Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

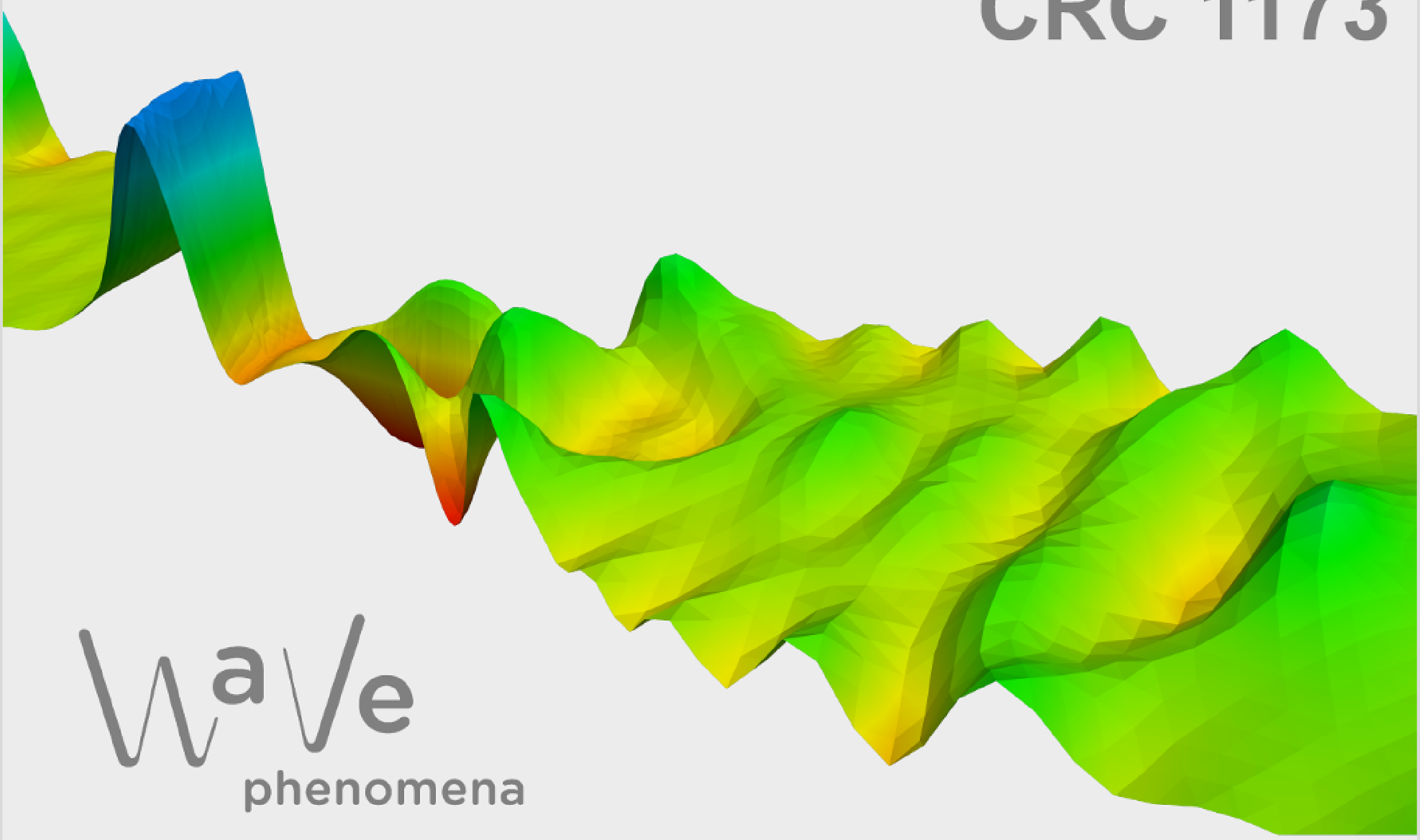
The validity of the Derivative NLS approximation for systems with cubic nonlinearities

Max Heß, Guido Schneider

CRC Preprint 2022/58, November 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

The validity of the Derivative NLS approximation for systems with cubic nonlinearities

Max Heß and Guido Schneider

Institut für Analysis, Dynamik und Modellierung,

Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

November 14, 2022

Abstract

The (generalized) Derivative Nonlinear Schrödinger (DNLS) equation can be derived as an envelope equation via multiple scaling perturbation analysis from dispersive wave systems. It occurs when the cubic coefficient for the associated NLS equation vanishes for the spatial wave number of the underlying slowly modulated wave packet. It is the purpose of this paper to prove that the DNLS equation makes correct predictions about the dynamics of a Klein-Gordon model with a cubic nonlinearity. The proof is based on energy estimates and normal form transformations. New difficulties occur due to a total resonance and due to a second order resonance.

1 Introduction

The (generalized) Derivative Nonlinear Schrödinger (DNLS) equation

$$i\partial_T A = \nu_1 \partial_X^2 A + \nu_2 A|A|^2 + i\nu_3 |A|^2 \partial_X A + i\nu_4 A^2 \partial_X \overline{A} + \nu_5 A|A|^4, \qquad (1)$$

with $T \geq 0$, $X \in \mathbb{R}$, $A(X,T) \in \mathbb{C}$, and coefficients $\nu_j \in \mathbb{R}$ for $j = 1, \ldots, 5$, is a canonical dispersive equation that can be obtained in a number of weakly nonlinear scaling regimes. It has been derived for instance as a long wave limit equation from the 1D compressible MHD equations in the presence of the Hall effect, cf. [Mj76, CLPS99, JLPS20]. Here, we are interested in its appearance as envelope equation describing slow modulations in time and space of an oscillating wave packet $e^{i(k_0 x - \omega_0 t)}$. It takes the role of the NLS equation if the cubic coefficient of the NLS equation vanishes for the chosen wave number k_0 of the underlying slowly modulated wave packet, cf. Remark 1.3. Hence the DNLS equation describes a degenerated situation. This situation appears for instance in the water wave problem, cf. [AS81], for a curve in the two-dimensional parameter plane of surface tension and basic wave number k_0.

It is the purpose of this paper to answer the question whether the DNLS approximation makes correct predictions about the dynamics of the original dispersive system for which the DNLS equation has been derived as an envelope equation in this sense. To our knowledge no such approximation result is documented in the existing literature for initial conditions in Sobolev spaces. As a first step in this direction we consider a simple model equation as original system, namely a nonlinear Klein Gordon equation with a cubic nonlinearity

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

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

$$\varrho(ik) = \frac{k^2 - 1}{k^2 + 1}, \qquad \text{resp.} \qquad \varrho(\partial_x) = -(1 - \partial_x^2)^{-1}(1 + \partial_x^2). \tag{3}$$

In order to derive the DNLS equation we make the ansatz

$$u(x,t) \approx \varepsilon^{1/2} \psi_A(x,t) = \varepsilon^{1/2} A(\varepsilon(x - c_g t), \varepsilon^2 t) e^{i(k_0 x - \omega_0 t)} + c.c., \tag{4}$$

where $c.c.$ denotes the complex conjugate, c_g the linear group velocity, $k_0 = 1$ the basic spatial wave number, and ω_0 the basic temporal wave number. Moreover, $0 < \varepsilon \ll 1$ is a small perturbation parameter. See Figure 1.

Substitution of this ansatz into (2) and equating the coefficients in front of $e^{i(k_0 x - \omega_0 t)}$ to zero gives the linear dispersion relation $\omega_0^2 = k_0^2 + 1$ at $\mathcal{O}(\varepsilon^{1/2})$ and the linear group velocity $c_g = \omega'(k_0) = k_0/\omega_0$ at $\mathcal{O}(\varepsilon^{3/2})$. Using the expansion

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

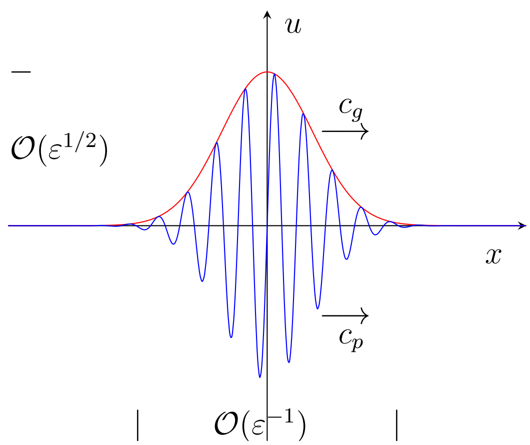


Figure 1: The envelope (advancing with the linear group velocity c_g) of the oscillating wave packet (advancing with the phase velocity $c_p = \omega_0/k_0$) is described by the amplitude A which solves the DNLS equation (1).

yields

$$\varrho(\partial_x)(3\varepsilon^{3/2}A|A|^2 e^{i(x-\omega_0 t)}) = -3i\varepsilon^{5/2}(\partial_X(A|A|^2))e^{i(x-\omega_0 t)} + \mathcal{O}(\varepsilon^{7/2}),$$

and so at $\mathcal{O}(\varepsilon^{5/2})$ we find the DNLS equation

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

which is a special case of the (generalized) DNLS equation (1). In particular, there is no $A|A|^4$ term in the DNLS equation due to our special choice of nonlinearity in (2). It is the goal of this paper to prove that the DNLS equation (5) makes correct predictions about the dynamics of our Klein-Gordon model (2), i.e., to prove that the following approximation property holds.

Theorem 1.1. *Let $s_A \geq 12$ and $A \in C([0,T_0],H^{s_A})$ be a solution of the DNLS equation* (5). *Then there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ we have solutions u of the Klein-Gordon model* (2) *such that*

$$\sup_{t\in[0,T_0/\varepsilon^2]}\sup_{x\in\mathbb{R}}\left|u(x,t) - \left(\varepsilon^{1/2}A(\varepsilon(x-c_g t),\varepsilon^2 t)e^{i(k_0 x-\omega_0 t)} + c.c.\right)\right| \leq C\varepsilon^{3/2}.$$

Remark 1.2. The proof of Theorem 1.1 is a nontrivial task since solutions of order $\mathcal{O}(\varepsilon^{1/2})$ have to be estimated on an $\mathcal{O}(\varepsilon^{-2})$-time scale. Since we have a cubic nonlinearity a simple application of Gronwall's inequality would only give estimates on an $\mathcal{O}(\varepsilon^{-1})$-time scale. In fact, such approximation results should not be taken for granted. There are a number of counter examples where formally derived envelope equations make wrong predictions about the dynamics of the original systems, cf. [Sch95, SSZ15, HS20].

