Validity of the nonlinear Schrödinger approximation for quasilinear dispersive systems

Von der Fakultät für Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Max Edgar Heß

aus Stuttgart

Hauptberichter:	**apl. Prof. Dr. Wolf-Patrick Düll**
Mitberichter:	**Prof. C. Eugene Wayne, PhD**
	Prof. Dr. Guido Schneider

Tag der mündlichen Prüfung: 27. November 2019

Institut für Analysis, Dynamik und Modellierung der Universität Stuttgart

2019

Contents

Zusammenfassung

Wir rechtfertigen die Nichtlineare Schrödinger Approximation für eine Klasse von quasilinearen dispersiven Systemen. Wir erlauben nichttriviale Resonanzen und erlauben dem quasilinearen quadratischen Term ein beliebiges Maß an Regularität zu verlieren, solange er nicht mehr Regularität als der lineare Term des Systems verliert. Dies ist das erste Mal, dass die Nichtlineare Schrödinger Approximation für quasilineare dispersive Systeme gerechtfertigt wird, wo der quasilineare Term mehr als eine Ableitung verlieren darf.

Wir leiten die NLS Gleichung über Multi-Skalen-Analysis und das Zeigen von Residuumsabschätzungen her. Wir rechtfertigen die NLS Approximation auf ihrer natürlichen Zeitskala, indem wir Fehlerabschätzungen beweisen. Für die Fehlerabschätzungen verwenden wir eine abgewandelte Energie, die auf gewissen Normalformtransformationen beruht. Diese Energie wird weiter angepasst um die Schließung der Fehlerabschätzungen zu ermöglichen.

Wir geben zudem ein Beispiel dafür, wie unsere Techniken auf allgemeinere quasilineare dispersive Systeme angewandt werden können, indem wir Fehlerabschätzungen für ein reduziertes System zeigen, welches über das zweidimensionale Wasserwellen-Problem mit endlicher Tiefe und Oberflächenspannung motiviert ist.

Abstract

We derive and justify the nonlinear Schrödinger approximation for a class of quasilinear dispersive systems. We allow nontrivial resonances to happen and set no bound on the amount of regularity the quadratic quasilinear term is allowed to lose, apart from not losing more regularity than the linear term of the system does. This is the first time the nonlinear Schrödinger approximation is justified for quasilinear dispersive systems, where the quasilinear term is allowed to lose more than one derivative.
We rigorously derive the NLS equation via multiple scaling analysis and showing residual estimates. We justify the NLS approximation on its natural timescale by proving error estimates. For the error estimates we use a modified energy based on some normal form transformations. This energy gets modified even further in order to allow the closing of the error estimates.
We also give an example how our techniques can be applied to more general quasilinear dispersive systems by showing error estimates for a reduced system, which is motivated by the 2D water wave problem with finite depth and surface tension.

Danksagung und Erklärung

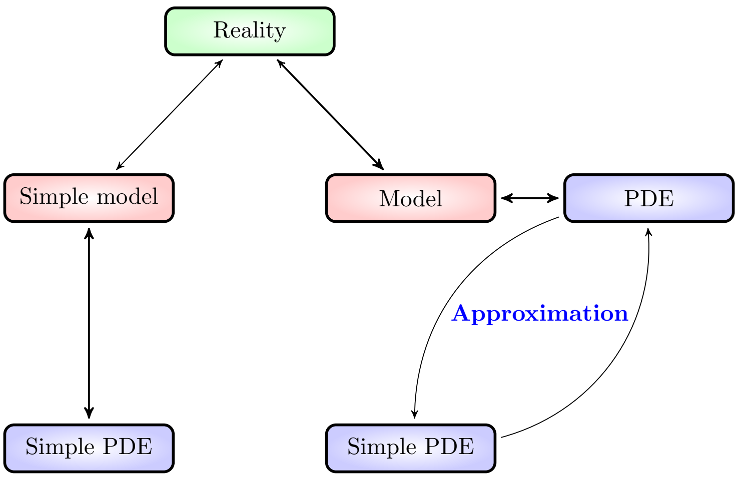
Ich bedanke mich bei meinem Betreuer apl. Prof. Düll, der mich überhaupt erst davon überzeugt hat zu promovieren. Ferner bedanke ich mich auch bei Prof. Schneider für die Unterstützung meiner Promotion. Zudem möchte ich mich bei Prof. Wayne für die herzliche Gastfreundschaft und die netten Diskussionen während meines Aufenthalts an der Boston University bedanken.
Mein Dank gilt der Deutschen Forschungsgemeinschaft DFG, die meine Arbeit unter dem Zeichen DU 1198/2 finanziell unterstützt hat. Ich bedanke mich ebenso bei allen Lesern, insbesondere bei den Gutachtern meiner Dissertation.

