Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

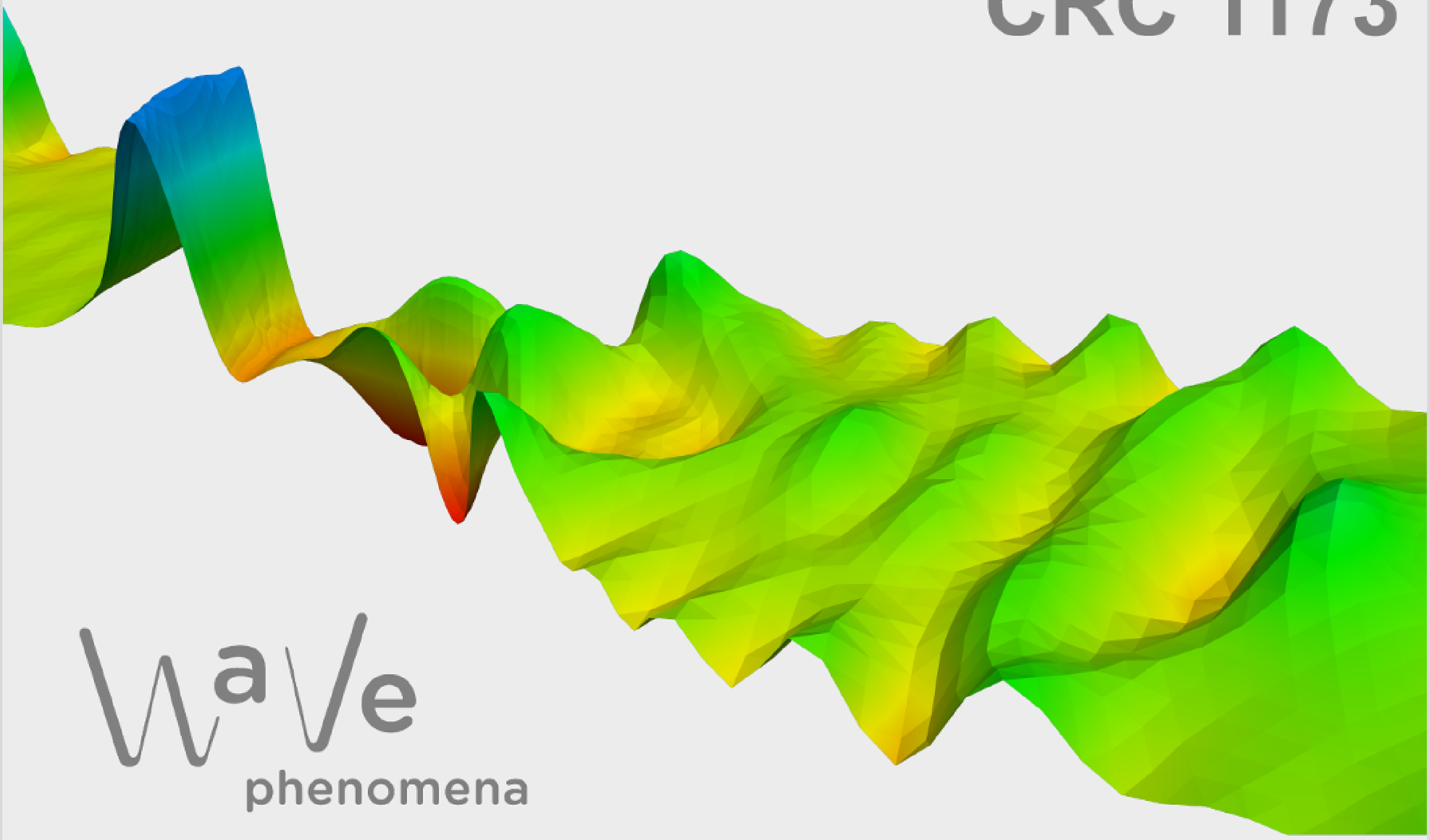
A robust way to justify the Derivative NLS approximation

Max Heß, Guido Schneider

CRC Preprint 2022/50, September 2022

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave phenomena

Participating universities

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

TU WIEN

Funded by

DFG

ISSN 2365-662X

A robust way to justify the Derivative NLS approximation

Max Heß and Guido Schneider
Institut für Analysis, Dynamik und Modellierung,
Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

September 30, 2022

Abstract

The Derivative Nonlinear Schrödinger (DNLS) equation can be derived as an amplitude equation via multiple scaling perturbation analysis for the description of the slowly varying envelope of an underlying oscillating and traveling wave packet in dispersive wave systems. It appears in the degenerated situation when the cubic coefficient of the similarly derived NLS equation vanishes. It is the purpose of this paper to prove that the DNLS approximation makes correct predictions about the dynamics of the original system under rather weak assumptions on the original dispersive wave system if we assume that the initial conditions of the DNLS equation are analytic in a strip of the complex plane. The method is presented for a Klein-Gordon model with a cubic nonlinearity.

1 Introduction

The Nonlinear Schrödinger (NLS) equation

$$i\partial_T A = \nu_1 \partial_X^2 A + \widetilde{\nu}_2 A|A|^2, \tag{1}$$

with coefficients $\nu_1, \widetilde{\nu}_2 \in \mathbb{R}$, can be derived for the description of small modulations in time and space of oscillatory wave packets in dispersive wave systems, such as the quadratic ($f(u) = u^2$) or cubic ($f(u) = u^3$) Klein-Gordon equation

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

the water wave problem, systems from nonlinear optics, etc. For the cubic Klein-Gordon equation the ansatz for the derivation of the NLS equation is given by

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

where c_g is the linear group velocity, k_0 the basic spatial wave number, ω_0 the basic temporal wave number, and $0 < \varepsilon \ll 1$ a small perturbation parameter.

Various NLS approximation results have been proven in the last decades. Such a result can trivially be established for a dispersive wave system with no quadratic terms by using Gronwall's inequality, cf. [KSM92]. However, in case of quadratic nonlinearities such a result is non-trivial since terms of order $\mathcal{O}(\varepsilon)$ have to be controlled on the long $\mathcal{O}(1/\varepsilon^2)$-time scale. The idea to get rid of this problem is to use so-called normal form transformations. By a near identity change of variables the terms of order $\mathcal{O}(\varepsilon)$ can be eliminated if non-resonance conditions are satisfied, cf. [Kal88]. The last years saw various attempts to weaken these non-resonance conditions in order to control appearing resonances, cf. [Sch05], and to make the theory applicable to quasilinear systems, cf. [WC17, Due17, DH18], such as the water wave problem, cf. [TW12, DSW16, IT19, Due21].

It turned out that in case of initial conditions for the NLS equation which are analytic in a strip of the complex plane almost no non-resonance conditions are necessary, cf. [Sch98, DHSZ16]. It is the purpose of this paper to explain that the method developed in [Sch98, DHSZ16] can be used in the justification of the Derivative Nonlinear Schrödinger (DNLS) approximation, too. Interestingly, it allows to get rid of a problem which is not present in the justification of the NLS approximation, see below.

The DNLS equation

$$i\partial_T A = \nu_1 \partial_X^2 A + \nu_2 A|A|^2 + i\nu_3 |A|^2 \partial_X A + i\nu_4 A^2 \partial_X \overline{A} + \nu_5 A|A|^4, \qquad (2)$$

with $T \geq 0$, $X \in \mathbb{R}$, $A(X,T) \in \mathbb{C}$, and coefficients $\nu_j \in \mathbb{R}$ for $j = 1, \ldots, 5$, appears when the cubic coefficient $\widetilde{\nu}_2 = \widetilde{\nu}_2(k_0)$ in (1) vanishes for the chosen basic wave number k_0. This situation appears for instance in the water wave problem for certain values of surface tension and basic spatial wave number k_0, cf. [AS81]. The DNLS equation can be derived with an ansatz

$$u(x,t) = \varepsilon^{1/2} A(\varepsilon(x - c_g t), \varepsilon^2 t) e^{i(k_0 x - \omega_0 t)} + c.c.. \qquad (3)$$

The justification is more difficult and from a mathematical point of view even more interesting than that for the NLS approximation since for original dispersive wave systems with a quadratic nonlinearity, in the equation for the error, terms of order $\mathcal{O}(\varepsilon^{1/2})$ have to be controlled on a long $\mathcal{O}(1/\varepsilon^2)$-time scale. Even for dispersive wave systems with a cubic nonlinearity, in the equation for the error, terms of order $\mathcal{O}(\varepsilon)$ have to be controlled on a long $\mathcal{O}(1/\varepsilon^2)$-time scale and so as a first step in establishing an approximation theory for the DNLS approximation we start with the most simple toy problem, namely a nonlinear Klein-Gordon equation with a special cubic nonlinearity,

$$\partial_t^2 u = \partial_x^2 u - u + \varrho(\partial_x) u^3. \tag{4}$$

Herein, $x \in \mathbb{R}$, $t \in \mathbb{R}$, $u(x,t) \in \mathbb{R}$, and

$$\varrho(ik) = \frac{k^2-1}{k^2+1}, \qquad \text{resp.} \qquad \varrho(\partial_x) = -(1-\partial_x^2)^{-1}(1+\partial_x^2).$$

Plugging the ansatz (3) with $k_0 = 1$ into (4) and equating the coefficients in front of $e^{i(k_0 x - \omega_0 t)}$ to zero gives at $\mathcal{O}(\varepsilon^{1/2})$ the linear dispersion relation $\omega_0^2 = k_0^2 + 1$ and at $\mathcal{O}(\varepsilon^{3/2})$ the linear group velocity $c_g = k_0/\omega_0$. Using the expansion

$$\varrho(i + \varepsilon\partial_X) = \frac{-(i+\varepsilon\partial_X)^2 - 1}{-(i+\varepsilon\partial_X)^2 + 1} = -i\varepsilon\partial_X + \mathcal{O}(\varepsilon^2)$$

gives at $\mathcal{O}(\varepsilon^{5/2})$ the DNLS equation

$$2i\omega_0 \partial_T A = (1-c_g^2)\partial_X^2 A - 3i\partial_X(A|A|^2). \tag{5}$$

It is the goal of this paper to prove that the DNLS equation (5) makes correct predictions about the dynamics of the Klein-Gordon model (4). It turns out that there are two new difficulties which have to be overcome and which were not present in the justification analysis of other modulation equations so far, namely the problem of a total resonance and the problem of a second order resonance, see below and Section 3. As already said, we do so, by adapting a method developed in [Sch98, DHSZ16] for justifying the NLS approximation under rather weak non-resonance conditions, however, with the drawback that the initial conditions for the NLS equation have to be chosen analytic in a strip of the complex plane. This approach will be combined with some energy estimates in order to get rid of some resonance structure which is not present for the NLS approximation. We call the following approach to justify the DNLS approximation robust since our approximation result holds, as already said, under rather weak non-resonance conditions.