Remark 1.3. The classical Nonlinear Schrödinger (NLS) equation

$$i\partial_T \tilde{A} = \nu_1 \partial_X^2 \tilde{A} + \widetilde{\nu}_2 \tilde{A}|\tilde{A}|^2, \tag{6}$$

with coefficients $\nu_1, \widetilde{\nu}_2 \in \mathbb{R}$, can be derived for the description of dispersive wave systems, such as the cubic Klein-Gordon equation

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

with $x, t, u(x,t) \in \mathbb{R}$, the water wave problem, or systems from nonlinear optics. The ansatz for the derivation of the NLS equation is of the form

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

where c_g is the linear group velocity, k_0 the basic spatial wave number, ω_0 the basic temporal wave number, and where again $0 < \varepsilon \ll 1$ is a small perturbation parameter. Various NLS approximation results have been established in the last decades, cf. [Kal88, KSM92, TW12, Due21]. The DNLS equation (1) appears if the cubic coefficient $\widetilde{\nu}_2 = \widetilde{\nu}_2(k_0)$ in (6) vanishes for the chosen wave number k_0.

Remark 1.4. The DNLS equation (5) is completely integrable and can be solved through the inverse scattering method, cf. [KN78, JLPS20]. Local existence and uniqueness of smooth solutions in Sobolev spaces H^s with $s > 3/2$ was established in [TF80]. For extensions of this result to solutions of lower regularity and global existence results see for instance [HO92, Tak99, CKS$^+$02, Wu15, WG17]. For $u_0 \in H^s$, $s < 1/2$, the Cauchy problem is ill-posed in the sense that uniform continuity with respect to the initial conditions fails [Tak99].

Remark 1.5. The basic idea to close the gap between the trivial $\mathcal{O}(\varepsilon^{-1})$- and the natural $\mathcal{O}(\varepsilon^{-2})$-time scale is to use normal form transformations. By these near identity change of variables a number of cubic terms in the nonlinearity can be transformed in $\mathcal{O}(u^5)$-terms. However, for our problem, resonances prevent the elimination of all cubic terms. In the equations for the error R there is a total resonance, see below, but the terms which cannot be eliminated due to this resonance (proportional to $\varepsilon|A|^2R$) can be estimated with simple energy estimates. The more serious difficulty comes from a resonance at the wave numbers $k = \pm k_0 = \pm 1$. This resonance is not linear, but of second order, and so the linear vanishing of the nonlinearity at these wave numbers is not sufficient for the elimination of these terms (proportional to $\varepsilon\varrho(\partial_x)(A^2e^{2i(k_0x-\omega_0t)}R)$). At the end of Section 2 it is outlined how to get rid of this problem.

Remark 1.6. Recently, the validity of the NLS approximation for quasilinear dispersive wave systems has attracted a lot of interest, cf. [WC17, DH18, Due17, Due21, Hes22]. The most simple quasilinear example where a DNLS equation can be derived is a quasilinear Klein Gordon model, given by

$$\partial_t^2 u = (\partial_x^2 - 1)u + (\partial_x^2 + 1)u^3 \,. \tag{7}$$

The difficulty in handling quasilinear dispersive wave systems lies in the fact that normal form transformations lose regularity. However, by a suitably constructed energy these difficulties can be avoided. Therefore, we strongly expect that the analysis made in [Hes22] about the validity of the NLS approximation transfers to (7) and the validity of the DNLS approximation. This will be the topic of future research.

Remark 1.7. Why is (2) a good model for starting a validity theory for the DNLS approximation? 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. Quadratic terms which are not present in our model can be completely eliminated by normal form transformations in Klein-Gordon models. Therefore, we do not expect new difficulties coming from these terms. We expect that the transfer of the present result to the water wave problem, cf. [AS81], will be a combination of the approach presented in this paper with the ones about the NLS approximation for the water wave problem, cf. [TW12, DSW16, Due21]. New difficulties will be a more complicated resonance structure and the quasilinearity of the water wave problem.

Remark 1.8. Recently, in [HS22] we established a DNLS approximation result for initial conditions of the DNLS equation in Gevrey spaces with a completely different method. The proof of [HS22] is based on a Cauchy-Kowalevskaya like approach in so called modulational Sobolev spaces. This approach allows us to weaken the non-resonance condition but has the disadvantages that analytic initial conditions are necessary and that the approximation time is T_1/ε^2 with T_1 possibly smaller than T_0.

Notation. The Fourier transform of a function $u \in L^2(\mathbb{R},\mathbb{K})$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ is denoted by $\mathcal{F}(u)(k) = \widehat{u}(k) = \frac{1}{2\pi}\int_{\mathbb{R}} u(x)e^{-ikx}dx$. $H^s(\mathbb{R},\mathbb{K})$ is the space of functions from $\mathbb{R}$ into $\mathbb{K}$, for which the norm $\|u\|_{H^s(\mathbb{R},\mathbb{K})} = (\int_{\mathbb{R}} |\widehat{u}(k)|^2(1+|k|^2)^s \mathrm{d}k)^{1/2}$ is finite. The space $L^p_s(\mathbb{R},\mathbb{K})$ is defined by $u \in L^p_s(\mathbb{R},\mathbb{K}) \Leftrightarrow u\sigma^s \in L^p(\mathbb{R},\mathbb{K})$, where $\sigma(x) = (1+x^2)^{1/2}$. Many, possibly different, constants are denoted by the same symbol C if they can be independently chosen 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 Outline of the proof

It is the purpose of this section to present the underlying ideas of the proof of Theorem 1.1. Our model problem is of the form

$$\partial_t^2 u = -\omega_{op}^2 u - \omega_{op}\rho_{op} u^3 \,, \tag{8}$$

with pseudo differential operators ω_{op} and ρ_{op} which are defined below. We write (8) as a first order system

$$\begin{aligned} \partial_t u &= -\, i\omega_{op} v \,, \\ \partial_t v &= -\, i\omega_{op} u - i\rho_{op} u^3 \,. \end{aligned} \tag{9}$$

Via the invertible transformation

$$V = \begin{pmatrix} v_{-1} \\ v_1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} \tag{10}$$

one obtains the first order system

$$\partial_t V = \Lambda V + N(V,V,V), \tag{11}$$

where ΛV stands for the linear terms for which the associated linear operator $\widehat{\Lambda}$ is now in diagonal form in Fourier space. Moreover, $N(V,V,V)$ stands for the nonlinear terms which can be written as a symmetric trilinear mapping. In Fourier space this system is written as

$$\partial_t \widehat{V} = \widehat{\Lambda}\widehat{V} + \widehat{N}(\widehat{V},\widehat{V},\widehat{V}), \tag{12}$$

where

$$\begin{aligned}\widehat{\Lambda}(k) &= \begin{pmatrix} -i\omega(k) & 0 \\ 0 & i\omega(k) \end{pmatrix}, \\ \widehat{N}(\widehat{V},\widehat{V},\widehat{V})(k,t) &= \frac{1}{2} i\rho(k) \begin{pmatrix} -(\widehat{v}_1 + \widehat{v}_{-1})^{*3} \\ (\widehat{v}_1 + \widehat{v}_{-1})^{*3} \end{pmatrix}(k,t),\end{aligned}$$

with

$$\omega(k) = \operatorname{sign}(k)\sqrt{k^2+1} \qquad \text{and} \qquad \rho(k) = -\frac{\varrho(ik)}{\omega(k)},$$

where ϱ has been defined in (3) and where $\cdot^{*3}$ stands for the two times convolution. By this definition of ω and ρ the components of the vector V will be real-valued in physical space.

In order to estimate the error $\varepsilon^{\beta} R$ that is made by a DNLS-approximation $\varepsilon^{1/2}\Psi$ we write a solution V of (11) as sum of the approximation and the error, i.e.,

$$V = \varepsilon^{1/2}\Psi + \varepsilon^{\beta} R \tag{13}$$

for a $\beta > 3/2$. If R can be uniformly bounded independently of ε, our theorem would follow. We find that the error function R satisfies

$$\partial_t R = \Lambda R + 3\varepsilon N(\Psi,\Psi,R) + \mathcal{O}(\varepsilon^2). \tag{14}$$

Hence, since Λ is a skew-symmetric operator, it remains to get rid of the $3\varepsilon N(\Psi,\Psi,R)$-term in order to uniformly bound the error function R on the long $\mathcal{O}(\varepsilon^{-2})$-time scale. We follow the existing literature about the justification of envelope equations and try to eliminate this term by a near identity change of variables, i.e.,

$$W = R + \varepsilon M(\Psi,\Psi,R),$$

with M a suitably chosen trilinear mapping. An elimination is only possible if a non-resonance condition is satisfied. In order to see which terms can be eliminated and which not, we split the approximation into a part $\varepsilon^{1/2}a_1$ which is concentrated at the wave number $k = k_0 = 1$ and into a part $\varepsilon^{1/2}a_{-1}$ which is concentrated at the wave number $k = -k_0 = -1$. Due to our definition of ω the DNLS approximation is of the form

$$\varepsilon^{1/2}\Psi = \begin{pmatrix} \varepsilon^{1/2}a_1 + \varepsilon^{1/2}a_{-1} \\ 0 \end{pmatrix} + \mathcal{O}(\varepsilon^{3/2}),$$

cf. Section 3 for more details. For the subsequent explanations we concentrate on the first coordinate $\widehat{R}_{-1}$ of the error function. We obtain a subsystem of the form

$$\partial_t \widehat{R}_{-1}(k,t) = -i\omega \widehat{R}_{-1}(k,t) + \varepsilon \sum_{j_1,j_2=\pm 1} \rho_{j_1,j_2,-1} \cdot (\widehat{a}_{j_1} * \widehat{a}_{j_2} * \widehat{R}_{-1})(k,t) + \ldots,$$

where $*$ denotes convolution, and where the $\rho_{j_1,j_2,j_3} = \rho_{j_1,j_2,j_3}(k)$ are smooth functions vanishing linearly at $k = \pm 1$. The non-resonance condition to eliminate the lowest order part of a term $\varepsilon\rho_{j_1,j_2,j_3} \cdot (\widehat{a}_{j_1} * \widehat{a}_{j_2} * \widehat{R}_{j_3})(k,t)$ can be reduced to

$$\omega(k) - \omega(j_1) - \omega(j_2) + j_3\omega(k - j_1 - j_2) \neq 0,$$

using the fact that $\widehat{a}_j$ is strongly concentrated in an $\mathcal{O}(\varepsilon)$-neighborhood of the wave number $k = j$. See below or [SU17, §11.5]. The left-hand side of this non-resonance condition appears in the denominator of the normal form transformation.

