le C_2. \tag{44}$$

Proof. We follow the same strategy as in the proof of Theorem 2.3, and approximate $v_0 \in H^s(\mathbb{R})$ by smooth $v_{n,0}$ using the same cutoff in Fourier space at wave-number n. The difference in the proofs lies in the norm $\|\cdot\|_{H^1_\varepsilon}$ instead of $\|\cdot\|_E$, yielding first (45), and subsequently $\varepsilon^{(s-1)/2}$ in (47), and $\delta = s-1$ after (49), which together yield the condition $s > 1$. We again set $\phi_n(t) = \phi_{v_n(t)}$ where $v_n(t)$ is the unique solution of (7) with initial data $v_{n,0}$, and start with the triangle inequality

$$\|u - \phi_v\|_{H^1_\varepsilon} \le \|u - \phi_n\|_{H^1_\varepsilon} + \|\phi_n - \phi_v\|_{H^1_\varepsilon}.$$

By Lipschitz continuity with respect to initial conditions for (7) (see Remark 2.4 a)) we have $\|v_n(t) - v(t)\|_{H^s} \le C\|v_n(0) - v(0)\|_{H^s}$, and hence

$$\|\phi_n - \phi_v\|^2_{H^1_\varepsilon} \le C(\varepsilon^2 \int |\partial_x v_n - \partial_x v|^2 \mathrm{d}x + \int |v_n - v|^2 \mathrm{d}x) \le C(\varepsilon^2 + \frac{1}{n^{2s}}), \tag{45}$$

since $\int |v_n(0) - v(0)|^2 \mathrm{d}x = \int_{|k|\ge n} \frac{k^4}{k^4} |\hat{v}_n(0) - \hat{v}(0)|^2 \mathrm{d}x \le n^{-2s}\|v_n(0) - v(0)\|^2_{H^s}$.

To bound $\|u - \phi_n\|_{H^1_\varepsilon}$ we again use the energy norm and the estimate (36). Letting $r_n := \varepsilon^{-s/2}(u - \phi_n)$, the equation for r_n becomes

$$\varepsilon^2 \partial_t^2 r_n = \partial_x^2 r_n - \varepsilon^{-2} r_n - f_n, \quad f_n = c_\varepsilon(3\phi_n^2 r_n + 3\varepsilon^{s/2}\phi_n r^2 + \varepsilon^s r_n^3) - \varepsilon^{-s/2}\mathrm{Res}(\phi_n), \tag{46}$$

hence $\frac{\mathrm{d}}{\mathrm{d}t}\|r_n\|^2_E = 2\varepsilon^2 \int (\partial_t r_n) f_n \mathrm{d}x \le 2\varepsilon^2 \|\partial_t r_n\|_{L^2} \|f\|_{L^2} \le 2\|r_n\|_E \|f_n\|_{L^2}$. With $C_{\mathrm{Res},n} = \|\varepsilon^{-s/2}\mathrm{Res}_n\|_{L^2}$ we obtain

$$\begin{aligned} \|f_n\|_{L^2} \le& C_v^2 \|r_n\|_{L^2} + C_v \varepsilon^{s/2} \|r_n^2\|_{L^2} + \varepsilon^s \|r_n^3\|_{L^2} + C_{\mathrm{Res},n} \\ \le& C_v^2 \|r_n\|_{L^2} + C_v C \varepsilon^{s/2} \|r_n\|^{3/2}_{L^2} \|\partial_x r_n\|^{1/2}_{L^2} + C\varepsilon^s \|r_n\|^2_{L^2} \|\partial_x r_n\|_{L^2} + C_{\mathrm{Res},n} \\ \le& C_v^2 \|r_n\|_E + C_v C \varepsilon^{(s-1)/2} \|r_n\|^2_E + C\varepsilon^{s-1} \|r_n\|^3_E + C_{\mathrm{Res},n} \\ \le& (C_v^2 + C_R \varepsilon^{(s-1)/2}) \|r_n\|_E + C_{\mathrm{Res},n}, \end{aligned} \tag{47}$$

as long as $\|r_n\|_E \le R$ with a constant C_R which depends on R, determined below, but not on ε.

As $\|v_n\|_{H^s}$ is bounded, from Lemma 2.2 we obtain

$$C_{\mathrm{Res},n} \le \varepsilon^{1-s/2} C(1+\varepsilon^2\|v_n\|_{H^2}^2 + \varepsilon\|v_n\|_{H^4}) \le C(\varepsilon^{1-s/2} + n^2\varepsilon^{3-s/2} + \varepsilon^{2-s/2}n^{4-s}) \le C_{\mathrm{Res}} \quad (48)$$

if $s \le 2$ and $n = n(\varepsilon) = \varepsilon^{-1/2}$. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|r_n\|_E^2 \le \varepsilon^2\|\partial_t r_n\|_{L^2}\|f_n\|_{L^2} \le 2\|r_n\|_E\|f_n\|_{L^2} \le (C_v^2 + 1/2 + C_R\varepsilon^\delta)\|r_n\|_E^2 + C_{\mathrm{Res}}^2 \quad (49)$$

with $\delta = s-1$, and

$$\begin{aligned} r_n(0) &= \varepsilon^{-s/2}(v(0)-v_n(0))e_1 + \varepsilon^{1-s/2}g_1, \\ \frac{\mathrm{d}}{\mathrm{d}t}r_n(0) &= -\varepsilon^{-(2+s/2)}\omega_0(v(0)-v_n(0))e_1 - \varepsilon^{-s/2}\dot{v}_n e_1 + \varepsilon^{-(1+s/2)}g_2. \end{aligned}$$