Our analysis is based on the use of Gevrey spaces

$$G_\sigma^s = \{u : \mathbb{R} \to \mathbb{C} : \|u\|_{G_\sigma^s} := \|e^{\sigma(|k|+1)}(1+|k|^{2s})^{1/2}\widehat{u}(k)\|_{L^2(dk)} < \infty\}$$

which are defined for $\sigma, s \geq 0$. We recall that due to the Paley-Wiener theorem functions $u \in G_\sigma^s$ can be extended to a strip $\{z \in \mathbb{C} : |\mathrm{Im} z| < \sigma\}$ in the complex plane, cf. [Kat04].

Remark 1.1. The Fourier transform of the DNLS approximation (3) is given by

$$\widehat{u}(k,t) = \varepsilon^{1/2}\varepsilon^{-1}\widehat{A}(\frac{k-k_0}{\varepsilon}, \varepsilon^2 t)e^{-i\omega_0 t - ic(k-k_0)t} + \widehat{c.c.} \tag{6}$$

Hence, the Fourier transform is strongly concentrated at the wave numbers $\pm k_0 = \pm 1$ and so the evolution of $\widehat{A}$ is determined by the form of the dispersion relation and of the nonlinearity at the wave number $k_0 = 1$.

Then our approximation theorem is as follows.

Theorem 1.2. *Let $s_A \geq 12$, $\sigma_0 > 0$, and $A \in C([0,T_0], G_{\sigma_0}^{s_A})$ be a solution of the DNLS equation* (5). *Then there exist $\varepsilon_0 > 0$, $T_1 \in (0, T_0]$, and $C > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have solutions u of the Klein-Gordon model* (4) *such that*

$$\sup_{t\in[0,T_1/\varepsilon^2]} \sup_{x\in\mathbb{R}} |u(x,t) - (\varepsilon^{1/2}A(\varepsilon(x-ct), \varepsilon^2 t)e^{i(k_0 x - \omega_0 t)} + c.c.)| \leq C\varepsilon.$$

Remark 1.3. As already said, such an approximation result is non-trivial since solutions of order $\mathcal{O}(\varepsilon^{1/2})$ of (4) have to be controlled on an $\mathcal{O}(1/\varepsilon^2)$-time scale. Although we have a cubic nonlinearity a simple application of Gronwall's inequality would only give estimates on an $\mathcal{O}(1/\varepsilon)$-time scale.

Remark 1.4. Such an approximation result should not be taken for granted. There are counterexamples, cf. [Sch95, SSZ15, HS20], showing that there are amplitude equations which are derived in a formally correct way, but fail to make correct predications about the original system on the natural time scale of the approximation.

Remark 1.5. The approximation result is not optimal in the sense that error estimates can only be proven on the correct time scale, namely for $t \in [0, T_1/\varepsilon^2]$, but not necessarily for all $t \in [0, T_0/\varepsilon^2]$. Hence we can only guarantee that parts of the DNLS dynamics can be seen in the original system.

Remark 1.6. For completeness we remark that the DNLS equation is a well studied nonlinear dispersive system. Local well-posedness of smooth solutions in Sobolev spaces H^s with $s > 3/2$ was established by Tsutsumi and Fukuda [TF80]. See also [HO92, Tak99, CKS$^+$02, Wu15, WG17] for further improvements. The complete integrability of the DNLS equation has been established in [KN78]. For a recent overview see [JLPS20].

The plan of the paper is as follows. In Section 2 we write the Klein-Gordon model (4) as a first order system in Fourier space and derive the equations for the error made by an improved DNLS approximation. The improved DNLS approximation is $\mathcal{O}(\varepsilon^{3/2})$-close to the original DNLS approximation (3). In Appendix A we construct the improved DNLS approximation and estimate the remaining residual terms.

In Section 3 we perform some normal form transformations in order to get rid of the terms of order $\mathcal{O}(\varepsilon)$ in the equations for the error. It turns out that there are resonances present in the system and that not all terms of order $\mathcal{O}(\varepsilon)$ can be eliminated. There is a total resonance, i.e., there are cubic terms which cannot be eliminated for any wave number $k \in \mathbb{R}$. Moreover, there is a resonance at the wave numbers $\pm k_0 = \pm 1$. The denominator in the normal from transformation vanishes of second order and so this resonance will be called second order resonance in the following. Since the nonlinear terms which appear in the nominator only vanish linearly at the wave numbers $\pm k_0 = \pm 1$, the associated part of the normal form transform would be unbounded. Therefore, in Section 3 we eliminate all terms which are not associated to the total resonance or second order resonance. It turns out that the total resonant terms can be controlled with a simple energy estimate and so we concentrate on the handling of the terms associated to the second order resonance.

As a preparation in Section 4 we recall the estimates from the local existence and uniqueness proof of the DNLS equation in Gevrey spaces. The exponential localization of the solutions in Fourier space will allow us to use the derivative in front of the nonlinear term in (5) and to come subsequently to the correct time scale. By lowering the decay rates σ in the definition of the norm $\|\cdot\|_{G^s_\sigma}$ with time, we obtain an artificial smoothing which allows us to get rid of first order derivatives in the nonlinearity.

However, in order to use this idea in the original system (4) we have to get rid of the fact that the DNLS modes are concentrated at the wave number $k_0 = 1$ and that the nonlinear term vanishes at this wave number, too. Hence, in Section 5 we introduce a space where the Fourier modes are located at integer multiples of k_0 with an exponential decay proportional to $|k - mk_0|$. Again by lowering the decay rates with time we rebuilt the construction from Section 4 for (4), cf. Figure 1. This allows us to to come with our error estimates to the natural $\mathcal{O}(\varepsilon^{-2})$-time scale of the DNLS approximation.

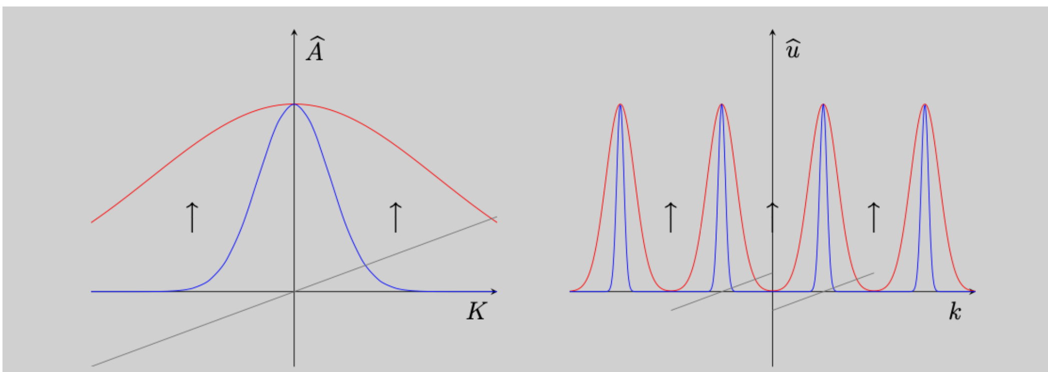


Figure 1: The left panel shows the mode distribution for the DNLS equation (5) and the vanishing of the nonlinearity at $K = 0$. The arrows indicate that the required decay rates are lowered in time. The right panel shows the mode distribution for the original system (4) and the vanishing of the nonlinearity at $k = \pm 1$. The arrows indicate that the required decay rates in between the integers are lowered in time.

We remark that the spaces which will be introduced in Section 5 have been used in [DHSZ16] for a completely different purpose, namely to control resonances which are bounded away from integer multiples of the basic wave number k_0. Hence, as in [DHSZ16] this allows us to weaken the non-resonance conditions and to allow for additional resonances away from the integer multiples of the basic wave number k_0. Since, depending on the wave numbers, then different parts of the error function are handled differently, in Section 6 we introduce some mode filters which allow us to separate these parts in Fourier space. Estimates for the normal form transformation in the chosen spaces can be found in Section 7.

The final error estimates can be found in Section 8. We use energy estimates for the transformed system since we still have to get rid of the totally resonant terms. All ideas from the previous sections can be incorporated in these energy estimates. We close the paper in Section 9 with a discussion about possible improvements, generalizations, and about the possible transfer to more complicated systems.

Notation. The Fourier transform of a function $u : \mathbb{R} \mapsto \mathbb{C}$ is given by

$$(\mathcal{F}u)[k] = \widehat{u}(k) = \frac{1}{2\pi} \int_{\mathbb{R}} u(x) e^{-ikx} dx.$$

The inverse Fourier transform of a function $\widehat{u} : \mathbb{R} \mapsto \mathbb{C}$ is given by

$$(\mathcal{F}^{-1}\widehat{u})[x] = u(x) = \int_{\mathbb{R}} \widehat{u}(k) e^{ikx} dk.$$

Multiplication $(uv)(x) = u(x)v(x)$ in physical space corresponds in Fourier space to the convolution

$$(\widehat{u} * \widehat{v})(k) = \int_{\mathbb{R}} \widehat{u}(k-l)\widehat{v}(l) dl.$$

In the following many possibly different constants are denoted with the same symbol C if they can be chosen independent of the small perturbation parameter $0 < \varepsilon \ll 1$.

Acknowledgement. The work of Guido Schneider is partially supported by the Deutsche Forschungsgemeinschaft DFG through the SFB 1173 "Wave phenomena" Project-ID 258734477.