Ich erkäre hiermit, dass ich diese Dissertation selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel benutzt habe. Alle Stellen, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, sind von mir durch Angabe der Quelle als Entlehnung kenntlich gemacht.

Chapter 1

Introduction

Being able to foresee a certain outcome, after some action was made or observed, is one of the most valuable skills to have in life. In order to make predictions, one creates a model in which all aspects that are considered the most relevant are covered. Thus, in order to make preciser predictions, almost always a more complicated model is needed. However, a model can quickly become so complicated that a solution is no longer available for the underlying mathematical equations. At this point, one can go back to modeling and attempt to create a simpler model, or, try to approximate the solution of these equations. What makes the second approach especially interesting is that the model does not need to be changed and, since only mathematics is involved, one may be able to prove how close a chosen approximation is to an original solution.



This thesis focuses on the Nonlinear Schrödinger (NLS) approximation for quasilinear dispersive systems. The NLS approximation can be used to describe wave packet like solutions of nonlinear dispersive systems. For nonlinear dispersive systems, the NLS equation

$$\partial_T A = i\nu_1, \partial_X^2 A + i\nu_2 A|A|^2 \,, \tag{1.1}$$

with $T, X \in \mathbb{R}$, $\nu_1 > 0$, $\nu_2 \in \mathbb{R}$ and $A(T,X) \in \mathbb{C}$, can usually be derived via multiple scaling analysis as an modulation equation that describes slow modulations in time and space of the envelope of a temporally and spatially oscillating wave packet.

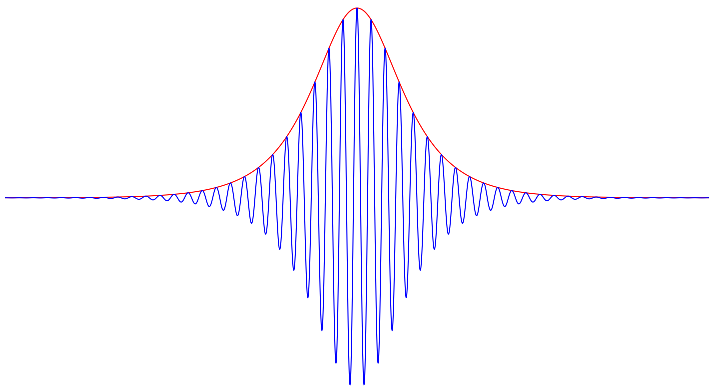


Figure 1.1: The NLS approximation, a temporally and spatially oscillating wave packet with an envelope that is described by the solution of a NLS equation.

Since nonlinear dispersive systems can often be very difficult to solve as well analytically as numerically and the NLS equation can be explicitly solved, the NLS approximation can be a great tool for understanding the dynamics of these systems. For this reason and in order to save computational costs, the NLS approximation is used in nonlinear optics [A01], mathematical physics [Z68], quantum mechanics [P11] and many other fields where one is interested in the evolution of wave packets, e.g. [SH94]. Especially in scenarios where one is only interested in the evolution of the envelope of a wave packet, as for instance the transport of information via light pulses in glass fiber, the NLS approximation drastically increases the efficiency of numerical simulations.

While the NLS approximation is very successful in many applications, a formally derived NLS approximation can make wrong predictions about the behavior of the original system, see [S05, SSZ15]. Thus, error estimates have to be proven in order to show that a NLS approximation is valid. A NLS approximation has to be justified. Only then, one can truly rely on the predictions of a NLS approximation.

In this thesis, we will first consider the nonlinear Schrödinger approximation for a class of quasilinear first order systems, where the nonlinearity is allowed to lose an arbitrary amount of regularity, but not more regularity than the linear term does. Later we consider systems of a more general form, where the nonlinearity of the diagonalized first order system is allowed to lose one derivative. We motivate these systems by the water wave equations.