Like for the Hopf bifurcation cubic terms cannot be eliminated if $j_1 = -j_2$ (and here $j_3 = -1$ since $j_3 = 1$ is not resonant). For this choice the left-hand side of the non-resonance condition vanishes identically. Hence a term of the form

$$\varepsilon\rho_{1,-1.-1} \cdot (\widehat{a}_1 * \widehat{a}_{-1} * \widehat{R}_{-1})(k,t)$$

cannot be eliminated at all. See also the fifth picture of Figure 2 and the third picture of Figure 3. Fortunately, these remaining totally resonant terms can be estimated by energy estimates.

There is another resonance at the wave numbers $k = \pm k_0 = \pm 1$ which makes more trouble. For $j_1 = j_2 =: j = \pm 1$ and $j_3 = 1$ we have

$$\omega(j) - \omega(j) - \omega(j) + \omega(-j) = 0.$$

The fact that the nonlinear terms vanish at this wave number, too, leads to a zero in the nominator of the normal form transformation. One could have the hope that the two zeroes cancel but in fact the left-hand side of the non-resonance condition vanishes quadratically. We have

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

See also the second and sixth picture of Figure 2. Therefore, an elimination of

$$\varepsilon\rho_{j,j,-1} \cdot (\widehat{a}_j * \widehat{a}_j * \widehat{R}_{-1})(k,t)$$

near the wave number $k = j$ is not possible with an $\mathcal{O}(\varepsilon)$-perturbation of the identity.

However, by adding and subtracting irrelevant terms of order $\kappa\mathcal{O}(\varepsilon^2)$ to and from the equations for the error, the quadratic singularity in the denominator of the normal form transformation can be shifted away from zero which turns out be sufficient for an elimination of these terms. In the end we have a kernel in the normal form transformation proportional to $\varepsilon ik/(\kappa\varepsilon^2 + k^2)$ which, however, is only $O(1)$ and not $\mathcal{O}(\varepsilon)$. Nevertheless, by choosing $\kappa = O(1)$ sufficiently large this normal form transformation is still invertible. The drawback of the fact that this perturbation of identity is of order $\mathcal{O}(1)$ is that in the equations for the error new terms of order $\mathcal{O}(\varepsilon)$ are created. However, in the energy estimates these new terms have a kernel proportional to $(\varepsilon ik)^2/(\kappa\varepsilon^2 + k^2)$ which is of order $\mathcal{O}(\varepsilon^2)$. And so finally in the equations for the error all terms can be controlled by energy estimates or can be brought to order $\mathcal{O}(\varepsilon^2)$ by a normal form transformation.

In addition to this plan, we construct a higher order DNLS approximation and estimate the remaining residual terms in Section 3. This is necessary for choosing $\beta > 3/2$. In Section 4, after analyzing the occurring resonances we perform the normal form transformation for eliminating the non-resonant terms. In Section 5 the final energy estimates will be provided.

3 The higher order DNLS approximation

The residual

$$\mathrm{Res}_u(u) = -\partial_t^2 u + \partial_x^2 u - u - (1-\partial_x^2)^{-1}(1+\partial_x^2)u^3$$

contains all terms which do not cancel after inserting a DNLS approximation $u = \varepsilon^{1/2}\Psi$ into the equation (2). Hence, the smaller the residual, the better the approximation $\varepsilon^{1/2}\Psi$ can be. In order to have a residual which is sufficiently small for our purposes, we construct an improved approximation $\varepsilon^{1/2}\Psi$ which differs from our original DNLS approximation

$$\varepsilon^{1/2}\psi_A(x,t) = \varepsilon^{1/2}A(\varepsilon(x-c_gt),\varepsilon^2t)e^{i(k_0x-\omega_0t)} + c.c.,$$

by a number of higher order terms. With $\mathbf{E} = e^{i(k_0x-\omega_0t)}$ we find that

$$\begin{aligned} \mathrm{Res}_u(\varepsilon^{1/2}\psi_A) \;=\;& \varepsilon^{1/2}(\omega_0^2-k_0^2-1)A\mathbf{E} + c.c. \\ &+\varepsilon^{3/2}(2i\omega_0-2ick_0)(\partial_XA)\mathbf{E} + c.c. \\ &+\varepsilon^{5/2}(-2i\omega_0\partial_TA + (1-c_g^2)\partial_X^2A - 3i\partial_X(A|A|^2))\mathbf{E} + c.c. \\ &+\varepsilon^{3/2}\varrho(3i)A^3\mathbf{E}^3 + c.c. - \varepsilon^{5/2}(\partial_T^2A)\mathbf{E} + c.c. + \mathcal{O}(\varepsilon^{7/2}), \end{aligned}$$

where the $\mathcal{O}(\varepsilon^{7/2})$-terms come from a further expansion of ϱ. The first line in the residual cancels due to the dispersion relation. The second line in the residual cancels due to our choice of the linear group velocity c_g. The third line cancels since A is chosen to be a solution of the DNLS equation (5).

For the subsequent proof we can allow for residual terms of order $\mathcal{O}(\varepsilon^{7/2+\delta_1})$ for a $\delta_1 > 0$ in some H^s-space. In order to achieve such an estimate we have to get rid of the lower order terms in the fourth line of the above residual. In detail, by adding higher order terms to the original DNLS approximation $\varepsilon^{1/2}\psi_A$ we construct an improved DNLS approximation $\varepsilon^{1/2}\Psi$ for (11), such that the residual is of formal order $\mathcal{O}(\varepsilon^5)$. Due to the scaling of the L^2-norm with respect to the scaling $X = \varepsilon x$ we lose an order $\varepsilon^{-1/2}$ from the formal order such that $\mathcal{O}(\varepsilon^5)$ is the first integer order larger than the formal order $\mathcal{O}(\varepsilon^{4+\delta_1})$.

i) Before we do so, we modify the original DNLS approximation $\varepsilon^{1/2}\psi_A$ by replacing A in the definition of $\varepsilon^{1/2}\psi_A$ by

$$A_c\big(\varepsilon(\cdot-c_gt),\varepsilon^2t\big) = \mathcal{F}^{-1}\Big[\chi_{[-\delta,\delta]}(\cdot)\mathcal{F}\big[A\big(\varepsilon(\cdot-c_gt),\varepsilon^2t\big)\big](\cdot)\Big],$$

where $\chi_{[-\delta,\delta]}$ is the characteristic function on the interval $[-\delta,\delta]$ and $\delta \in (0,k_0/20)$ is a fixed chosen constant that is independent of $0 < \varepsilon \ll 1$. Due to the estimate

$$\|\chi_{[-\delta,\delta]}\varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot) - \varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot)\|_{L^2_m} \leq C(\delta)\,\varepsilon^{m+M-1/2}\|f\|_{H^{m+M}}, \tag{15}$$

cf. [SU17, §11.5], and the fact that $A \in C([0,T_0],H^{s_A}(\mathbb{R},\mathbb{C}))$ solves the DNLS equation (5), we have that the error made by replacing A by A_c, remains small, i.e.,

$$\varepsilon^{5/2}\big\|-2i\omega_0\partial_TA_c + (1-c_g^2)\partial_X^2A_c - 3i\partial_X(A_c|A_c|^2)\big\|_{L^2} = \mathcal{O}(\varepsilon^5).$$

Since A_c only has a bounded support in Fourier space, the use of A_c instead of A subsequently allows us to control the terms in the normal form transformations more efficiently.

ii) The higher order DNLS approximation that we use for (11) is given by

$$\varepsilon^{1/2}\Psi = \varepsilon^{1/2}(a_1 + a_{-1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \varepsilon^{3/2}\Psi_q \,, \tag{16}$$

where

$$a_1(x,t) = \; A_c(\varepsilon(x-c_g t), \varepsilon^2 t)\, \mathbf{E} \,, \quad a_{-1}(x,t) = \overline{A_c(\varepsilon(x-c_g t), \varepsilon^2 t)}\, \mathbf{E}^{-1} \,,$$

and

$$\begin{aligned} \varepsilon^{3/2}\Psi_q(x,t) \;\; = \;\; & \sum_{n=1,2,3,4} \varepsilon^{1/2+n} \begin{pmatrix} A_{1,n}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E} + c.c. \\ B_{1,n}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E} + c.c. \end{pmatrix} \\ & + \sum_{n=0,1,2} \varepsilon^{5/2+n} \begin{pmatrix} A_{3,n}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E}^3 + c.c. \\ B_{3,n}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E}^3 + c.c. \end{pmatrix} \\ & + \varepsilon^{9/2} \begin{pmatrix} A_{5,0}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E}^5 + c.c. \\ B_{5,0}(\varepsilon(x-c_g t), \varepsilon^2 t)\mathbf{E}^5 + c.c. \end{pmatrix}, \end{aligned}$$

with $\mathbf{E}, \omega_0$ and c_g as before. Since $A_c\mathbf{E}$ has a small support in Fourier space near the wave number k_0 and since we have a polynomial nonlinearity, the A_l^n and B_l^n in the end can be chosen such that the support of $A_l^n\mathbf{E}^l$ and $B_l^n\mathbf{E}^l$ in Fourier space lies in a small neighborhood of the wave number lk_0.

Equating the coefficients at $\varepsilon^{5/2}\mathbf{E}$ in the first component of the residual $\mathrm{Res}_u(\varepsilon^{1/2}\Psi)$ to zero gives the DNLS equation for A_c, respectively A.

Equating the coefficients at $\varepsilon^{5/2+n}\mathbf{E}$ or $\varepsilon^{5/2+n}\mathbf{E}^{-1}$ in the first component to zero gives that the $A_{1,n}$ and $A_{-1,n}$ are determined by solving linear, but inhomogeneous, Schrödinger equations, in which the inhomogeneous terms in the end only depend on A_c.

Equating the coefficients at $\varepsilon^{1/2+n}\mathbf{E}$ or $\varepsilon^{1/2+n}\mathbf{E}^{-1}$ in the second component to zero gives that the $B_{1,n}$ and $B_{-1,n}$ are determined by solving linear, but inhomogeneous, algebraic equations, in which the inhomogeneous terms in the end only depend on A_c.