2 Equations for the error

The proof of Theorem 1.2 is based on the following ideas. The Fourier transformed cubic Klein-Gordon model (4) is given by

$$\partial_t^2 \widehat{u}(k,t) = -\omega^2(k)\widehat{u}(k,t) - \omega(k)\rho(ik)\widehat{u}^{*3}(k,t) \tag{7}$$

where $\omega(k) = \text{sign}(k)\sqrt{1+k^2}$ and $\rho(k) = -\frac{\varrho(ik)}{\omega(k)}$. By this choice of ω and ρ the subsequent variables will be real-valued in physical space. We write (7) as a first order system

$$\begin{aligned} \partial_t \widehat{u}_1(k,t) &= -i\omega(k)\widehat{u}_2(k,t), \\ \partial_t \widehat{u}_2(k,t) &= -i\omega(k)\widehat{u}_1(k,t) - i\rho(k)\widehat{u}_1^{*3}(k,t). \end{aligned}$$

This system is diagonalized with

$$2\widehat{v}_{-1} = \widehat{u}_1 + \widehat{u}_2, \qquad 2\widehat{v}_1 = \widehat{u}_1 - \widehat{u}_2.$$

We obtain

$$\partial_t V = \Lambda V + N(V, V, V),$$

where in Fourier space

$$\widehat{\Lambda}(k) = \begin{pmatrix} -i\omega(k) & 0 \\ 0 & i\omega(k) \end{pmatrix}$$

is a skew symmetric operator and

$$\widehat{N}(\widehat{V}, \widehat{V}, \widehat{V})(k,t) = \frac{1}{2} i\rho(k) \begin{pmatrix} -(\widehat{v}_1 + \widehat{v}_{-1})^{*3} \\ (\widehat{v}_1 + \widehat{v}_{-1})^{*3} \end{pmatrix} (k,t)$$

a symmetric trilinear mapping. The DNLS approximation is of the form

$$\varepsilon^{1/2}\psi = \begin{pmatrix} \varepsilon^{1/2} a_1 + \varepsilon^{1/2} a_{-1} + \varepsilon^{3/2}\psi_{s,1} \\ \varepsilon^{3/2}\psi_{s,-1} \end{pmatrix}, \tag{8}$$

with a_j concentrated at the wave number $k = j$ and higher order approximation terms $\varepsilon^{3/2}\psi_{s,\pm 1}$. See Appendix A for the detailed construction. The error $\varepsilon^{\widetilde{\beta}} R = V - \varepsilon^{1/2}\psi$ with $\widetilde{\beta} > 3/2$ made by the DNLS approximation satisfies

$$\partial_t R = \Lambda R + \varepsilon L_c(R) + \varepsilon^2 L_s(R) + \varepsilon^{\widetilde{\beta}+1/2} L_r(R) + \varepsilon^{-\widetilde{\beta}} \mathrm{Res}(\varepsilon^{1/2}\psi) \tag{9}$$

where

$$\begin{aligned}
\widehat{L_c(R)}(k,t) &= \frac{3}{2} i\rho(k) \begin{pmatrix} -(\widehat{a}_1 + \widehat{a}_{-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \\ (\widehat{a}_1 + \widehat{a}_{-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \end{pmatrix} (k,t), \\
\widehat{L_s(R)}(k,t) &= 3i\rho(k) \begin{pmatrix} -(\widehat{a}_1 + \widehat{a}_{-1}) * (\widehat{\Psi}_{q,1} + \widehat{\Psi}_{q,-1}) * (\widehat{R}_1 + \widehat{R}_{-1}) \\ (\widehat{a}_1 + \widehat{a}_{-1}) * (\widehat{\Psi}_{q,1} + \widehat{\Psi}_{q,-1}) * (\widehat{R}_1 + \widehat{R}_{-1}) \end{pmatrix} (k,t) \\
&\quad + \varepsilon \frac{3}{2} i\rho(k) \begin{pmatrix} -(\widehat{\Psi}_{q,1} + \widehat{\Psi}_{q,-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \\ (\widehat{\Psi}_{q,1} + \widehat{\Psi}_{q,-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \end{pmatrix} (k,t), \\
\widehat{L_r(R)}(k,t) &= \frac{3}{2} i\rho(k) \begin{pmatrix} -(\widehat{\Psi}_1 + \widehat{\Psi}_{-1}) * (\widehat{R}_1 + \widehat{R}_{-1})^{*2} \\ (\widehat{\Psi}_1 + \widehat{\Psi}_{-1}) * (\widehat{R}_1 + \widehat{R}_{-1})^{*2} \end{pmatrix} (k,t) \\
&\quad + \varepsilon^{\widetilde{\beta}-1/2} \frac{1}{2} i\rho(k) \begin{pmatrix} -(\widehat{R}_1 + \widehat{R}_{-1})^{*3} \\ (\widehat{R}_1 + \widehat{R}_{-1})^{*3} \end{pmatrix} (k,t).
\end{aligned}$$

and where $\varepsilon^{-\widetilde{\beta}}\mathrm{Res}(\varepsilon^{1/2}\psi)$ are the so called residual terms. These are the terms which do not cancel after inserting the DNLS approximation into the nonlinear Klein-Gordon equation (4). In the following we concentrate on estimating the error made by the DNLS approximation and postpone the standard construction of an improved approximation and the estimates for the residual to Appendix A. The improved approximation will be chosen in such a way that the term $\varepsilon^{-\widetilde{\beta}}\mathrm{Res}(\varepsilon^{1/2}\psi)$ is of order $\mathcal{O}(\varepsilon^2)$.

In the following in our notation we keep $\widetilde{\beta} > 3/2$ in order to show that by using improved approximations the error can be made arbitrarily small. All terms on the the right-hand side of (9) are at least of order $\mathcal{O}(\varepsilon^2)$ except of the first two terms. Since Λ is skew symmetric the first term on the right-hand side of (9) makes no problems, too. However, the second term $\varepsilon L_c(R)$ of order $\mathcal{O}(\varepsilon)$ makes serious problems in estimating the error on the long $\mathcal{O}(1/\varepsilon^2)$-time scale.

3 Normal form transformations and the resonance structure

The approach to get rid of the dangerous term $\varepsilon L_c(R)$ in (9) are normal form transformations. By these near identity change of variables

$$R = w + \varepsilon M(\psi, \psi, R),$$

with M a trilinear mapping, the $\mathcal{O}(\varepsilon)$-terms can be transformed into $\mathcal{O}(\varepsilon^2)$-terms, if a number of non-resonance conditions are satisfied. Since the $\widehat{a}_j$ are strongly concentrated at the wave numbers $k = j$ the non-resonance condition

$$r_{jj_1j_2j_3}(k) = -j\omega(k) - \omega(j_1) - \omega(j_2) + j_3\omega(k - j_1 - j_2) \neq 0$$

has to be satisfied for all $k \in \mathbb{R}$ for the elimination of a term $\widehat{a}_{j_1} * \widehat{a}_{j_2} * R_{j_3}$ from the equation for R_j, cf. [SU17]. The non-resonance conditions can be analyzed graphically. We find no resonances except of

(TR) For $(j, j_1, j_2, j_3) = (j, j_1, -j_1, j)$ the resonance function $r_{jj_1j_2j_3}(k)$ vanishes identically. Thus, the associated terms in $\varepsilon L_c(R)$ cannot be eliminated by a normal form transformation.

(SOR) For $(j, j_1, j_2, j_3) = (-1, j_1, j_1, -1)$, cf. Figure 2, there is a resonance at $k = j_1$, which is of second order, in detail

$$\omega(k) - 2\omega(j_1) - \omega(k - 2j_1) = 2\omega''(j_1)(k - j_1)^2 + \mathcal{O}(|k - j_1|^3)$$

for k near j_1. This second order resonance would appear in the denominator of the normal form transformation. It cannot be balanced by the term ρ in the nominator of the normal form transformation which only linearly vanishes at $k = \pm 1$. Thus, the normal form transformation would be singular near the wave numbers $k = \pm 1$.

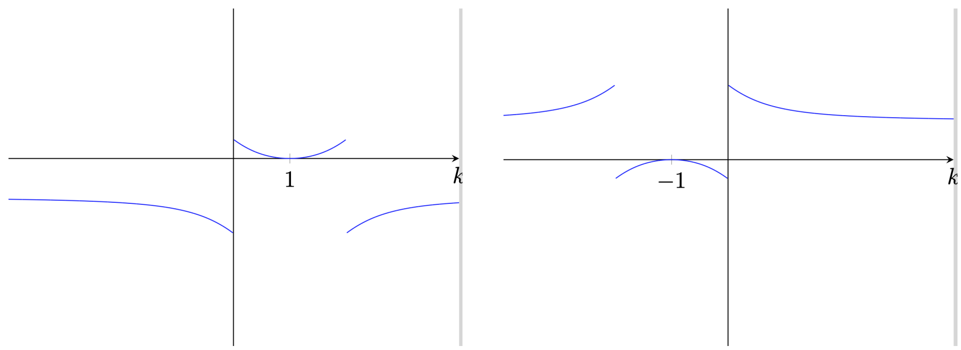


Figure 2: Plots of $r_{jj_1j_2j_3}(k)$ for the resonances with the second order touching. The left panel shows the case $(j, j_1, j_2, j_3) = (-1, 1, 1, -1)$ and the right panel shows the case $(j, j_1, j_2, j_3) = (-1, -1, -1, 1)$.

Thus, beside the normal form transformations which we use to get rid of the non-resonant terms, we need an idea to get rid of the terms which cannot be eliminated at all due to the total resonance **(TR)**, and we need an idea to get rid of the terms which cannot be eliminated in a small neighborhood of the wave numbers $k = \pm 1$ due to the second order resonance **(SOR)**.

It turned that the problem with the total resonance **(TR)** can be solved rather easily by using energy estimates. For handling the second order resonance **(SOR)** we use the fact that in lowest order the system near the wave numbers $k = \pm 1$ is given by the DNLS equation. Our approach to solve this problem is similar to the approach chosen for instance in [KN86, Sch96] for the justification of the KdV approximation. By this approach we not only get rid of the quasilinearity of the DNLS equation but also gain the missing $\mathcal{O}(\varepsilon)$-order to come to the long $\mathcal{O}(1/\varepsilon^2)$-time scale. In order to explain this approach we have a look at the the solution theory of the DNLS equation in Gevrey spaces in Section 4 first.

Solving the NLS equation in Gevrey spaces was also the basis of the approach which has been used in [Sch98, DHSZ16] for justifying the NLS approximation under rather weak non-resonance conditions. This underlying idea of the approach is introduced in Section 5. Interestingly, it allows us to get rid of the second order resonances which are not present in the justification of the NLS approximation,

And so in the end the transfer of the method developed in [Sch98, DHSZ16] from the NLS approximation to the DNLS approximation not only gains the missing $\mathcal{O}(\varepsilon)$-order in order to come to the long $\mathcal{O}(1/\varepsilon^2)$-time scale but also allows us to justify the DNLS approximation under rather weak non-resonance conditions. Since we need to control the total resonant terms, too, we use energy estimates instead of the variation of constant formula, and so the L^1-based spaces from [DHSZ16] are replaced here by L^2-based spaces in Fourier space.