1.1 Quasilinear dispersive systems

In the first part of this thesis, we consider the Nonlinear Schrödinger approximation for a class of first order systems

$$\begin{aligned} \partial_t u &= -\, i\omega v\,, \\ \partial_t v &= -\, i\omega u - i\rho u^2 \end{aligned} \tag{1.2}$$

with $u : \mathbb{R}\times\mathbb{R} \to \mathbb{R} : (x,t) \mapsto u(x,t)$ and $v : \mathbb{R}^2 \to \mathbb{R}$, where the pseudo differential operators ω and ρ are given by some odd functions $\rho : \mathbb{R} \to \mathbb{R}$ and $\omega : \mathbb{R} \to \mathbb{R}$ in Fourier space.
I.e. in Fourier space, we have

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

Such a first-order system is also equivalent to the equation

$$\partial_t^2 u = -\omega^2 u - \rho\omega u^2\,. \tag{1.3}$$

When there is some $k_0 > 0$ such that the three conditions

$$\omega''(k_0) \neq 0, \tag{1.4}$$

$$\omega'(k_0) \neq \pm\omega'(0) \quad \text{and} \quad \rho(0) = 0, \qquad \text{or} \qquad \lim_{k\to 0^+} \omega(k) \neq 0\,, \tag{1.5}$$

$$m\omega(k_0) \neq \pm\omega(mk_0) \qquad \text{for } m = \pm 2, \ldots, \pm 5\,, \tag{1.6}$$

are fulfilled, we can derive the Nonlinear Schrödinger equation

$$\partial_T A = i\,\frac{\omega''(k_0)}{2}\,\partial_X^2 A + i\nu_2(k_0)A|A|^2\,, \tag{1.7}$$

with $\nu_2(k_0) \in \mathbb{R}$, as a lowest order modulation equation. The explicit formulas for $\nu_2(k_0)$ can be found in section 2.1.
For the derivation, we use an ansatz of the form

$$u = \varepsilon\psi_{NLS} + \mathcal{O}(\varepsilon^2)\,,$$

where

$$\varepsilon\psi_{NLS}(x,t) = \varepsilon A\big(\varepsilon(x - c_g t), \varepsilon^2 t\big)e^{i(k_0 x - \omega_0 t)} + c.c. \tag{1.8}$$

is the Nonlinear Schrödinger approximation for solutions of (1.2).
Here, the complex-valued amplitude A is the solution of the NLS equation (1.7) and $0 < \varepsilon \ll 1$ is a small perturbation parameter. The basic temporal wave number $\omega_0 := \omega(k_0)$ of the underlying carrier wave $e^{i(k_0 x - \omega_0 t)}$ is associated to the basic spatial wave number $k_0 > 0$. The group velocity of the wave packet is $c_g := \omega'(k_0)$ and $c.c.$ simply denotes the complex conjugate.

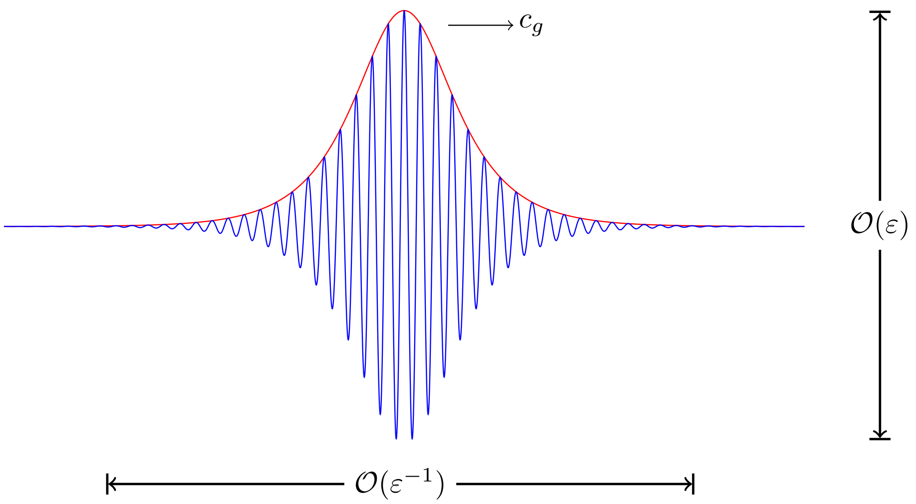


Figure 1.2: The NLS approximation ψ_{NLS} is an oscillating wave packet with an envelope that is described by the solution A of the NLS equation (1.7). As time goes on the envelope of height $\mathcal{O}(\varepsilon)$ and width $\mathcal{O}(\varepsilon^{-1})$ is moving to the right with the group velocity c_g.

The NLS approximation (1.8) describes slow modulations in time and space of a spatially and temporarily oscillating wave packet. The slow time scale of the NLS approximation is $T = \varepsilon^2 t$ and the slow spatial scale $X = \varepsilon(x - c_g t)$, i.e. the time scale of the modulations is $\mathcal{O}(\varepsilon^{-2})$ and the spatial scale of the modulations $\mathcal{O}(\varepsilon^{-1})$.