Equating the coefficients in front of $\varepsilon^{3/2+n}$ at the other $\mathbf{E}^j$ with $j \notin \{-1, 1\}$ to zero gives again linear algebraic equations for the $A_{j,n}$ and $B_{j,n}$

which can be explicitly solved with respect to $A_{j,n}$ and $B_{j,n}$ since the coefficients in front of the $A_{j,n}$ and $B_{j,n}$ do not vanish, i.e., because of

$$j\omega(k_0) \pm \omega(jk_0) \neq 0 \qquad\qquad \text{for } j \in \{\pm 3, \pm 5\}\,.$$

For more details we refer to literature about the classical NLS approximation.

iii) The properties of $\varepsilon^{1/2}\Psi$ which we need for the proof of the error estimates are summarized in the the following lemma.

Lemma 3.1. *Let $s_A \geq 12$ and $A \in C([0,T_0], H^{s_A}(\mathbb{R},\mathbb{C}))$ be a solution of the DNLS equation* (5) *with*

$$\sup_{T\in[0,T_0]} \|A\|_{H^{s_A}} \leq C_A.$$

Then for all $s \geq 0$ there exist $C_{\text{Res}}, C_\Psi, \varepsilon_0 > 0$ depending on C_A such that for all $\varepsilon \in (0,\varepsilon_0)$ the approximation $\varepsilon^{1/2}\Psi$ defined in (16) *satisfies*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\text{Res}_u(\varepsilon^{1/2}\Psi)\|_{H^s} \;\leq\; C_{\text{Res}}\,\varepsilon^5, \tag{17}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} \Big\|\varepsilon^{1/2}\Psi - \varepsilon^{1/2}\psi_A \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Big\|_{C_b^0} \;\leq\; C_\Psi\,\varepsilon^{3/2}, \tag{18}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} (\|\widehat{a}_1\|_{L^1_{s+1}(\mathbb{R},\mathbb{C})} + \|\widehat{a}_{-1}\|_{L^1_{s+1}(\mathbb{R},\mathbb{C})} + \|\widehat{\Psi}_q\|_{L^1_{s+1}(\mathbb{R},\mathbb{C})}) \;\leq\; C_\Psi\,. \tag{19}$$

Proof. Since the proofs of such estimates are documented in the existing literature about the NLS approximation, cf. [SU17, §11], a number of times, we refer to these papers and refrain from recalling the proof. □

Remark 3.2. a) The bound (19) allows us to estimate

$$\|a_1 f\|_{H^s} \leq C\|a_1\|_{C_b^s}\|f\|_{H^s} \leq C\|\widehat{a}_1\|_{L^1_s}\|f\|_{H^s},$$

without loss of powers in ε as it would be the case if in this estimate $\|a_1\|_{C_b^s}$ would be replaced by $\|a_1\|_{H^s} = \|\widehat{a}_1\|_{L^2_s}$, due to the scaling properties of the L^2-norm, namely $\|A(\varepsilon\cdot)\|_{L^2} = \varepsilon^{-1/2}\|A(\cdot)\|_{L^2}$.

b) Our construction of $\varepsilon\Psi$ with a bounded support in Fourier space has the additional consequence that the estimates (17) and (19) are true for all $s \geq 0$. First all estimates are shown in L^2. Since all appearing terms have a bounded support in Fourier space, we have the equivalence of the L^2-norm and each H^s-norm for these finitely many terms.

c) The necessary regularity $s_A \geq 12$ comes the fact that the equation for $A_{1,4}$ is solved in L^2 with an inhomogeneity containing $\partial_X^3 A_{1,3}$, that the equation for $A_{1,3}$ is solved in H^3 with an inhomogeneity containing $\partial_X^3 A_{1,2}$, etc. This can be reduced to $s_A \geq 6$ since for our purposes only A_1, $A_{1,1}$, and $A_{1,2}$ are necessary. The regularity can be reduced further by arguments which can be found in [SU17, §11]. However, since this is not the main goal of this paper we refrain from optimizing the value of s_A.

iv) In the following we prove that the improved approximation $\varepsilon\Psi$ makes correct predictions about the dynamics of the original system, i.e., error estimates are established for the improved approximation $\varepsilon\Psi$. The bound (18) and the triangle inequality then imply that these error estimates hold for the original DNLS approximation $\varepsilon\psi_A$, too. Hence, Theorem 1.1 is a direct consequence of the subsequent approximation result.

Theorem 3.3. *Let $\beta \in (3/2, 5/2)$. For $s_A \geq 12$ and all $C_1, T_0 > 0$ there exist $\varepsilon_0, C_2 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ of the DNLS equation* (5) *with*

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

the following holds. For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions

$$V \in C\big([0, T_0/\varepsilon^2], \big(H^1(\mathbb{R}, \mathbb{R})\big)^2\big)$$

of the diagonalized first order system (11) *which satisfy*

$$\sup_{t \in [0,T_0/\varepsilon^2]} \|V(\cdot, t) - \varepsilon^{1/2}\Psi(\cdot, t)\|_{H^1} \leq C_2 \varepsilon^{\beta}.$$

Remark 3.4. The parameter β could be chosen arbitrarily large by adding more and more higher order terms to the improved approximation $\varepsilon^{1/2}\Psi$ and by choosing then s_A sufficiently large.

Proof of Theorem 1.1: Sobolev's embedding theorem $H^1 \subset C_b^0$ and estimate (18) yield

$$
\begin{aligned}
& \sup_{t\in[0,T_0/\varepsilon^2]} \sup_{x\in\mathbb{R}} \left| u - \left(\varepsilon^{1/2} A(\varepsilon(x-c_g t), \varepsilon^2 t) e^{i(k_0 x - \omega_0 t)} + c.c.\right) \right| \\
\leq \quad & \sup_{t\in[0,T_0/\varepsilon^2]} \left\| V - \varepsilon^{1/2} \begin{pmatrix} \psi_A \\ 0 \end{pmatrix} \right\|_{C_b^0} \\
\leq \quad & \sup_{t\in[0,T_0/\varepsilon^2]} \left\| V - \varepsilon^{1/2} \Psi \right\|_{C_b^0} + \sup_{t\in[0,T_0/\varepsilon^2]} \left\| \varepsilon^{1/2}\Psi - \varepsilon^{1/2} \begin{pmatrix} \psi_A \\ 0 \end{pmatrix} \right\|_{C_b^0} \\
\leq \quad & \mathcal{O}(\varepsilon^{3/2}).
\end{aligned}
$$

Thus, Theorem 1.1 is a direct consequence of Theorem 3.3. □

4 The normal form transformations

In this section we compute the normal form transformations which would be necessary to eliminate various cubic terms. Since the totally resonant terms will be estimated with energy estimates, in the end we will not apply the normal form transformations but include them subsequently in our energy.

In order to prove Theorem 3.3, we follow the outline presented in Section 2. The error

$$\varepsilon^{\beta} R = V - \varepsilon^{1/2} \Psi \,, \tag{20}$$

satisfies

$$\partial_t R = \Lambda R + \varepsilon L_c(R) + \varepsilon^2 L_s(R) + \varepsilon^{\beta+1/2} L_r(R) + \varepsilon^{-\beta} \mathrm{Res}(\varepsilon^{1/2}\Psi) \tag{21}$$

where

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

$$\widehat{L_s(R)}(k,t) = 3i\rho(k) \begin{pmatrix} -(\widehat{a}_1+\widehat{a}_{-1}) * (\widehat{\Psi}_{q,1}+\widehat{\Psi}_{q,-1}) * (\widehat{R}_1+\widehat{R}_{-1}) \\ (\widehat{a}_1+\widehat{a}_{-1}) * (\widehat{\Psi}_{q,1}+\widehat{\Psi}_{q,-1}) * (\widehat{R}_1+\widehat{R}_{-1}) \end{pmatrix}(k,t)$$
$$+\varepsilon \frac{3}{2} i\rho(k) \begin{pmatrix} -(\widehat{\Psi}_{q,1}+\widehat{\Psi}_{q,-1})^{*2} * (\widehat{R}_1+\widehat{R}_{-1}) \\ (\widehat{\Psi}_{q,1}+\widehat{\Psi}_{q,-1})^{*2} * (\widehat{R}_1+\widehat{R}_{-1}) \end{pmatrix}(k,t),$$

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

Herein, $\widehat{R} = (\widehat{R}_{-1}, \widehat{R}_1)$, and

$$\varepsilon^{1/2} \begin{pmatrix} \widehat{\Psi}_{-1} \\ \widehat{\Psi}_1 \end{pmatrix} = \varepsilon^{1/2}(\widehat{a}_1+\widehat{a}_{-1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \varepsilon^{3/2} \begin{pmatrix} \widehat{\Psi}_{q,-1} \\ \widehat{\Psi}_{q,1} \end{pmatrix}.$$

Notation. We remind that the index j in R_j, Ψ_j, and $\Psi_{q,j}$ denotes the j-th coordinate, where in a_j the index j denotes the wave number where this part of the approximation is concentrated.

For $\beta > 3/2$ except of the first two terms all terms on the the right-hand side of (21) are at least of order $\mathcal{O}(\varepsilon^2)$. Since Λ is skew symmetric, the first term ΛR makes no problem in estimating the error on the long $\mathcal{O}(\varepsilon^{-2})$-time scale. However, handling the second term $\varepsilon L_c(R)$ is non-trivial. A simple application of Gronwall's inequality would only give estimates on an $\mathcal{O}(\varepsilon^{-1})$-time scale. The idea to close the gap between the trivial $\mathcal{O}(\varepsilon^{-1})$- and the natural $\mathcal{O}(\varepsilon^{-2})$-time scale is to use so called normal form transformations. As already said, by these near identity change of variables a number of cubic terms in the nonlinearity can be transformed in terms of higher order.

4.1 Analysis of the non-resonance condition

It is well known that for the elimination of a term $\widehat{a}_{j_1} * \widehat{a}_{j_2} * R_{j_3}$ by a normal form transformation from the equation for R_j the non-resonance condition

$$-j\omega(k) - \omega(j_1) - \omega(j_2) + j_3\omega(k - j_1 - j_2) \neq 0 \qquad (22)$$

has to be satisfied for all $k \in \mathbb{R}$.