4 The DNLS equation in Gevrey spaces

As already said we solve the DNLS equation in Gevrey spaces G^s_σ equipped with the norm

$$\|u\|_{G^s_\sigma} := \|e^{\sigma(|\xi|+1)}(1+|\xi|^{2s})^{1/2}\widehat{u}(\xi)\|_{L^2(d\xi)} . \tag{10}$$

In our presentation of the properties of these spaces we follow [BKS20]. For the local existence and uniqueness of solutions we use that G^s_σ is an algebra for $s > 1/2$, i.e., if $u, v \in G^s_\sigma$ then $uv \in G^s_\sigma$ and

$$\|uv\|_{G^s_\sigma} \leq C_s \|u\|_{G^s_\sigma} \|v\|_{G^s_\sigma}, \tag{11}$$

where the constant C_s is independent of $\sigma > 0$. Since the DNLS equation is a quasilinear system we need the following improved, so called tame, estimate

$$\|uv\|_{G^s_\sigma} \leq C_s(\|u\|_{G^s_\sigma}\|v\|_{G^\kappa_\sigma} + \|u\|_{G^\kappa_\sigma}\|v\|_{G^s_\sigma}) \tag{12}$$

which holds for all $\kappa \geq 0$. The elements of G^s_σ form a proper subset of the space of functions which are analytic in a strip of the complex plane of width $< 2\sigma$, symmetric around the real axis, equipped with the sup-norm due to the Paley-Wiener theorem, cf. [Kat04].

For the DNLS equation we have the following local existence and uniqueness result.

Theorem 4.1. *Let $s > 1$ and $\sigma_A > 0$. Then, for every $R > 0$, there exist $\eta = \eta(R, s, \sigma_A)$ such that for every $A_0 \in G^s_{\sigma_A}$, with $\|A_0\|_{G^s_{\sigma_A}} \leq R$, there exists a unique local solution $A(T) \in G^s_{\sigma(T)}$ of the DNLS equation* (5) *with $\sigma(T) := \sigma_A - \eta T$, $T \in [0, \sigma_A/\eta]$, and $\sup_{T\in[0,\sigma_A/\eta]} \|A(T)\|_{G^s_{\sigma(T)}} \leq R$.*

Proof. By rescaling T, X, and A the DNLS equation (5) is brought in its normal form

$$\partial_T A = i\partial_X^2 A - \partial_X(A|A|^2).$$

Next we set

$$A(\cdot, T) = S(T)B(\cdot, T) = e^{2\sigma(T)(1+M)}B(\cdot, T),$$

where $M = \sqrt{-\partial_x^2}$. Then B satisfies

$$\partial_T B = -\eta(1 + M)B + i\partial_X^2 B - \partial_X(S^{-1}(T)((S(T)B)|S(T)B|^2)).$$

We denote the scalar product in H^s with $(\cdot, \cdot)_s$ and obtain

$$\partial_T(B, B)_s = -\eta((1 + M)^{1/2}B, (1 + M)^{1/2}B)_s + g(B),$$

where

$$|g(B)| \leq C\|B\|^2_{H^s}\|B\|^2_{H^{s+1/2}} \leq C\|B\|^2_{H^s}((1 + M)^{1/2}B, (1 + M)^{1/2}B)_s$$

such that

$$\partial_T(B, B)_s \leq (-\eta + C\|B\|^2_{H^s})((1 + M)^{1/2}B, (1 + M)^{1/2}B)_s.$$

Hence $(B, B)_s$ decays in time if $\eta > 0$ is chosen so large that initially

$$(-\eta + C\|B|_{T=0}\|^2_{H^s}) < 0.$$

With these a priori estimates the existence and uniqueness of solutions follows by standard arguments, cf. [Kat75]. □

The existence of the solutions of the DNLS equation (5) which is assumed in Theorem 1.2 is guaranteed by the following corollary.

Corollary 4.2. *Let $s_A \geq 12$ and $\sigma_A > 0$. Then, for every $R > 0$, there exist $\eta = \eta(R, s_A, \sigma_A)$ such that for every $A_0 \in G^{s_A}_{\sigma_A}$, with $\|A_0\|_{G^{s_A}_{\sigma_A}} \leq R$ and $\sigma_0 \in [0, \sigma_A)$ there exists a $T_0 > 0$ and unique local solution $A \in C([0, T_0], G^{s_A}_{\sigma_0})$ of the DNLS equation* (5) *with $\sup_{T\in[0,\sigma_A/\eta]} \|A(T)\|_{G^s_{\sigma(T)}} \leq R$.*

Proof. As above choose $\sigma(T) := \sigma_A - \eta T$ and stop for $\sigma(T_0) = \sigma_0$. □

5 Modulational Gevrey spaces

In the last section we have seen that with an initial exponential decay in Fourier space for $|K| \to \infty$ we can create an artificial smoothing which allows us to control the derivative in front of the nonlinear terms of the DNLS equation. In the nonlinear Klein-Gordon equation (4) the DNLS equation is the lowest order approximation for the modes located at the wave number $k = 1$, in particular the derivative in front of the nonlinear terms of the DNLS equation correspond to the vanishing of the nonlinear terms of the nonlinear Klein-Gordon equation (4) at the wave number $k = 1$. For the DNLS approximation the associated modes decay with an exponential rate around the wave number $k = 1$, see the left panel of Figure 1. However, by nonlinear interaction small peaks with width of order $\mathcal{O}(\varepsilon)$ are created at odd integer multiples of the basic wave number $k_0 = 1$. See the right panel of Figure 1. This means that the solutions of the nonlinear Klein-Gordon equation (4) will have a Fourier mode distribution which is bounded from above by a multiple of $\dfrac{1}{\vartheta_{\alpha/\varepsilon}}$, where

$$\vartheta_\beta(k) := \exp\Big(\beta \inf_{m\in\mathbb{Z}_{odd}} |k - mk_0|\Big)$$

or equivalently

$$\frac{1}{\vartheta_\beta(k)} = \sup_{m\in\mathbb{Z}_{odd}} e^{-\beta|k-mk_0|}$$

for $\beta \geq 0$. We define a number of spaces to combine these facts with the ideas from Section 4 for the DNLS equation (2) in order to handle the nonlinear Klein-Gordon equation (4). For estimating the solutions of the original system we use energy estimates and so the nonlinear Klein-Gordon equation (4) will be solved in the L^2-based space

$$\mathcal{M}^s_\beta = \{u : \mathbb{R} \to \mathbb{C} : \|u\|_{\mathcal{M}_\beta} = \|\widehat{u}(k)\vartheta_\beta(k)(1 + k^2)^{s/2}\|_{L^2(dk)} < \infty\}.$$

As a consequence for $u \in \mathcal{M}^s_\beta$, with $\beta > 0$, the modes bounded away from integer multiples of the basic wave number $k_0 = 1$ are exponentially small w.r.t. ε, i.e., these modes are of order $\mathcal{O}(e^{-r/\varepsilon})$ for $0 < \varepsilon \ll 1$ with an $\mathcal{O}(1)$-bound $r > 0$. Due to the L^2-scaling properties, the DNLS approximation is of order $\mathcal{O}(1)$ in the $\mathcal{M}^s_\beta$-spaces and not of the formal order $\mathcal{O}(\varepsilon^{1/2})$. Therefore, we additionally define the spaces

$$\mathcal{W}^s_\beta = \{u : \mathbb{R} \to \mathbb{C} : \|u\|_{\mathcal{W}^s_\beta} = \|\widehat{u}(k)\vartheta_\beta(k)(1 + k^2)^{s/2}\|_{L^1(dk)} < \infty\}$$

for which the DNLS approximation is of order $\mathcal{O}(\varepsilon^{1/2})$.

For the subsequent error estimates we need that these spaces are closed under point-wise multiplication.

Lemma 5.1. *For all* $\beta \geq 0$ *and* $s > 1/2$ *we have*

$$\|uv\|_{\mathcal{M}^s_\beta} \leq \|u\|_{\mathcal{M}^s_\beta} \|v\|_{\mathcal{M}^s_\beta}.$$

Moreover, for all $\beta, s \geq 0$ *we have*

$$\|uv\|_{\mathcal{M}^s_\beta} \leq \|u\|_{\mathcal{W}^s_\beta} \|v\|_{\mathcal{M}^s_\beta}. \tag{13}$$

Proof. The estimates immediately follow from

$$\begin{aligned} \|uv\|_{\mathcal{M}^s_\beta} &= \|(\widehat{u} * \widehat{v})\vartheta_\beta\|_{L^2_s} \\ &\leq \|\widehat{u}\vartheta_\beta\|_{L^1} \|\widehat{v}\vartheta_\beta\|_{L^2_s} + \|\widehat{u}\vartheta_\beta\|_{L^2_s} \|\widehat{v}\vartheta_\beta\|_{L^1} \\ &\leq \|u\|_{\mathcal{W}^0_\beta} \|v\|_{\mathcal{M}^s_\beta} + \|u\|_{\mathcal{M}^s_\beta} \|v\|_{\mathcal{W}^0_\beta} \end{aligned}$$

due to Young's inequality for convolutions, Sobolev's embedding

$$\|u\|_{\mathcal{W}^s_\beta} \leq C \|v\|_{\mathcal{M}^{s+\delta}_\beta}$$

for $\delta > 1/2$, and the inequality

$$\begin{aligned} \frac{1}{\vartheta_\beta(k-l)\vartheta_\beta(l)} &= \sup_{m \in \mathbb{Z}_{odd}} \left(e^{-\beta|k-l-mk_0|} \right) \sup_{m \in \mathbb{Z}_{odd}} \left(e^{-\beta|l-mk_0|} \right) \\ &\leq \sup_{m \in \mathbb{Z}_{odd}} \left(e^{-\beta|k-mk_0|} \right) = \frac{1}{\vartheta_\beta(k)}. \end{aligned}$$

□

Remark 5.2. Subsequently, (13) will be used to estimate combinations of the approximation with the error. The DNLS approximation will be estimated in the space $\mathcal{W}^s_\beta$ where it is of order $\mathcal{O}(\varepsilon^{1/2})$ and the error will be estimated in the space $\mathcal{M}^s_\beta$.