With $\|r_n(0)\|_E^2 = \|r_n(0)\|_{L^2}^2 + \varepsilon^2\|\partial_x r_n(0)\|_{L^2}^2 + \varepsilon^4\|\frac{\mathrm{d}}{\mathrm{d}t}r_n(0)\|_{L^2}^2$ we obtain that

$$\begin{aligned} \|r_n(0)\|_E \le C\Big(&\varepsilon^{-s/2}\|v(0)-v_n(0)\|_{L^2} + \varepsilon^{1-s/2}\|g_1\|_{L^2} + \varepsilon^{1-s/2}\|\partial_x v(0) - \partial_x v_n(0)\|_{L^2} + \varepsilon^{2-s/2}\|\partial_x g_1\|_{L^2} \\ &+ \varepsilon^{1-s/2}\|\partial_y g_1\|_{L^2} + \varepsilon^{-s/2}\|v(0)-v_n(0)\|_{L^2} + \varepsilon^{2-s/2}\|\dot{v}_n\|_{L^2} + \varepsilon^{1/2-\delta}\|g_2\|_{L^2}\Big) \end{aligned}$$

is also bounded for $n = n(\varepsilon) = \varepsilon^{-1/2}$. As above we use that for any fixed $C_R > 0$ there exists an $\varepsilon_0 > 0$ such that

$$C_R\varepsilon^\delta \le 1/2 \text{ for all } 0 < \varepsilon \le \varepsilon_0. \quad (50)$$

Then, by Gronwall's inequality, for $0 \le t \le t_0$,

$$\|r_n\|_E^2 \le (\|r(0)\|_E^2 + C_{\mathrm{Res}}^2)\mathrm{e}^{(C_v^2+1)t_0} =: R^2, \quad (51)$$

and the remainder of the proof works exactly as in the proof of Theorem 2.3. □

Acknowledgements: The authors are grateful for the support by *The Royal Society* (International Exchange reference IE131675).

References

[All08] G. Allaire. Periodic homogenization and effective mass theorems for the Schrödinger equation. In *Quantum transport*, volume 1946 of *Lecture Notes in Math.*, pages 1–44. Springer, Berlin, 2008.

[AP05] G. Allaire and A. Piatnitski. Homogenization of the Schrödinger equation and effective mass theorems. *Commun. Math. Phys.*, 258(1):1–22, 2005.

[AP06] G. Allaire and M. Palombaro. Localization for the Schrödinger equation in a locally periodic medium. *SIAM J. Math. Anal.*, 38(1):127–142, 2006.

[APR09] G. Allaire, M. Palombaro, and J. Rauch. Diffractive behavior of the wave equation in periodic media: weak convergence analysis. *Annali di Matematica*, 188:561–589, 2009.

[APR11] G. Allaire, M. Palombaro, and J. Rauch. Diffractive geometric optics for Bloch wave packets. *Arch. Ration. Mech. Anal.*, 202(2):373–426, 2011.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[BSTU06] K. Busch, G. Schneider, L. Tkeshelashvili, and H. Uecker. Justification of the Nonlinear Schrödinger equation in spatially periodic media. *ZAMP*, 57:1–35, 2006.

[Car07] Rémi Carles. Wkb analysis for nonlinear schrdinger equations with potential. *Comm. Math. Phys.*, 269:195–221, 2007.

[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.

[GS01] M. D. Groves and G. Schneider. Modulating pulse solutions for a class of nonlinear wave equations. *Comm. Math. Phys.*, 219(3):489–522, 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.

[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.

[Lan98] David Lannes. Dispersion effects for nonlinear geometrical optics with rectification. *Asymptotic Analysis*, 18:111–146, 1998.

[MN02] N. Masmoudi and K. Nakanishi. From nonlinear Klein-Gordon equation to a system of coupled nonlinear Schrödinger equations. *Math. Ann.*, 324(2):359–389, 2002.

[MNO02] S. Machihara, K. Nakanishi, and T. Ozawa. Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations. *Math. Ann.*, 322(3):603–621, 2002.

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

[Sch05] G. Schneider. Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. *J. Diff. Eq.*, 216(2):354–386, 2005.

[Sha85] J. Shatah. Normal forms and quadratic nonlinear Klein–Gordon equations. *Comm. Pure Appl. Math.*, 38:685–696, 1985.

[Spa06] Chr. Sparber. Effective mass theorems for nonlinear Schrödinger equations. *SIAM J. Appl. Math.*, 66(3):820–842, 2006.

[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.

[SU07] G. Schneider and H. Uecker. The mathematics of light pulses in dispersive media. *Jahresberichte der DMV*, 109:139–161, 2007.

[SW04] T. Schäfer and C. E. Wayne. Propagation of ultra-short optical pulses in cubic nonlinear media. *Phys. D*, 196(1-2):90–105, 2004.

[Tao06] T. Tao. *Nonlinear dispersive equations*, volume 106 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 2006.