Remark 4.1. The non-resonance condition (22) with $j_1, j_2 \in \{-1, 1\}$ is obtained from an expression

$$-j\omega(k) - \omega(k_1) - \omega(k_2) + j_3\omega(k - k_1 - k_2) \neq 0 \tag{23}$$

for wave numbers $k, k_1, k_2 \in \mathbb{R}$. Since a_1 is concentrated in an $\mathcal{O}(\varepsilon)$-neighborhood of $k = 1$ and a_{-1} in an $\mathcal{O}(\varepsilon)$-neighborhood of $k = -1$, in the end the non-resonance condition (23) can be replaced by the non-resonance condition (22). The error made by this replacement is of order $\mathcal{O}(\varepsilon)R$ such that finally only additional $\mathcal{O}(\varepsilon^2)R$-terms occur in the equations for the error when going from (23) to (22). This could be made rigorous for instance by applying the subsequent Lemma 4.3. See also Section 5.

The graphical analysis of the non-resonance condition (22) can be found in Figure 2 and Figure 3. Resonances correspond to zeroes in these figures. So two kind of resonances occur.

i) For $(j, j_1, j_2, j_3) = (j, j_1, -j_1, j)$, we have a total resonance, i.e., the left-hand side of (22) vanishes identically, and a normal form transformation is not possible. See the fifth picture of Figure 2 and the third picture of Figure 3.

ii) For $(j, j_1, j_2, j_3) = (-1, j_1, j_1, -1)$ there is a resonance in $k = j_1$, which is not linear but of second order. Indeed, a Taylor expansion of the non-resonance condition (22) for k near j_1 shows

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

See the second and sixth picture of Figure 2 and the explanations made in Section 2.

4.2 The total resonance

We start with the total resonance and write (21) as

$$\partial_t R_j = ji\omega R_j + j\frac{3}{2}\varepsilon\, i\rho\Big(\sum_{j_1, j_2, j_3 \in \{\pm 1\}} a_{j_1} a_{j_2} R_{j_3}\Big) + \mathcal{O}(\varepsilon^2)\,. \tag{24}$$

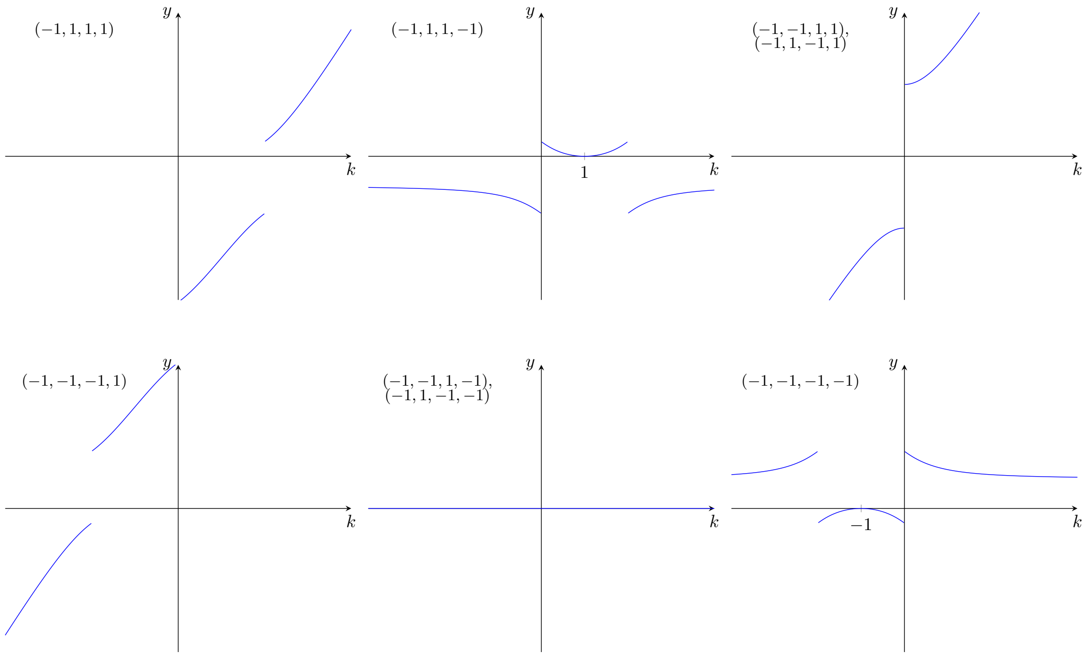


Figure 2: Plots of the left-hand sides of (22) depending on $(-1, j_1, j_2, j_3)$. No resonances are present in the first, third, and fourth picture. Picture five shows the total resonance. The second and the sixth picture show the resonance at $k = 1$ and $k = -1$ with the second order touching.

It turns out that the totally resonant terms can be controlled by energy estimates. For the evolution of the L^2-energy of the error we find

$$\begin{aligned} \partial_t \|R_j\|_{L^2}^2 &= 2 \int_{\mathbb{R}} R_j \partial_t R_j \,\mathrm{d}x \\ &= 2j \int_{\mathbb{R}} R_j i\omega R_j \,\mathrm{d}x + 3j\varepsilon \sum_{j_1, j_2, j_3 \in \{\pm 1\}} \int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1} a_{j_2} R_{j_3}\big) \,\mathrm{d}x + \mathcal{O}(\varepsilon^2) \\ &= 3j\varepsilon \sum_{j_1, j_2, j_3 \in \{\pm 1\}} \int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1} a_{j_2} R_{j_3}\big) \,\mathrm{d}x + \mathcal{O}(\varepsilon^2)\,, \end{aligned} \tag{25}$$

using the skew symmetry of $i\omega$ and the fact that $R_j(x, t) \in \mathbb{R}$ for $j = \pm 1$. By taking a closer look at the last integrals, it turns out that they are of order

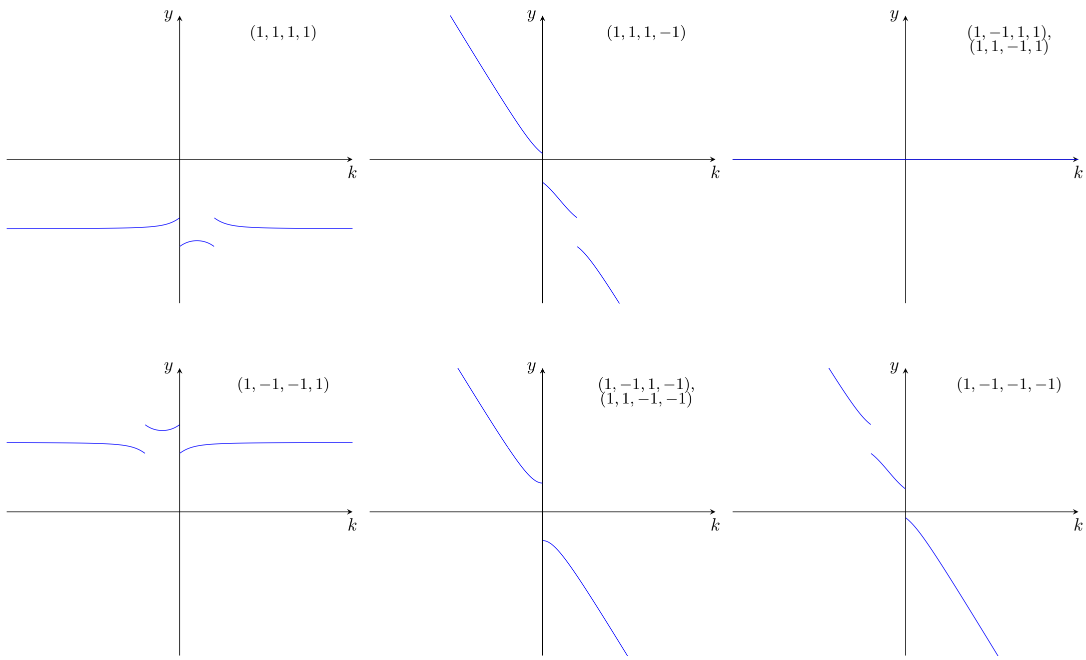


Figure 3: Plots of the left-hand sides of (22) depending on $(1, j_1, j_2, j_3)$. The third picture shows the total resonance. In all other pictures there is no resonance.

$\mathcal{O}(\varepsilon^2)$ for the totally resonant terms, i.e., for $(j, j_1, j_2, j_3) = (j, j_1, -j_1, j)$. Hence, an elimination of the nonlinear terms corresponding to $(j, j_1, j_2, j_3) = (j, j_1, -j_1, j)$ by a normal form transformation is not necessary.

Lemma 4.2. *We have*

$$|\varepsilon \int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1} a_{-j_1} R_j\big) \,\mathrm{d}x| \leq C\varepsilon^2 \|R_j\|_{L^2}^2 \,. \tag{26}$$

Proof. The key ingredients of the proof are that $i\rho$ is a skew symmetric operator and that $\partial_x(a_1 a_{-1}) = \partial_x(|A_c(\varepsilon\cdot)|^2) = \mathcal{O}(\varepsilon)$. In detail, we get with

Plancherel's identity and the fact that $a_{j_1}a_{-j_1}$ is real-valued that

$$\begin{aligned}
&\varepsilon \int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1}a_{-j_1}R_j\big) \,\mathrm{d}x \\
&= \frac{\varepsilon}{2}\int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1}a_{-j_1}R_j\big) \,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}} (i\rho R_j) a_{j_1}a_{-j_1}R_j \,\mathrm{d}x \\
&= \frac{\varepsilon}{2}\int_{\mathbb{R}} R_j \, i\rho\big(a_{j_1}a_{-j_1}R_j\big) \,\mathrm{d}x - \frac{\varepsilon}{2}\int_{\mathbb{R}} R_j a_{j_1}a_{-j_1}(i\rho R_j) \,\mathrm{d}x \\
&= \pi\frac{\varepsilon}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} i(\rho(k)-\rho(m)) \, \overline{\widehat{R}_j(k)}\widehat{|A_c|^2}\big(\frac{k-m}{\varepsilon}\big)\widehat{R}_j(m) \,\mathrm{d}m \,\mathrm{d}k \,.
\end{aligned}$$

Since ρ is Lipschitz-continuous we have $|\rho(k)-\rho(m)| \leq C|k-m|$ such that the assertion follows by applying the subsequent Lemma 4.3 and using estimate (19). ☐

Lemma 4.3. *Fix $p \in \mathbb{R}$ and $n \in \mathbb{N}$. Assume that $g \in C^{n+1}(\mathbb{R},\mathbb{C})$ has a Fourier transform with a bounded support and that $f \in H^s(\mathbb{R},\mathbb{C})$ for $s \geq 0$. Then*

$$\Big\| \int (\cdot - l - p)^n \, \widehat{g}\Big(\frac{\cdot - \ell - p}{\varepsilon}\Big)\widehat{f}(\ell)\, d\ell \,\Big\|_{L^2_s} \leq C\varepsilon^n \|f\|_{H^s} \,. \tag{27}$$