The initial value problem for (4) for initial conditions of order $\mathcal{O}(\varepsilon^{1/2})$ can be solved in these spaces on a time interval of length $\mathcal{O}(1/\varepsilon)$ using the variation of constant formula, using the fact that we have a cubic nonlinearity and the fact that these spaces are closed under multiplication. In order to bound the error not only on the short $\mathcal{O}(1/\varepsilon)$-time scale but also on the natural $\mathcal{O}(1/\varepsilon^2)$-time scale of the DNLS approximation we use the spaces $\mathcal{M}^s_\beta$, but now with time-dependent β. Similar to above we choose

$$\beta(t) = \sigma_0/\varepsilon - \eta\varepsilon t, \tag{14}$$

with constants σ_0, $\eta > 0$, which can be chosen independently of $0 < \varepsilon \ll 1$. Note that (14) is the scaled version of $\sigma(T) = \sigma_0 - \eta T$ defined in Theorem 4.1 and that $T_1 = \sigma_0/\eta$. If $\widehat{A}$ is initially in a space $G^{s+1}_{\sigma_0}$, then according to Remark 1.1 the DNLS approximation is initially in a space $\mathcal{W}^s_{\sigma_0/\varepsilon}$. This is the reason why $\beta(t)$ starts with σ_0/ε. The decay $-\eta\varepsilon t$ allows to consider t on the natural $\mathcal{O}(1/\varepsilon^2)$-time scale of the DNLS approximation. It turns out that subsequently choosing $\eta = \mathcal{O}(1)$ is sufficient for our purposes.

In the subsequent sections we explain in detail how with this approach all problems to come to the long $\mathcal{O}(1/\varepsilon^2)$-time scale, found in Section 3, can be solved.

6 Separation of the modes

In order to obtain a bound for the error on the long $\mathcal{O}(1/\varepsilon^2)$-time scale, independently of the small perturbation parameter $0 < \varepsilon \ll 1$, we have to get rid of the term $\varepsilon L_c(R)$ in (9). Except at the resonant wave numbers this term is oscillatory and can be removed by a near identity change of variables. In the last sections we explained our strategy to get rid of the total resonance **(TR)** and of the second order resonance **(SOR)**.

Hence for the error estimates we separate the modes in three parts. The first part which is denoted by R_n has a support near the odd integer multiples of the basic wave number $k_0 = 1$ excluding neighborhoods around the basic wave numbers $\pm k_0 = \pm 1$. It will be handled with normal form transformations and energy estimates. The second part which is denoted by R_r has a support which is bounded away from the odd integer multiples of the basic wave number $k_0 = 1$ and will linearly be exponentially damped by our choice of time-dependent weights. The third part which is denoted by R_c has support near the basic wave numbers $\pm k_0 = \pm 1$ and will be handled with the ideas which have been explained above in Section 4 and Section 5 and with energy estimates. The index n stands for normal form, r for rest, and c for critical.

In detail, we define for a $\delta_r > 0$ small, but independent of $0 < \varepsilon \ll 1$, the mode filter

$$\widehat{E}_r(k) = \begin{cases} 1, & \text{for } \inf_{n \in \mathbb{Z}_{odd}} |k-n| > \delta_r \\ 0, & \text{else,} \end{cases}$$

the mode filter

$$\widehat{E}_c(k) = \begin{cases} 1, & \text{for } \inf_{n \in \{-1,1\}} |k-n| \leq \delta_r \\ 0, & \text{else,} \end{cases}$$

and finally the mode filter $\widehat{E}_n = 1 - \widehat{E}_r - \widehat{E}_c$.

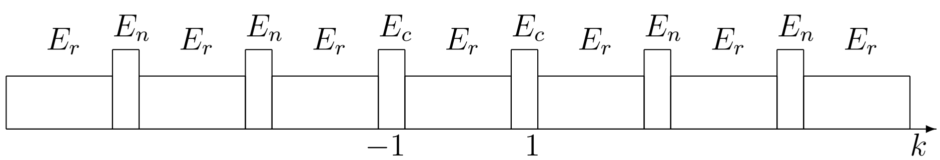


Figure 3: The support of the mode filters E_r, E_n, and E_c in Fourier space.

We use these projections to separate the error $R = R_r + R_n + R_c$ in three parts, namely $R_r = E_r R_r$, $R_c = E_c R_c$, and $R_n = E_n R_n$. These new variables satisfy

$$\begin{aligned} \partial_t R_r &= \Lambda R_r + \varepsilon E_r L_c(R) + \varepsilon^2 E_r G, & (15) \\ \partial_t R_n &= \Lambda R_n + \varepsilon E_n L_c(R) + \varepsilon^2 E_n G, & (16) \\ \partial_t R_c &= \Lambda R_c + \varepsilon E_c L_c(R) + \varepsilon^2 E_c G, & (17) \end{aligned}$$

where

$$\varepsilon^2 G = \varepsilon^2 L_s(R) + \varepsilon^{\widetilde{\beta}+1/2} L_r(R) + \varepsilon^{-\widetilde{\beta}} \mathrm{Res}(\varepsilon^{1/2}\psi).$$

7 The normal form transform

As already said, in order to come to the long $\mathcal{O}(1/\varepsilon^2)$-time scale, we have to get rid of the terms

$$\begin{aligned} &\varepsilon \widehat{E_j L_c(R)}(k,t) \\ &= \varepsilon \frac{3}{2} i\rho(k) \widehat{E}_j(k) \begin{pmatrix} -(\widehat{a}_1 + \widehat{a}_{-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \\ (\widehat{a}_1 + \widehat{a}_{-1})^{*2} * (\widehat{R}_1 + \widehat{R}_{-1}) \end{pmatrix} (k,t), \end{aligned}$$

for $j = n, r, c$ in (15)-(17). In a first step we simplify the $\varepsilon E_j L_c(R)$ for $j = n, r, c$ by eliminating all non-resonant terms by normal form transformations. The $\varepsilon E_j L_c(R)$ for $j = n, r, c$ are sums of trilinear mappings w.r.t. a_{j_1}, a_{j_2} and R_{j_3}, with $j_1, j_2, j_3 \in \{-1, 1\}$.

Remark 7.1. In order to eliminate a trilinear term $\varepsilon B(a_{j_1}, a_{j_2}, R_{j_3})$ of the form

$$\begin{aligned} & \widehat{B(a_{j_1}, a_{j_2}}, R_{j_3}) \\ & = \int\int b(k, k-k_1, k_1-k_2, k_2) \widehat{a}_{j_1}(k-k_1) \widehat{a}_{j_2}(k_1-k_2) \widehat{R}_{j_3}(k_2) dk_2 dk_1 \end{aligned}$$

from the equation of R_j by a near identity transformation $w_j = R_j + \varepsilon M(a_{j_1}, a_{j_2}, R_{j_3})$ we have to choose

$$\begin{aligned} & \widehat{M(a_{j_1}, a_{j_2}}, R_{j_3}) \\ & = \int\int m(k, k-k_1, k_1-k_2, k_2) \widehat{a}_{j_1}(k-k_1) \widehat{a}_{j_2}(k_1-k_2) \widehat{R}_{j_3}(k_2) dk_2 dk_1 \end{aligned}$$

with

$$m(k, k-k_1, k_1-k_2, k_2) = \frac{b(k, k-k_1, k_1-k_2, k_2)}{-j\omega(k) - \omega(j_1) - \omega(j_2) + j_3\omega(k-j_1-j_2)},$$

cf. [SU17, §11]. For the terms which will be eliminated the nominator b is bounded and the denominator is bounded away from zero.

By the analysis of the denominator, made in Section 3, we can eliminate all terms except of the total resonant terms and second order resonant terms. See Remark 7.4 for more details. If we do so, after the transformation we obtain a system

$$\begin{aligned} \partial_t w_r &= \Lambda w_r + \varepsilon E_r L_c(w_r) + \varepsilon^2 H_r, & (18) \\ \partial_t w_n &= \Lambda w_n + \varepsilon B_1(a_1, a_{-1}, w_n) + \varepsilon^2 H_n, & (19) \\ \partial_t w_c &= \Lambda w_c + \varepsilon B_2(a_1, a_{-1}, w_c) & (20) \\ & \quad + \varepsilon B_3(a_{-1}, a_{-1}, w_c) + \varepsilon B_4(a_1, a_1, w_c) + \varepsilon^2 H_c, \end{aligned}$$

where the B_j are smooth trilinear mappings in their arguments and $\varepsilon^2 H_{r,n,c} = \mathcal{O}(\varepsilon^2)$ with properties specified below. The properties of the normal form transformation are summarized in the following lemma.

Lemma 7.2. *Let $s > 1/2$ and $\sigma_0 \geq 0$. The transformation*

$$\mathcal{T}^\varepsilon : \begin{cases} (\mathcal{M}^s_{\sigma/\varepsilon})^3 \to (\mathcal{M}^s_{\sigma/\varepsilon})^3, \\ (R_n, R_r, R_c) \mapsto (w_n, w_r, w_c), \end{cases}$$

is a small perturbation of identity. For all $\sigma \in [0, \sigma_0]$ the mapping is analytic. For all $C_1 > 0$ there exists an $\varepsilon_0 > 0$ such for all $\varepsilon \in (0, \varepsilon_0)$ and all $\sigma \in [0, \sigma_0]$ the following holds. For all (w_n, w_r, w_c) with $\|(w_n, w_r, w_c)\|_{\mathcal{M}^s_{\sigma/\varepsilon}} \leq C_1$ there exists an analytic inverse. All bounds are independent of $\varepsilon \in (0, \varepsilon_0)$ and $\sigma \in [0, \sigma_0]$.

With this lemma we immediately have

Corollary 7.3. *Let $s > 1/2$ and*

$$M = \|w\|_{\mathcal{M}^s_{\beta(t)}} = \|w_r\|_{\mathcal{M}^s_{\beta(t)}} + \|w_n\|_{\mathcal{M}^s_{\beta(t)}} + \|w_c\|_{\mathcal{M}^s_{\beta(t)}},$$

with $\beta(t)$ defined in (14). There exist constants C_1, $C_3 > 0$ independent of M and $\varepsilon \in (0, \varepsilon_0]$, with $\varepsilon_0 > 0$ from Lemma 7.2, and a monotonically increasing function $C_2(M) > 0$, independent of $\varepsilon \in (0, \varepsilon_0]$, such that

$$\|\varepsilon^2 H_j\|_{\mathcal{M}^s_{\beta(t)}} \leq C_1 \varepsilon^2 \|w\|_{\mathcal{M}^s_{\beta(t)}} + C_2(M) \varepsilon^{\widetilde{\beta}+1/2} \|w\|^2_{\mathcal{M}^s_{\beta(t)}} + C_3 \varepsilon^2,$$

for $j = r, n, c$.

Remark 7.4. We have transformed the term $\varepsilon E_r L_c(R)$ in a term $\varepsilon E_r L_c(w_r)$. The last term contains totally resonant terms, for which we know that the denominator in the above normal form transformations would vanish identically. Hence, they couldn't be eliminated in this way. In order to explain subsequently a few possible improvements to our approach we handle the totally resonant terms in $\varepsilon E_r L_c(w_r)$ differently than the other totally resonant terms. The term $\varepsilon E_r L_c(w_r)$ will initially be exponentially small w.r.t ε, and to grow to an order $\mathcal{O}(\varepsilon)$ it will take an $\mathcal{O}(1/\varepsilon^2)$-time scale. This observation allows us for general dispersive systems to reduce the number of necessary non-resonance conditions. The other totally resonant terms $B_1(a_1, a_{-1}, w_n)$ and $B_2(a_1, a_{-1}, w_c)$, are initially not exponentially small, and so they will be controlled by energy estimates in the following. For the second order resonant terms $B_3(a_{-1}, a_{-1}, w_c)$ and $B_4(a_1, a_1, w_c)$ the denominator in the above normal form transformation would vanish quadratically for $k = \pm 1$. Since the nominator only vanishes linearly at these wave numbers the normal form transform would be unbounded. Therefore, the second order resonant terms will be handled with the ideas presented in Section 4 and Section 5.

8 The final error estimates

In order to estimate the solutions of the equations (18)-(20) for the error we use the modulational Gevrey spaces introduced in Section 5. Introducing the new weighted variables

$$\widehat{W}_j(k) = \widehat{w}_j(k)\vartheta_\beta(k)$$

for $j = r, n, c$ allows us to work in classical Sobolev spaces. We find

$$\begin{aligned}
\partial_t W_r &= \Lambda W_r + \Gamma W_r + \varepsilon E_r \widetilde{L}_c(W_r) + \varepsilon^2 \widetilde{H}_r, & (21)\\
\partial_t W_n &= \Lambda W_n + \Gamma W_n + \varepsilon \widetilde{B}_1(a_1, a_{-1}, W_n) + \varepsilon^2 \widetilde{H}_n, & (22)\\
\partial_t W_c &= \Lambda W_c + \Gamma W_c + \varepsilon \widetilde{B}_2(a_1, a_{-1}, W_c) & (23)\\
&\quad + \varepsilon \widetilde{B}_3(a_{-1}, a_{-1}, w_c) + \varepsilon \widetilde{B}_4(a_1, a_1, W_c) + \varepsilon^2 \widetilde{H}_c,
\end{aligned}$$

where the operator Γ is defined in Fourier space by

$$\widehat{\Gamma W}(k) = -\eta\varepsilon(\inf_{m\in\mathbb{Z}_{odd}} |k - mk_0|)\widehat{W}(k). \qquad (24)$$

The trilinear mappings B_j from (18)-(20) transform into the $\widetilde{B}_j$ which are again smooth trilinear mappings in their arguments. They are estimated below in detail. The remaining terms H_j from (18)-(20) transform into the $\widetilde{H}_j$ whose properties are specified in the subsequent lemma.

Lemma 8.1. *Let $s > 1/2$ and*

$$M = \|W\|_{H^s} = \|W_r\|_{H^s} + \|W_n\|_{H^s} + \|W_c\|_{H^s}.$$

There are constants C_1, $C_3 > 0$ independent of M and $\varepsilon \in (0, \varepsilon_0]$, with $\varepsilon_0 > 0$ from Lemma 7.2, and a monotonically increasing function $C_2(M) > 0$, independent of $\varepsilon \in (0, \varepsilon_0]$, such that

$$\|\varepsilon^2 \widetilde{H}_j\|_{H^s} \leq C_1\varepsilon^2\|W\|_{H^s} + C_2(M)\varepsilon^{\widetilde{\beta}+1/2}\|W\|^2_{H^s} + C_3\varepsilon^2,$$

for $j = r, n, c$.

Proof. The lemma is mainly a reformulation of Corollary 7.3. The missing details can be found below in the estimates for the $\widetilde{B}_j$. □

In order to estimate the solutions W_j of the equations (21)-(23) we use energy estimates, i.e., we multiply the equation for W_j with W_j for $j = r, n, c$ and take the H^s-scalar product $(\cdot,\cdot)_s$. We find

$$\begin{aligned}
\partial_t(W_r, W_r)_s &= 2\mathrm{Re}(s_1 + s_2 + s_3 + s_4), \\
\partial_t(W_n, W_n)_s &= 2\mathrm{Re}(s_5 + s_6 + s_7 + s_8), \\
\partial_t(W_c, W_c)_s &= 2\mathrm{Re}(s_9 + s_{10} + s_{11} + s_{12} + s_{13} + s_{14}),
\end{aligned}$$

with

$$\begin{aligned}
s_1 &= (W_r, \Lambda W_r)_s, & s_2 &= (W_r, \Gamma W_r)_s, \\
s_3 &= (W_r, \varepsilon E_r \widetilde{L}_c(W_r))_s, & s_4 &= (W_r, \varepsilon^2 \widetilde{H}_r)_s,
\end{aligned}$$

with

$$\begin{aligned}
s_5 &= (W_n, \Lambda W_n)_s, & s_6 &= (W_n, \Gamma W_n)_s, \\
s_7 &= (W_n, \varepsilon \widetilde{B}_1(a_1, a_{-1}, W_n))_s, & s_8 &= (W_n, \varepsilon^2 \widetilde{H}_n)_s,
\end{aligned}$$

and

$$\begin{aligned}
s_9 &= (W_c, \Lambda W_c)_s, & s_{10} &= (W_c, \Gamma W_c)_s, \\
s_{11} &= (W_c, \varepsilon \widetilde{B}_2(a_1, a_{-1}, W_c))_s, & s_{12} &= (W_c, \varepsilon \widetilde{B}_3(a_{-1}, a_{-1}, W_c))_s, \\
s_{13} &= (W_c, \varepsilon \widetilde{B}_4(a_1, a_1, W_c))_s, & s_{14} &= (W_c, \varepsilon^2 \widetilde{H}_c,)_s.
\end{aligned}$$

In between the terms $s_1, \ldots, s_{14}$ there are terms which vanish identically and terms which do not make any difficulties since they have an ε^2 in front. The dangerous terms are the ones which only have an ε in front, namely $s_3, s_7, s_{11}, s_{12}, s_{13}$. They will be estimated through integration by parts, such as the totally resonant terms s_7 and s_{11}, or by the damping terms s_2, s_6, s_{10}, such as s_3 or the second order resonant terms s_{12}, s_{13}.

We start with the terms which vanish identically, namely
$\mathbf{s_1}, \mathbf{s_5}, \mathbf{s_9}$: Due to the skew symmetry of Λ we immediately have

$$s_1 = s_5 = s_9 = 0.$$

For the terms which have an ε^2 in front we proceed as follows and find:
$\mathbf{s_4}, \mathbf{s_8}, \mathbf{s_{14}}$: Using Lemma 8.1, the Cauchy-Schwarz inequality, and $a \leq 1 + a^2$ in the last term yields

$$|s_j| \;\; \leq \;\; C_1 \varepsilon^2 \|W\|_{H^s}^2 + C_2(M) \varepsilon^{\widetilde{\beta}+1/2} \|W\|_{H^s}^3 + C_3 \varepsilon^2 (1 + \|W\|_{H^s}^2),$$

for $j = 4, 8, 14$ using the notation from Lemma 8.1.

Next we go on with the rest of the W_r-equation.

$\mathbf{s_2}$: is the good term in the W_r-equation. There is a $\sigma > 0$ independent of $0 < \varepsilon^2 \ll 1$ such that

$$2\mathrm{Re}(s_2) = 2\mathrm{Re}((W_r, \Gamma W_r)_s \leq -\alpha(\eta)\varepsilon(W_r, W_r)_s.$$

We have $\alpha(\eta) \to \infty$ for $\eta \to \infty$, cf. (24).

$\mathbf{s_3}$: is the dangerous term in the W_r-equation which, however, can be estimated by the s_2-term. For the s_3-term we have

$$|s_3| = |(W_r, \varepsilon E_r \widetilde{L}_c(W_r))_s| \leq C_{s_3}\varepsilon(W_r, W_r)_s$$

for a contant $C_{s_3} = C_{s_3}(\psi)$ independent of $0 < \varepsilon^2 \ll 1$.

Next we go on with the rest of the W_n-equation.

$\mathbf{s_6}$: is the good term in the W_n-equation. However, for our purposes it is sufficient to have $2\mathrm{Re}(s_6) \leq 0$.

$\mathbf{s_7}$: In Fourier space we have to control terms of the form

$$\begin{aligned}
&\varepsilon \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \overline{\widehat{W}_n(k)} i\vartheta(k)\rho(k)\big(\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\vartheta^{-1}(l)\widehat{W}_n(l)\big)dmdldk \\
&\qquad + \varepsilon \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{W}_n(k)\overline{i\vartheta(k)\rho(k)\big(\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\vartheta^{-1}(l)\widehat{W}_n(l)\big)}dmdldk \\
&= \varepsilon \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \overline{\widehat{W}_n(k)} i\vartheta(k)\rho(k)\big(\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\vartheta^{-1}(l)\widehat{W}_n(l)\big)dmdldk \\
&\qquad - \varepsilon \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{W}_n(l)\overline{i\vartheta(l)\rho(l)\big(\widehat{a}_{j_1}(l-m)\widehat{a}_{-j_1}(m-k)\vartheta^{-1}(k)\widehat{W}_n(k)\big)}dmdkdl \\
&= \varepsilon \int_{\mathbb{R}}\int_{\mathbb{R}} \overline{\widehat{W}_n(k)}\big(Q(k,k-l,l)\widehat{W}_n(l)\big)dkdl
\end{aligned}$$

with

$$\begin{aligned}
Q(k,k-l,l) &= \int_{\mathbb{R}} i\vartheta(k)\rho(k)\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\vartheta^{-1}(l) \\
&\qquad - \overline{i\vartheta(l)\rho(l)\widehat{a}_{j_1}(l-m)\widehat{a}_{-j_1}(m-k)\vartheta^{-1}(k)}dm \\
&= \int_{\mathbb{R}} i\vartheta(k)\rho(k)\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\vartheta^{-1}(l) \\
&\qquad - \overline{i\vartheta(l)\rho(l)}\widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)\overline{\vartheta^{-1}(k)}dm \\
&= \varrho_1(k,l)\int_{\mathbb{R}} \widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)dm,
\end{aligned}$$

with

$$\varrho_1(k,l) = (i\vartheta(k)\rho(k)\vartheta^{-1}(l) - i\vartheta(l)\rho(l)\vartheta^{-1}(k)),$$