Proof. See the calculations below [SU17, Lemma 11.5.4]. ☐

Remark 4.4. With the same argument we also have for all $\ell \in \mathbb{N}_0$ that

$$\varepsilon \int_{\mathbb{R}} (\partial_x^\ell R_j)\, \partial_x^\ell\big(i\rho\big(a_{j_1}a_{-j_1}R_j\big)\big) \,\mathrm{d}x = \mathcal{O}(\varepsilon^2)\|R_j\|^2_{H^\ell} \,. \tag{28}$$

4.3 The non-resonent terms

In order to get rid of the remaining terms of $\mathcal{O}(\varepsilon)$ in (21) which are not close to the second order resonances, we use the normal form transform

$$R_j \mapsto W_j := R_j + \varepsilon \sum_{j_1,j_2,j_3\in\{\pm 1\}} \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1}, a_{j_2}, R_{j_3}). \tag{29}$$

In Section 4.2 we have already seen that we cannot eliminate the totally resonant terms and so we set

$$\mathcal{M}_{j,j_1,-j_1,j} = 0.$$

Hence, in the following assume $(j, j_1, j_2, j_3) \neq (j, j_1, -j_1, j)$. Since no further resonances occur for $(j, j_1, j_2, j_3) \neq (-1, j_1, j_1, -1)$, for these indices we set

$$\begin{aligned} &\widehat{\mathcal{M}}_{j,j_1,j_2,j_3}(a_{j_1}, a_{j_2}, R_{j_3})(k) \\ &\quad = \int_{\mathbb{R}} \widehat{m}_{j,j_1,j_2,j_3}(k)\widehat{a}_{j_1}(k-m)\widehat{a}_{j_2}(m-n)\widehat{R}_{j_3}(n)\,\mathrm{d}n\,\mathrm{d}m\,, \end{aligned} \tag{30}$$

with

$$\widehat{m}_{j,j_1,j_2,j_3}(k) = -j\frac{3}{2}\frac{\rho(k)}{-j\omega(k) - \omega(j_1) - \omega(j_2) + j_3\omega(k - j_1 - j_2)}\,. \tag{31}$$

4.4 At the second order resonance

Hence it remains to get rid of the indices $(j, j_1, j_2, j_3) = (-1, j_1, j_1, -1)$, i.e., of the second order resonance terms. As already explained above the normal form transformation would become singular when approaching the second order resonance at the wave numbers $j_1 = \pm 1$. Since $\rho(k) = \mathcal{O}(|k - j_1|)$ for k close to the wave numbers $j_1 = \pm 1$ the normal form transformation would possess a first order singularity at the wave numbers $j_1 = \pm 1$.

We get rid of this problem by adding and subtracting both

$$\kappa\varepsilon^2 a_1 a_1 R_{-1} \qquad \text{and} \qquad \kappa\varepsilon^2 a_{-1} a_{-1} R_{-1}$$

to the equation of R_{-1} with $\kappa = \mathcal{O}(1)$ for $\varepsilon \to 0$ chosen subsequently sufficiently large. In the second picture of Figure 2 by adding $\kappa\varepsilon^2 a_1 a_1 R_{-1}$ we can shift the second order resonance $\mathcal{O}(\varepsilon^2)$-away from the k-axis. The subtracted counterpart $-\kappa\varepsilon^2 a_1 a_1 R_{-1}$ is of order $\mathcal{O}(\varepsilon^2)$ and therefore can be easily estimated by Gronwall's inequality. Hence, we set

$$\begin{aligned} &\widehat{\mathcal{M}}_{-1,1,1,-1}(a_1, a_1, R_{-1})(k) \\ &\quad = \int_{\mathbb{R}} \widehat{m}_{-1,1,1,-1}(k)\widehat{a}_1(k-m)\widehat{a}_1(m-n)\widehat{R}_{-1}(n)\,\mathrm{d}n\,\mathrm{d}m\,, \end{aligned} \tag{32}$$

with

$$\widehat{m}_{-1,1,1,-1}(k) = \frac{3}{2}\frac{\rho(k)}{\omega(k) - \omega(1) - \omega(1) - \omega(k-2) + \kappa\varepsilon^2}\,. \tag{33}$$

We have

$$|\varepsilon\widehat{m}_{-1,1,1,-1}(k)| \approx \varepsilon + |\frac{\varepsilon(k-1)}{(k-1)^2 + \kappa\varepsilon^2}| = \mathcal{O}(1)$$

for $\varepsilon \to 0$, but $\mathcal{O}(\varepsilon + \kappa^{-1})$ for $\kappa \to \infty$. Therefore, although the part $\varepsilon\mathcal{M}_{-1,1,1,-1}$ is $\mathcal{O}(1)$ in the transformation (29), this transformation is still invertible for $\kappa > 0$ sufficiently large and $\varepsilon > 0$ sufficiently small.

We do exactly the same with $-\kappa\varepsilon^2 a_{-1}a_{-1}R_{-1}$ which allows us to shift the second order resonance $\mathcal{O}(\varepsilon^2)$-away from the k-axis in the sixth picture of Figure 2.

5 Energy estimates

This section contains a **Proof of Theorem 3.3:** Since the totally resonant terms have been estimated with energy estimates we proceed as in the existing literature, cf. [Sch05], and include the normal form transformations in our energy. Therefore, the subsequent energy estimates for the totally resonant terms and the non-resonant terms will be straightforward. The interesting aspect of this section is the handling of the second order resonance.

5.1 Equivalence of the energy and the H^ℓ-norm

We control the H^ℓ-norm of the error by estimating the L^2-norm of the error and of its ℓ-th derivative. But instead of using the L^2-norm of the transformed error W, we use an energy based on it. We set

$$\mathcal{E}_\ell = E_0 + E_\ell\,, \tag{34}$$

where

$$\begin{aligned} E_\ell = &\sum_{j\in\{\pm1\}} \left\|\partial_x^\ell R_j\right\|_{L^2}^2 \\ &+ \varepsilon \Big(\sum_{j,j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1}, a_{j_2}, R_{j_3})\,\mathrm{d}x + \text{c.c.}\Big) \\ &+ \varepsilon^2 \sum_{j_1\in\{\pm1\}} \left\|\partial_x^\ell \mathcal{M}_{-1,j_1,j_1,-1}(a_{j_1}, a_{j_1}, R_{-1})\right\|_{L^2}^2\,, \end{aligned} \tag{35}$$

and where c.c. denotes the complex conjugate. In comparison to the squared L^2-norm of the transformed error W and of its ℓ-th derivative, we have dropped all unnecessary $\mathcal{O}(\varepsilon^2)$-terms due to he fact that $\mathcal{O}(\varepsilon^2)$-terms in the energy cannot produce or eliminate $\mathcal{O}(\varepsilon)$-terms in the evolution equations for the energy. Dropping these unnecessary $\mathcal{O}(\varepsilon^2)$-terms already here, makes subsequent calculations simpler.

The next lemma guarantees that the square root of the energy $\mathcal{E}_\ell$ is equivalent to the H^ℓ-norm for sufficiently small $\varepsilon > 0$.

Lemma 5.1. *There exist $\varepsilon_0 > 0$, $C_1 > 0$, and $C_2 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have*

$$\left(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\right)^2 \leq C_1\, \mathcal{E}_\ell \leq C_2 \left(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\right)^2.$$

Proof. Although $\varepsilon \mathcal{M}_{-1,1,1,-1}$ is $\mathcal{O}(1)$ we made this part small by choosing $\kappa > 0$ sufficiently large but independent of $0 < \varepsilon \ll 1$. Hence the second and third line of the right-hand side of (35) are a small perturbation of the first line of the right-hand side of (35). □

5.2 The time derivative of the energy

We have

$$\partial_t E_\ell = \sum_{j \in \{\pm 1\}} \sum_{n=1}^{12} s_{n,j} + \text{c.c.} \tag{36}$$

where

$$
\begin{aligned}
s_{1,j} &= j\int_{\mathbb{R}} \partial_x^\ell R_j i\omega \partial_x^\ell R_j \,\mathrm{d}x,\\
s_{2,j} &= j\frac{3}{2}\varepsilon \sum_{j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \, i\rho\partial_x^\ell\big(a_{j_1}a_{j_2}R_{j_3}\big)\,\mathrm{d}x\\
s_{3,j} &= \varepsilon^2\int_{\mathbb{R}} \partial_x^\ell R_j \,\partial_x^\ell \mathcal{L}_j(R)\,\mathrm{d}x,\\
s_{4,j} &= \varepsilon \sum_{j_1,j_2,j_3\in\{\pm1\}} j\int_{\mathbb{R}} \partial_x^\ell i\omega R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x\\
s_{5,j} &= j\frac{3}{2}\varepsilon^2 \sum_{j_1,j_2,j_3,j_4,j_5,j_6\in\{\pm1\}} \int_{\mathbb{R}} \overline{i\rho\partial_x^\ell\big(a_{j_4}a_{j_5}R_{j_6}\big)}\partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x\\
s_{6,j} &= \varepsilon^3 \sum_{j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell \mathcal{L}_j(R) \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x\\
s_{7,j} &= \varepsilon \sum_{j_1,j_2,j_3\in\{\pm1\}} j_3\int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},i\omega R_{j_3})\,\mathrm{d}x\\
s_{8,j} &= j_3\frac{3}{2}\varepsilon^2 \sum_{j_1,j_2,j_3,j_4,j_5,j_6\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}\big(a_{j_1},a_{j_2},i\rho(a_{j_4}a_{j_5}R_{j_6})\big)\,\mathrm{d}x\\
s_{9,j} &= \varepsilon^3 \sum_{j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},\mathcal{L}_{j_3}(R))\,\mathrm{d}x\\
s_{10,j} &= \varepsilon \sum_{j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(\partial_t a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x\\
s_{11,j} &= \varepsilon \sum_{j_1,j_2,j_3\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},\partial_t a_{j_2},R_{j_3})\,\mathrm{d}x,\\
s_{12,j} &= \varepsilon^2 \sum_{j_1\in\{\pm1\}} \int_{\mathbb{R}} \overline{\partial_x^\ell \mathcal{M}_{-1,j_1,j_1,-1}(a_{j_1},a_{j_1},R_{-1})}\,\partial_t\partial_x^\ell \mathcal{M}_{-1,j_1,j_1,-1}(a_{j_1},a_{j_1},R_{-1})\,\mathrm{d}x
\end{aligned}
$$

with

$$
\mathcal{L}_j(R) := L_{s,j}(R) + \varepsilon^{\beta-3/2}L_{r,j}(R) + \varepsilon^{-\beta-2}\mathrm{Res}_j(\varepsilon^{1/2}\psi)\,,
$$

and $L_s(R) =: \big(L_{s,-1}(R), L_{s,1}(R)\big)^T$, $L_r(R) =: \big(L_{r,-1}(R), L_{r,1}(R)\big)^T$ and $\mathrm{Res}(\varepsilon^{1/2}\psi) =: \big(\mathrm{Res}_{-1}(\varepsilon^{1/2}\psi), \mathrm{Res}_1(\varepsilon^{1/2}\psi)\big)^T$, cf. (21). Note that here we could sometimes refrain from writing the complex conjungates since R_j is real-valued.