and where we used that the product $a_{j_1}a_{-j_1}$ is real. We have that $(k,l) \mapsto \varrho_1(k,l)$ is smooth and satisfies $\varrho_1(k,k) = 0$ such that finally $|\varrho_1(k,l)| \leq C|k-l|$. Since $\int_{\mathbb{R}} \widehat{a}_{j_1}(k-m)\widehat{a}_{-j_1}(m-l)dm$ is strongly concentrated at $k-l=0$ we gain another power of ε such that finally

$$s_7 = \mathcal{O}(\varepsilon^2).$$

Finally we come to the remaining terms of the W_c-equation.
$\mathbf{s_{11}}$: The totally resonant terms $\mathbf{s_{11}}$ are handled line for line as the totally resonant terms $\mathbf{s_7}$ and so we also have

$$s_{11} = \mathcal{O}(\varepsilon^2).$$

$\mathbf{s_{10}}$: is the good term which allows us to handle the second order resonant terms. We have

$$2\mathrm{Re}(s_{10}) = 2\mathrm{Re}(W_c, \Gamma W_c)_s \leq -\eta\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s$$

where $|k-1|_{op}$ is defined by the multiplier $|k-1|$ in Fourier space.
$\mathbf{s_{12}}, \mathbf{s_{13}}$: In the following $W_{c,1}$ denotes the part of W_c located at $k=1$ and $W_{c,-1}$ the part of W_c located at $k=-1$. Then the second order resonant terms are estimated by

$$\begin{aligned} |s_{12}| + |s_{13}| &\leq C|(W_{c.-1}, \varepsilon\widetilde{B}_3(a_{-1}, a_{-1}, W_{c,1}))_s| + C|(W_{c,1}, \varepsilon\widetilde{B}_4(a_1, a_1, W_{c,-1}))_s| \\ &\leq C\varepsilon\|\rho^{1/2}W_{c,-1}\|_{H^s}\|\rho^{1/2}\widetilde{B}_3(a_{-1}, a_{-1}, W_{c,1})\|_{H^s} \\ &\quad +C\varepsilon\|\rho^{1/2}W_{c,1}\|_{H^s}\|\rho^{1/2}\widetilde{B}_4(a_1, a_1, W_{c,-1})\|_{H^s}. \end{aligned}$$

The term $\|\rho^{1/2}\widetilde{B}_4(a_1, a_1, W_{c,-1})\|_{L^2}^2$ is in Fourier space of the form

$$\int\int\int |\rho^{1/2}(k)\widehat{a}_1(k-l)\widehat{a}_1(l-m)\widehat{W}_{c,-1}(m)|^2 dmdldk \leq s_{15} + s_{16},$$

with

$$\begin{aligned} s_{15} &= \int\int\int |(\rho(k) - \rho(m+2))||\widehat{a}_1(k-l)\widehat{a}_1(l-m)\widehat{W}_{c,-1}(m)|^2 dmdldk, \\ s_{16} &= \int\int\int |\widehat{a}_1(k-l)\widehat{a}_1(l-m)|^2|\rho(m+2)||\widehat{W}_{c,-1}(m)|^2 dmdldk. \end{aligned}$$

We use that

$$\rho(k) - \rho(m+2) = \rho(m+2+k-m-2) - \rho(m+2) = \mathcal{O}(|k-m-2|),$$

that $\int \widehat{a}_1(k-l)\widehat{a}_1(l-m)dl$ is strongly concentrated at $k-m \approx 2$ and Young's inequality to obtain a bound

$$s_{15} \leq C\varepsilon \|W_{c,-1}\|^2.$$

Less complicated is the bound

$$s_{16} \leq C\|\Gamma^{1/2} W_{c,-1}\|^2$$

since $\rho(m+2) = \mathcal{O}(|m+1|)$. Then finally we obtain

$$\begin{aligned} |s_{12}| + |s_{13}| &\leq C\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s + C\varepsilon^{3/2}\|\Gamma^{1/2}W_c\|_{H^s}\|W_c\|_{H^s} \\ &\leq 2C\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s + C\varepsilon^2\|W_c\|^2_{H^s} \\ &\leq C_\psi\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s + C_1\varepsilon^2\|W_c\|^2_{H^s} \end{aligned}$$

where we used $\varepsilon^{3/2}ab \leq \varepsilon a^2 + \varepsilon^2 b^2$. This defines the constrant C_ψ and we may increase the original constant C_1.
Summary: Collecting all estimates gives for

$$E_s = (W_r, W_r)_s + (W_n, W_n)_s + (W_c, W_c)_s$$

that

$$\begin{aligned} \partial_t E_s &\leq 2\mathrm{Re}(s_2) + 2|s_3| + 2|s_4| + 2\mathrm{Re}(s_7) + 2|s_8| \\ &\quad +2\mathrm{Re}(s_{10}) + 2\mathrm{Re}(s_{11}) + 2|s_{12}| + 2|s_{13}| + 2|s_{14}| \\ &\leq -\alpha(\eta)\varepsilon(W_r, W_r)_s + C_{s_3}\varepsilon(W_r, W_r)_s \\ &\quad -\eta\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s + C_\psi\varepsilon(\Gamma^{1/2}W_c, \Gamma^{1/2}W_c)_s \\ &\quad +2C_1\varepsilon^2 E_s + C_2(M)\varepsilon^{\widetilde{\beta}+1/2}E_s^{3/2} + C_3\varepsilon^2(1+E_s). \end{aligned}$$

The third and fourth line can be made negative by choosing η sufficiently large, but independent of the small perturbation parameter $0 < \varepsilon \ll 1$ such that we finally have

$$\partial_t E_s \leq 2C_1\varepsilon^2 E_s + C_2(M)\varepsilon^{\widetilde{\beta}+1/2}E_s^{3/2} + C_3\varepsilon^2(1+E_s).$$

Choosing $C_2(M)\varepsilon^{\widetilde{\beta}-3/2}E_s^{1/2} \leq 1$ yields

$$\partial_t E_s \leq (2C_1 + C_3 + 1)\varepsilon^2 E_s + C_3\varepsilon^2. \tag{25}$$

Applying Gronwall's inequality yields for all $t \in [0, T_1/\varepsilon^2]$ that

$$E_s(t) \leq C_3 T_1 e^{(2C_1+C_3+1)T_1} =: M$$

independent of $\varepsilon \in (0, \varepsilon_0)$ where $\varepsilon_0 > 0$ had to be chosen so small that $C_2(M)\varepsilon^{\widetilde{\beta}-3/2}M^{1/2} \leq 1$. Therefore, we are done. $\square$

9 Discussion

In this section we make a number of remarks about possible improvements and generalizations.

Remark 9.1. For an arbitrary, but fixed, $\widetilde{\beta} > 3/2$ the improved approximation can be constructed in such a way that the residual term $\varepsilon^{-\widetilde{\beta}}\text{Res}(\varepsilon^{1/2}\psi)$ is of order $\mathcal{O}(\varepsilon^2)$, cf. Appendix A. Hence the error made by this improved approximation is of order $\mathcal{O}(\varepsilon^{\widetilde{\beta}})$ in some Sobolev norm.

Remark 9.2. It is obvious by the proof that the assumption on the solutions of the DNLS equation (5), namely $A \in C([0, T_0], G_{\sigma_0}^{s_A})$, can be replaced by the weaker assumption that we take a solution constructed in Theorem 4.1 with $s = s_A$, $\sigma_A = \sigma_0$, and $A(T) \in G_{\sigma(T)}^{s}$ with $\sigma(T) = \sigma_0 - \eta T$.

Remark 9.3. From (15) to (18) we have eliminated all terms of order $\mathcal{O}(\varepsilon)$ except of the totally resonant ones. This is not necessary since finally $\varepsilon E_r \widetilde{L}_c(W_r)$ appears in an equation where W_r is exponentially damped. Hence other terms can be kept and other resonances can be handled as long as they are bounded away from odd integer multiples of the basic wave number $k_0 = 1$, cf. [DHSZ16].

Remark 9.4. We have demonstrated that the DNLS approximation makes correct predictions about the dynamics of our chosen nonlinear Klein-Gordon equation (4). The question about possible generalizations and about the possible transfer to more complicated systems occurs. First of all we would like to mention that the problems with the total resonance and the second order resonance occur for all non-trivial systems for which the DNLS approximation can be derived. On the one hand with this respect our system is not more complicated than necessary. On the other hand the chosen nonlinear Klein-Gordon equation (4) is sufficiently complicated to contain all principle difficulties which have to be overcome.

Remark 9.5. Other additional difficulties one could think of have been handled in our situation before. For instance quadratic terms in the original systems can be eliminated completely with a normal form transform for Klein-Gordon models. For other more complicated original systems additional quadratic or quintic resonances can occur. It is not obvious how existing methods to handle such resonances interplay with the presented approach of this paper. The same is true for quasilinear systems such as the water wave problem. This will be the topic of future research.

Remark 9.6. It is the topic of parallel research to prove a DNLS approximation result for initial conditions which are not analytic in a strip of the complex plane but only live in a Sobolev space. In this case the totally resonant terms have to be handled with energy estimates again. New ideas are needed to handle the second order resonant terms. Moreover, all other terms of order $\mathcal{O}(\varepsilon)$ in the error equations (9) have to be eliminated by normal form transformations, i.e., no other resonances can be allowed. This is different to the situation in this paper where additional resonances bounded away from integer multiples of the basic wave number $k_0 = 1$ can be allowed due to the exponential smallness of these modes initially, cf. Remark 9.3.

A Higher order DNLS approximation and estimates for the residual

In the previous sections we concentrated on estimating the error made by a DNLS approximation. In order to do so we used an improved approximation introduced in (8) with the property that the residual term $\varepsilon^{-\widetilde{\beta}}\mathrm{Res}(\varepsilon^{1/2}\psi)$ is of order $\mathcal{O}(\varepsilon^2)$ in the equations for the error (9). Since similar constructions can be found in several papers, cf. [SU17], we took this part out of the line of proof. Nevertheless, we will provide a few details in this appendix.