5.3 Bounds on $s_1, \ldots, s_{12}$

In the end we have to estimate $s_1, \ldots, s_{12}$ by an $\mathcal{O}(\varepsilon^2)$-bound in order to get estimates for the error on the long $\mathcal{O}(\varepsilon^{-2})$-time scale. Since $\mathcal{M}_{j,j_1,j_2,j_3}$ can be of order $\mathcal{O}(\varepsilon^{-1})$, getting the estimates is more than a pure counting of powers of ε.

5.3.1 Some trivial bounds

i) We have

$$s_{1,j} + c.c. = 0$$

due to the skew-symmetry of $i\omega_j$.

ii) Although $\mathcal{M}_{j,j_1,j_2,j_3}$ can be of order $\mathcal{O}(\varepsilon^{-1})$ there are three terms which can be handled by a pure counting of powers of ε and so we get

$$|s_{3,j} + s_{6,j} + s_{9,j} + c.c.| \leq C\varepsilon^2(\mathcal{E}_\ell + 1))$$

since $\beta > 3/2$ and since the residual is small enough, see (17). Due to Corollary 5.1, we can estimate these terms against $\mathcal{E}_\ell$.

5.3.2 Separating the resonant terms

We recall that the complete analysis with the resonances and normalform transformations was done to control $\mathcal{O}(\varepsilon)$-terms in the equations of the error. They are contained in $s_{2,j}$. Since we handle different parts differently we split them in totally resonant terms, second order resonant terms, and non-resonant terms

$$s_{2,j} = s_{TR,j} + s_{SOR,j} + s_{NON,j}$$

where

$$\begin{aligned}
s_{TR,j} &= j\frac{3}{2}\varepsilon \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \, i\rho \partial_x^\ell \big(a_{j_1} a_{-j_1} R_j\big)\, \mathrm{d}x, \\
s_{SOR,-1} &= -\frac{3}{2}\varepsilon \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{-1} \, i\rho \partial_x^\ell \big(a_{j_1} a_{j_1} R_{-1}\big)\, \mathrm{d}x, \\
s_{SOR,1} &= 0, \\
s_{NON,j} &= j\frac{3}{2}\varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j \, i\rho \partial_x^\ell \big(a_{j_1} a_{j_2} R_{j_3}\big)\, \mathrm{d}x.
\end{aligned}$$

5.3.3 Bounds for the totally resonant terms

The totally resonant terms are collected in $s_{TR,j}$. The $\mathcal{O}(\varepsilon^2)$-bound has been established in Section 4.2.

5.3.4 Recovering the normal form transform

The terms $s_{4,j}$, $s_{7,j}$, $s_{10,j}$, and $s_{11,j}$ will be used to get rid of the non-resonant terms collected in $s_{SOR,j}$ and $s_{NON,j}$ by recovering the normal form transformation from the previous section.

i) We first split $s_{4,j}$ into

$$\begin{aligned} s_{4,-1,SOR} &= \varepsilon \sum_{j_1\in\{\pm 1\}} -\int_{\mathbb{R}} \partial_x^\ell i\omega R_{-1}\partial_x^\ell \mathcal{M}_{-1,j_1,j_1,-1}(a_{j_1},a_{j_1},R_{-1})\,\mathrm{d}x, \\ s_{4,j,NON} &= \varepsilon \sum_{rest} j\int_{\mathbb{R}} \partial_x^\ell i\omega R_j\partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x \end{aligned}$$

ii) Next we split $s_{7,j}$ by writing

$$i\omega(n)\widehat{R}_{j_3} = i\omega(k-j_1-j_2)\widehat{R}_{j_3} + (i\omega(n) - i\omega(k-j_1-j_2))\widehat{R}_{j_3}$$

into

$$s_{7,j} = t_{7,j,SOR} + r_{7,j,SOR} + t_{7,j,NON} + r_{7,j,NON},$$

with

$$\begin{aligned} t_{7,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j\partial_x^\ell \mathcal{K}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})\,\mathrm{d}x, \\ r_{7,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j\partial_x^\ell \mathcal{Q}_{j,j_1,j_2,j_3}\big(a_{j_1},a_{j_2},R_{j_3}\big)\,\mathrm{d}x. \end{aligned}$$

where

$$\begin{aligned} \widehat{\mathcal{K}}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})(k) = \int_{\mathbb{R}} & j_3 i\omega(k-j_1-j_2) \\ & \times \widehat{m}_{j,j_1,j_2,j_3}(k)\widehat{a}_{j_1}(k-m)\widehat{a}_{j_2}(m-n)\widehat{R}_{j_3}(n)\,\mathrm{d}n\,\mathrm{d}m\,, \\ \widehat{\mathcal{Q}}_{j,j_1,j_2,j_3}(a_{j_1},a_{j_2},R_{j_3})(k) = \int_{\mathbb{R}} & j_3(i\omega(n) - i\omega(k-j_1-j_2)) \\ & \times \widehat{m}_{j,j_1,j_2,j_3}(k)\widehat{a}_{j_1}(k-m)\widehat{a}_{j_2}(m-n)\widehat{R}_{j_3}(n)\,\mathrm{d}n\,\mathrm{d}m, \end{aligned}$$

and similarly for $t_{7,j,SOR}$ and $r_{7,j,SOR}$.

iii) Finally, we split $s_{10,j}$ and $s_{11,j}$ by writing

$$\partial_t a_j = -ji\omega_0 a_j + (\partial_t a_j + ji\omega_0 a_j) \tag{37}$$

into

$$\begin{aligned} s_{10,j} &= t_{10,j,SOR} + r_{10,j,SOR} + t_{10,j,NON} + r_{10,j,NON}, \\ s_{11,j} &= t_{11,j,SOR} + r_{11,j,SOR} + t_{11,j,NON} + r_{11,j,NON} \end{aligned}$$

with

$$\begin{aligned} t_{10,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(-j_1 i\omega_0 a_{j_1}, a_{j_2}, R_{j_3})\,\mathrm{d}x, \\ r_{10,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}\big((\partial_t a_{j_1} + j_1 i\omega_0 a_{j_1}), a_{j_2}, R_{j_3}\big)\,\mathrm{d}x, \\ t_{11,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}(a_{j_1}, -j_2 i\omega_0 a_{j_2}, R_{j_3})\,\mathrm{d}x, \\ r_{11,j,NON} &= \varepsilon \sum_{rest} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}\big(a_{j_1}, (\partial_t a_{j_2} + j_2 i\omega_0 a_{j_2}), R_{j_3}\big)\,\mathrm{d}x., \end{aligned}$$

and similarly for $t_{10,j,SOR}$, $r_{10,j,SOR}$, $t_{11,j,SOR}$, and $r_{11,j,SOR}$.

5.3.5 The normal form transformation for the non-resonant terms

The terms $s_{4,j,NON}$, $t_{7,j,NON}, t_{10,j,NON}$ and $t_{11,j,NON}$ are the ones which we used in Section 4.3 for the elimination of $s_{NON,j}$ and where we needed the validity of the non-resonance condition (22). Therefore, by construction we have

$$s_{NON,j} + s_{4,j,NON} + t_{7,j,NON} + t_{10,j,NON} + t_{11,j,NON} = 0. \tag{38}$$

5.3.6 Estimates for the remaining terms in the non-resonant case

Next we need to show

$$r_{7,j,NON} + r_{10,j,NON} + r_{11,j,NON} = \mathcal{O}(\varepsilon^2)(\mathcal{E}_\ell + 1)\,. \tag{39}$$

Since

$$\begin{aligned} \partial_t a_j + ji\omega_0 a_j &= -\varepsilon c_g (\partial_X A_c)(\varepsilon(x - c_g t), \varepsilon^2 t)\, e^{ji(k_0 x - \omega_0 t)} \\ &\quad + \varepsilon^2 (\partial_T A_c)(\varepsilon(x - c_g t), \varepsilon^2 t)\, e^{ji(k_0 x - \omega_0 t)}\,, \end{aligned} \tag{40}$$

Estimate (39) follows for $r_{10,j,NON} + r_{11,j,NON}$ by are pure counting of powers of ε. For estimating $r_{7,j,NON}$ respectively $\widehat{\mathcal{Q}}_{j,j_1,j_2,j_3}(a_{j_1}, a_{j_2}, R_{j_3})$ we use

$$\begin{aligned} i(\omega(n) - \omega(k - j_1 - j_2)) &= i\omega'(k - j_1 - j_2)(n - k + j_1 + j_2) \\ &\quad + r_1(n - k + j_1 + j_2), \end{aligned}$$

with $r(n - k + j_1 + j_2) = \mathcal{O}\big((n - k + j_1 + j_2)^2\big)$. Since

$$n - k + j_1 + j_2 = -\big((k - m) - j_1\big) - \big((m - n) - j_2\big) ,$$

and

$$((k - m) - j_1)\widehat{a}_{j_1}(k - m) = \mathcal{O}(\varepsilon) , \qquad ((m - n) - j_2)\widehat{a}_{j_2}(m - n) = \mathcal{O}(\varepsilon)$$

we found the missing power of ε also for $r_{7,j,NON}$.

5.3.7 The normal form transformation for the SOR terms

We proceed as explained in Section 4.4. We add and subtract

$$\begin{aligned} s_{13,j} &= \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell (\kappa \varepsilon^2 a_1 a_1 R_{-1}) \, \mathrm{d}x, \\ s_{14,j} &= \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell (\kappa \varepsilon^2 a_{-1} a_{-1} R_{-1}) \, \mathrm{d}x, \end{aligned}$$

with the obvious estimate

$$-s_{13,j} + s_{14,j} = \mathcal{O}(\varepsilon^2)\mathcal{E}_\ell.$$

From the construction in Section 4.4 we have

$$s_{SOR,j} + s_{4,j,SOR} + t_{7,j,SOR} + t_{10,j,SOR} + t_{11,j,SOR} + s_{13,j} - s_{14,j} = 0. \qquad (41)$$

5.3.8 Estimates for the remaining terms in the SOR case

It remains to show

$$r_{7,j,SOR} + r_{10,j,SOR} + r_{11,j,SOR} = \mathcal{O}(\varepsilon^2)(\mathcal{E}_\ell + 1) . \qquad (42)$$

Due to the second order resonance this estimate is more complicated than for the non-resonant terms. We start with $r_{7,j,SOR}$, respectively

$$\begin{aligned} \widehat{\mathcal{Q}}_{-1,j_1,j_1,-1}(a_{j_1}, a_{j_1}, R_{-1})(k) = -\int_{\mathbb{R}} &(i\omega(n) - i\omega(k - j_1 - j_1)) \\ &\times \widehat{m}_{-1,j_1,j_1,-1}(k)\widehat{a}_{j_1}(k - m)\widehat{a}_{j_1}(m - n)\widehat{R}_{-1}(n) \, \mathrm{d}n \, \mathrm{d}m, \end{aligned}$$

Since $\varepsilon\widehat{m}_{-1,j_1,j_1,-1}(k) = \mathcal{O}(1)$ we obtain from the estimates for the non-resonant terms $r_{7,j,NON} + r_{10,j,NON} + r_{11,j,NON}$ only $\mathcal{O}(\varepsilon)$. For the second power w.r.t. ε additional work is necessary. We are left with

$$\begin{aligned} & r_{7,j,SOR} + r_{10,j,SOR} + r_{11,j,SOR} \\ & = \varepsilon^2 \sum_{j_1\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_{-1} \partial_x^\ell \mathcal{P}_{-1,j_1,j_1,-1}\big(\tilde{a}_{j_1}, a_{j_1}, R_{-1}\big)\,\mathrm{d}x \\ & \quad + \mathcal{O}(\varepsilon^2)(\mathcal{E}_\ell + 1)\,, \end{aligned}$$

where

$$\tilde{a}_{j_1}(x,t) = (\partial_X A_c)(\varepsilon(x - c_g t), \varepsilon^2 t)\, e^{j_1 i(k_0 x - \omega_0 t)}\,,$$

and

$$\begin{aligned} & \widehat{\mathcal{P}}_{-1,j_1,j_1,-1}\big(\tilde{a}_{j_1}, a_{j_1}, R_{-1}\big)(k) \\ & = \int_{\mathbb{R}} \widehat{p}_{-1,j_1,j_1,-1}(k)\widehat{\tilde{a}}_{j_1}(k-m)\widehat{a}_{j_1}(m-n)\widehat{R}_{-1}(n)\mathrm{d}n\mathrm{d}m, \end{aligned}$$

with

$$\widehat{p}_{-1,j_1,j_1,-1}(k) = 2(\omega'(k-2j_1) - c_g)\widehat{m}_{-1,j_1,j_1,-1}(k)\,.$$

Due to

$$\begin{aligned} 2\big(\omega'(k-2j_1) - c_g\big) &= 2\big(\omega'(k-2j_1) - \omega'(1)\big) \\ &= 2\big(\omega'(k-2j_1) - \omega'(-j_1)\big) = \mathcal{O}(k - j_1)\,, \end{aligned}$$

we now obtain

$$\widehat{p}_{-1,j_1,j_1,-1}(k) = 2(\omega'(k-2j_1) - c_g)\widehat{m}_{-1,j_1,j_1,-1}(k) = \mathcal{O}(1) \qquad (43)$$

and can conclude

$$r_{7,j,SOR} + r_{10,j,SOR} + r_{11,j,SOR} = \mathcal{O}(\varepsilon^2)(\mathcal{E}_\ell + 1)\,. \qquad (44)$$

5.3.9 Estimates for $s_{5,j}$

In the integrals $s_{5,j}$ and $s_{8,j}$ the presence of the operator ρ prevents a loss of ε-powers near the resonances. Let us consider the integral in $s_{5,j}$ first. With Plancherel's identity we can write

$$s_{5,j} = -j3\pi\,\varepsilon^2 \sum_{j_1,j_2,j_3,j_4,j_5,j_6\in\{\pm1\}} s_{5,j,j_1,j_2,j_3,j_4,j_5,j_6}$$

where

$$s_{5,j,j_1,j_2,j_3,j_4,j_5,j_6} = \int_{\mathbb{R}} i\rho(k)\widehat{m}_{j,j_1,j_2,j_3}(k)\overline{(ik)^\ell \widehat{a}_{j_4}(k-k_1)\widehat{a}_{j_5}(k_1-k_2)R_{j_6}(k_2)} \times (ik)^\ell a_{j_1}(k-k_3)a_{j_2}(k_3-k_4)R_{j_3}(k_4)\,\mathrm{d}k_1\mathrm{d}k_2\mathrm{d}k_3\mathrm{d}k_4\mathrm{d}k\,.$$

Since $\rho(j_1)=0$ we have for the kernel in Fourier space

$$m_{j,j_1,j_2,j_3}(k)\rho(k) = \mathcal{O}(1)\,,$$

even in the second order resonance case, where

$$m_{-1,j_1,j_1,-1}(k)\rho(k) \approx C\frac{(k\pm 1)^2}{(k\pm 1)^2+\kappa\varepsilon^2}.$$

Therefore,

$$s_{5,j} = \varepsilon^2\mathcal{O}(\mathcal{E}_\ell + 1).$$

5.3.10 Estimates for $s_{8,j}$

For estimating the integral in $s_{8,j}$ we use the Cauchy-Schwarz inequality. We use Plancherel's identity and Lemma 4.3 to estimate

$$\varepsilon^2\left\|\partial_x^\ell \mathcal{M}_{j,j_1,j_2,j_3}\big(a_{j_1}, a_{j_2}, i\rho(a_{j_4}a_{j_5}R_{j_6})\big)\right\|_{L^2}. \tag{45}$$

Using the fact that the a_j are concentrated at the wave number j, as above, this can be written as a term with a kernel

$$m_{j,j_1,j_2,j_3}(k)\rho(k-j_1-j_2)$$

plus a term which can be estimated by $\varepsilon^2\mathcal{O}(\mathcal{E}_\ell+1)$. In the non-resonant case we have $m_{j,j_1,j_2,j_3}(k) = \mathcal{O}(1)$ and so the estimate follows. In the second order resonant case we use

$$m_{-1,j_1,j_1,-1}(k)\rho(k-2j_1) \approx \frac{i(k-j_1)}{(k-j_1)^2+\kappa\varepsilon^2}(k-j_1)\times\ldots = \mathcal{O}(1).$$

Therefore,

$$s_{8,j} = \varepsilon^2\mathcal{O}(\mathcal{E}_\ell + 1).$$

5.3.11 Estimates for $s_{12,j}$

Now consider the term $s_{12,j}$. As above we get

$$\begin{aligned}&\partial_t\widehat{\mathcal{M}}_{-1,j_1,j_1,-1}(a_{j_1},a_{j_2},R_{j_3})(k)\\&\quad=\big(-i\omega(j_1)-i\omega(j_1)-i\omega(k-2j_1)\big)\,\widehat{\mathcal{M}}_{-1,j_1,j_1,-1}(a_{j_1},a_{j_1},R_{-1})(k)+\mathcal{O}(\varepsilon)\,.\end{aligned}$$

By construction of the normal form transformation we have

$$\begin{aligned}&\big(-2i\omega(j_1)-i\omega(k-2j_1)\big)\,m_{-1,j_1,j_1,-1}(k)\\&\quad=\ -i\omega(k)m_{-1,j_1,j_1,-1}(k)+\frac{3}{2}i\rho(k).\end{aligned}$$

We split the integral according to the two terms on the right-hand side. Exploiting the skew symmetry of the operator $i\omega$ the integral belonging to the first right-hand side term vanishes. The second right-hand side term can be estimated exactly as $s_{5,j}$. Therefore,

$$s_{8,j}=\varepsilon^2\mathcal{O}(\mathcal{E}_\ell+1).$$

5.4 Gronwall's inequality

Summarizing all results, we obtain

$$\partial_t\mathcal{E}_\ell\leq\varepsilon^2\mathcal{O}(\mathcal{E}_\ell+1), \tag{46}$$

or more detailed, there exist constants C_1, $C_3>0$ independent of $\mathcal{E}_\ell$ and $\varepsilon\in(0,1]$ and a monotonically increasing function $C_2(\mathcal{E}_\ell)>0$ independent of $\varepsilon\in(0,1]$ such that

$$\partial_t\mathcal{E}_\ell\leq C_1\varepsilon^2\mathcal{E}_\ell+C_2(\mathcal{E}_\ell)\varepsilon^{\beta+1/2}\mathcal{E}_\ell+C_3\varepsilon^2,$$

A standard application of Gronwall's inequality gives the $\mathcal{O}(1)$-boundedness of $\mathcal{E}_\ell$ for all $t\in[0,T_0/\varepsilon^2]$ as long as $\varepsilon_0>0$ is chosen sufficiently small.

Thus, for sufficiently small $\varepsilon_0>0$ there is some constant C_R such that

$$\sup_{[0,T_0/\varepsilon^2]}\left\|\begin{pmatrix}R_{-1}\\R_1\end{pmatrix}\right\|_{H^\ell}\leq C_R$$

and so Theorem 3.3 follows. □

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.

[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.

[CLPS99] S. Champeaux, D. Laveder, T. Passot, and P.-L. Sulem. Remarks on the parallel propagation of small-amplitude dispersive. Alfvén waves. *Nonlinear Processes in Geophysics*, 6:169–178, 1999.

[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.

[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.

[Hes22] Max Hess. Validity of the nonlinear Schrödinger approximation for quasilinear dispersive systems with more than one derivative. *Math. Methods Appl. Sci.*, 45(3):1725–1751, 2022.

[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.

[HS22] Max Hess and Guido Schneider. A robust way to justify the derivative nls approximation. *Preprint*, 2022.

[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.

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

[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.

[Mj76] Einar Mjlhus. On the modulational instability of hydromagnetic waves parallel to the magnetic field. *Journal of Plasma Physics*, 16(3):321–334, 1976.

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

[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.