First we show how to construct an higher order approximation. We define the residual

$$\mathrm{Res}_u(u) = -\partial_t^2 u + \partial_x^2 u - u + \varrho(\partial_x) u^3 \tag{26}$$

for the nonlinear Klein-Gordon equation (4). For the computation of an higher order approximation we need a Taylor expansion of

$$\varrho(ik) = \frac{k^2-1}{k^2+1}, \qquad \text{resp.} \qquad \varrho(\partial_x) = -(1-\partial_x^2)^{-1}(1+\partial_x^2)$$

at $k_0 = 1$ and at other odd integer multiples of $k_0 = 1$. We find for instance

$$\varrho(i+\varepsilon\partial_X) = -i\varepsilon\partial_X + \frac{1}{2}\varepsilon^2\partial_X^2 + \frac{1}{4}\varepsilon^4\partial_X^4 + \frac{1}{4}i\varepsilon^5\partial_X^5 + \frac{1}{8}\varepsilon^6\partial_X^6 + \mathcal{O}(\varepsilon^7).$$

The improved ansatz is given

$$\begin{aligned} \varepsilon^{1/2}\psi(x,t) = & \sum_{n\in\mathbb{N}_{odd},|n|\leq N} \sum_{j=0}^{N} \varepsilon^{p(n)+j} A_{n,j}(X,T)\mathbf{E}^n \\ = & \ \varepsilon^{1/2}A_{1,0}(X,T)\mathbf{E} + \varepsilon^{3/2}A_{1,1}(X,T)\mathbf{E} + \varepsilon^{3/2}A_{3,0}(X,T)\mathbf{E}^3 + c.c. + h.o.t., \end{aligned}$$

with $\mathbf{E} = e^{i(k_0x-\omega_0t)}$, $X = \varepsilon(x-c_gt)$, $T = \varepsilon^2 t$, and $p(n) = (||n|-1|+1)/2$, with a fixed chosen $N \in \mathbb{N}$. For expository reasons we restrict ourselves in the following to the three terms explicitly displayed in the ansatz. They represent the three essential types of approximation equations which occur.

Plugging the ansatz into (26) and equating the coefficients in front of $\mathbf{E}$ to zero gives as before the linear dispersion relation $\omega_0^2 = k_0^2 + 1$ at $\mathcal{O}(\varepsilon^{1/2})$ and the linear group velocity $c_g = k_0/\omega_0$ at $\mathcal{O}(\varepsilon^{3/2})$. At $\mathcal{O}(\varepsilon^{5/2})$ we again obtain the DNLS equation

$$2i\omega_0\partial_T A_{1,0} = (1-c_g^2)\partial_X^2 A_{1,0} - 3i\partial_X(A_{1,0}|A_{1,0}|^2). \tag{27}$$

At $\mathcal{O}(\varepsilon^{7/2})$ we obtain the equation for $A_{1,1}$, namely

$$\begin{aligned} 2i\omega_0\partial_T A_{1,1} \ = \ & (1-c_g^2)\partial_X^2 A_{1,1} - 6i\partial_X(A_{1,1}|A_{1,0}|^2) - 3i\partial_X((A_{1,0})^2 A_{-1,1}) \\ & + \frac{3}{2}\partial_X^2(A_{1,0}|A_{1,0}|^2). \end{aligned}$$

This is a linearized DNLS equation with some inhomogeneity. All $A_{1,j}$ for $j \geq 2$ satisfy linearized DNLS equations with inhomogeneities, too, whose solutions exist in Gevrey spaces as long as the solutions of the DNLS equation (27) do exist.

Equating the coefficients in front of $\mathbf{E}^3$ to zero gives the determining approximation equation for $A_{3,0}$ at $\mathcal{O}(\varepsilon^{3/2})$, namely

$$-9\omega_0^2 A_{3,0} = -9k_0^2 A_{3,0} - A_{3,0} + \varrho(3i)(A_{1,0})^3$$

which can be solved w.r.t. $A_{3,0}$ since $9\omega_0^2 - 9k_0^2 - 1 \neq 0$. All $A_{n,j}$ with $|n| \geq 3$ satisfy algebraic equations which are linear in the $A_{n,j}$ with a non-vanishing coefficient in front. Again the solutions exist in Gevrey spaces as long as the solutions of the DNLS equation (27) do exist.

The set of equations is structured in such a way that they can be solved one after the other. Therefore, more and more terms cancel in the residual and so the residual Res_u can be made of order $\mathcal{O}(\varepsilon^{\widetilde{\beta}+2})$. Since the residual term $\varepsilon^{-\widetilde{\beta}}\mathrm{Res}(\varepsilon^{1/2}\psi)$ in the equations for the error (9) arises as $\mathcal{O}(1)$-bounded transformation from Res_u the same is true for $\mathrm{Res}(\varepsilon^{1/2}\psi)$ in (9), i.e.,

$$\sup_{t\in[0,T_1/\varepsilon^2]} \|\varepsilon^{-\widetilde{\beta}}\mathrm{Res}(\varepsilon^{1/2}\psi)\|_{\mathcal{M}^s_{\beta(t)}} \leq C\varepsilon^2$$

in (9), cf. Corollary 7.3. Moreover, we have

$$\sup_{t\in[0,T_1/\varepsilon^2]} \|\varepsilon^{1/2}\psi\|_{\mathcal{W}^s_{\beta(t)}} \leq C\varepsilon^{1/2}.$$

More details can be added following the existing literature about the validity of the NLS approximation, cf. [SU17] for an overview.

References

[AS81] Mark J. Ablowitz and Harvey Segur. *Solitons and the inverse scattering transform*, volume 4. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1981.

[BKS20] Thomas J. Bridges, Anna Kostianko, and Guido Schneider. A proof of validity for multiphase Whitham modulation theory. *Proc. R. Soc. Lond., A, Math. Phys. Eng. Sci.*, 476(2243):19, 2020. Id/No 20200203.

[CKS+02] James E. Colliander, Markus Keel, Gigliola Staffilani, Hideo Takaoka, and Terence C. Tao. A refined global well-posedness result for Schrödinger equations with derivative. *SIAM J. Math. Anal.*, 34(1):64–86, 2002.

[DH18] Wolf-Patrick Duell and Max Heß. Existence of long time solutions and validity of the nonlinear Schrödinger approximation for a quasilinear dispersive equation. *J. Differ. Equations*, 264(4):2598–2632, 2018.

[DHSZ16] Wolf-Patrick Duell, Alina Hermann, Guido Schneider, and Dominik Zimmermann. Justification of the 2D NLS equation for a fourth order nonlinear wave equation - quadratic resonances do not matter much in case of analytic initial conditions. *J. Math. Anal. Appl.*, 436(2):847–867, 2016.

[DSW16] Wolf-Patrick Duell, Guido Schneider, and C. Eugene Wayne. Justification of the nonlinear Schrödinger equation for the evolution of gravity driven 2D surface water waves in a canal of finite depth. *Arch. Ration. Mech. Anal.*, 220(2):543–602, 2016.

[Due17] Wolf-Patrick Duell. Justification of the nonlinear Schrödinger approximation for a quasilinear Klein-Gordon equation. *Commun. Math. Phys.*, 355(3):1189–1207, 2017.

[Due21] Wolf-Patrick Duell. Validity of the nonlinear Schrödinger approximation for the two-dimensional water wave problem with and without surface tension in the arc length formulation. *Arch. Ration. Mech. Anal.*, 239(2):831–914, 2021.

[HO92] Nakao Hayashi and Tohru Ozawa. On the derivative nonlinear Schrödinger equation. *Physica D*, 55(1-2):14–36, 1992.

[HS20] Tobias Haas and Guido Schneider. Failure of the N-wave interaction approximation without imposing periodic boundary conditions. *ZAMM Z. Angew. Math. Mech.*, 100(6):e201900230, 16, 2020.

[IT19] Mihaela Ifrim and Daniel Tataru. The NLS approximation for two dimensional deep gravity waves. *Sci. China, Math.*, 62(6):1101–1120, 2019.

[JLPS20] Robert Jenkins, Jiaqi Liu, Peter Perry, and Catherine Sulem. The derivative nonlinear Schrödinger equation: global well-posedness and soliton resolution. *Q. Appl. Math.*, 78(1):33–73, 2020.

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

[Kat75] Tosio Kato. The Cauchy problem for quasi-linear symmetric hyperbolic systems. *Arch. Ration. Mech. Anal.*, 58:181–205, 1975.

[Kat04] Yitzhak Katznelson. *An introduction to harmonic analysis.* Cambridge: Cambridge University Press, 2004.

[KN78] David J. Kaup and Alan C. Newell. An exact solution for a derivative nonlinear Schrödinger equation. *J. Math. Phys.*, 19:798–801, 1978.

[KN86] Tadayoshi Kano and Takaaki Nishida. A mathematical justification for Korteweg-de Vries equation and Boussinesq equation of water surface waves. *Osaka J. Math.*, 23:389–413, 1986.

[KSM92] Pius Kirrmann, Guido Schneider, and Alexander Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. R. Soc. Edinb., Sect. A, Math.*, 122(1-2):85–91, 1992.

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

[Sch96] Guido Schneider. Limits for the Korteweg-de Vries-approximation. *Z. Angew. Math. Mech.*, 76:341–344, 1996.

[Sch98] Guido Schneider. Justification of modulation equations for hyperbolic systems via normal forms. *NoDEA, Nonlinear Differ. Equ. Appl.*, 5(1):69–82, 1998.

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

[SSZ15] Guido Schneider, Danish Ali Sunny, and Dominik Zimmermann. The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions. *J. Dyn. Differ. Equations*, 27(3-4):1077–1099, 2015.

[SU17] Guido Schneider and Hannes Uecker. *Nonlinear PDEs. A dynamical systems approach*, volume 182. Providence, RI: American Mathematical Society (AMS), 2017.

[Tak99] Hideo Takaoka. Well-posedness for the one-dimensional nonlinear Schrödinger equation with the derivative nonlinearity. *Adv. Differ. Equ.*, 4(4):561–580, 1999.

[TF80] Masayoshi Tsutsumi and Isamu Fukuda. On solutions of the derivative nonlinear Schrödinger equation. Existence and uniqueness theorem. *Funkc. Ekvacioj, Ser. Int.*, 23:259–277, 1980.

[TW12] Nathan Totz and Sijue Wu. A rigorous justification of the modulation approximation to the 2D full water wave problem. *Commun. Math. Phys.*, 310(3):817–883, 2012.

[WC17] C. Eugene Wayne and Patrick Cummings. Modified energy functionals and the NLS approximation. *Discrete Contin. Dyn. Syst.*, 37(3):1295–1321, 2017.

[WG17] Yifei Wu and Zihua Guo. Global well-posedness for the derivative nonlinear Schrödinger equation in $H^{\frac{1}{2}}(\mathbb{R})$. *Discrete Contin. Dyn. Syst.*, 37(1):257–264, 2017.

[Wu15] Yifei Wu. Global well-posedness on the derivative nonlinear Schrödinger equation. *Anal. PDE*, 8(5):1101–1112, 2015.