In order to justify the NLS approximation, we need to make some further restrictions to our class of systems

$$\begin{aligned}\partial_t u &= -\, i\omega v\,,\\ \partial_t v &= -\, i\omega u - i\rho u^2.\end{aligned}$$

First off, we do not allow that the nonlinear terms of our system contain more derivatives than the linear ones. We demand

$$\deg^*(\rho) \leq \ \deg(\omega)\,. \tag{1.9}$$

Here we write $\deg^*(\gamma) \leq s$ for a function $\gamma : \mathbb{R} \to \mathbb{R}$ when there are some constants C, M such that

$$|\gamma(k)| \leq C(1+|k|)^s \quad \text{for} \quad |k| \geq M\,,$$

and $\deg(\gamma) = s$ when there is also some $c > 0$ such that

$$c(1+|k|)^s \leq |\gamma(k)| \leq C(1+|k|)^s \quad \text{for} \quad |k| \geq M.$$

Apart from (1.9), we set no further restriction on the amount of regularity that the quadratic term can lose.
The functions ω and ρ are allowed to have a jump in $k = 0$. However, one of the functions ω or $\operatorname{sign}(\cdot)\omega(\cdot)$, and, one of the functions ρ or $\operatorname{sign}(\cdot)\rho(\cdot)$ have to lie in $C^{m_\omega}(\mathbb{R})$, where $m_\omega = \max\{5, \lceil \deg(\omega) \rceil + 1\}$.
Furthermore, we demand, that for $n = 1, \ldots, m_\omega$, we have

$$\deg^*(\rho^{(n)}) \leq \deg^*(\rho^{(n-1)}) - 1 \tag{1.10}$$

as long as $\rho^{(n)} \neq 0$, and

$$\deg(\omega^{(n)}) = \deg(\omega^{(n-1)}) - 1 \tag{1.11}$$

as long as $\omega^{(n)} \neq 0$. I.e. we want the derivatives of ω and ρ to behave similarly as the ones of polynomials.
We additionally have to assume the local existence of real-valued solutions to our system (1.2) in H^s for some $s \geq \max\{\deg(\omega) + \deg^*(\rho) + 1, s_A\}$ with s_A as in the coming theorem. However we do not think of this as a real restriction since we expect that this local existence can be shown by using the results of [K75a, K75b] or proceeding similarly as in [A03].

In this thesis, we only justify the NLS approximation for cases where up to three resonances can occur. However, we expect that more resonances can be handled by using similar techniques as in [DS06]. We demand that for $j_1, j_2 \in \{\pm 1\}$ the only possible (real-valued) solutions of the equations

$$\omega(k) - j_1 j_2 \omega(k \mp k_0) + j_1 \omega(\pm k_0) = 0 \tag{1.12}$$

are $k = \pm k_0$ and $k = 0$. Solutions of (1.12) correspond to resonances in our normal form transforms.
We explicitly exclude resonances at infinity by demanding that there exists some constant $C > 0$ such that for all $|k| > C$ we have

$$\omega(k_0) \neq \pm k_0\, \omega'(k) \qquad \text{when } \deg(\omega) = 1\,, \tag{1.13}$$

$$\omega(k_0) \neq 0 \qquad \text{when } \deg(\omega) < 1\,, \tag{1.14}$$

$$\omega(k_0) \neq \pm 2\omega(k) \qquad \text{when } \deg(\omega) = 0\,. \tag{1.15}$$

We conditionally allow resonances happening in $k = 0^+$ or $k = 0^-$ by demanding that we always have

$$0 \neq \pm\omega(0^+) \neq 2\omega(k_0), \tag{1.16}$$

or

$$\omega'(k_0) \neq \pm\omega'(0),\ \rho(0) = 0 \quad \text{and } \omega(0^\pm) \neq 2\omega(k_0) + j\omega(2k_0) \qquad \text{for } j \in \{\pm 1\}\,. \tag{1.17}$$

In the case $\omega(0^+) = 0$, (1.17) is already implied by (1.5) and (1.6).
Under these conditions, we obtain:

Theorem 1.1.1. *Fix ω, ρ and $k_0 > 0$ as above and $s_A \geq 7$. For all $C_1, T_0 > 0$ there exists $\varepsilon_0 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ of the NLS equation* (1.7) *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

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

of equation (1.3) *which satisfy*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|u(\cdot, t) - \varepsilon\psi_{NLS}(\cdot, t)\|_{H^{s_A}(\mathbb{R},\mathbb{R})} \lesssim \varepsilon^{3/2}.$$

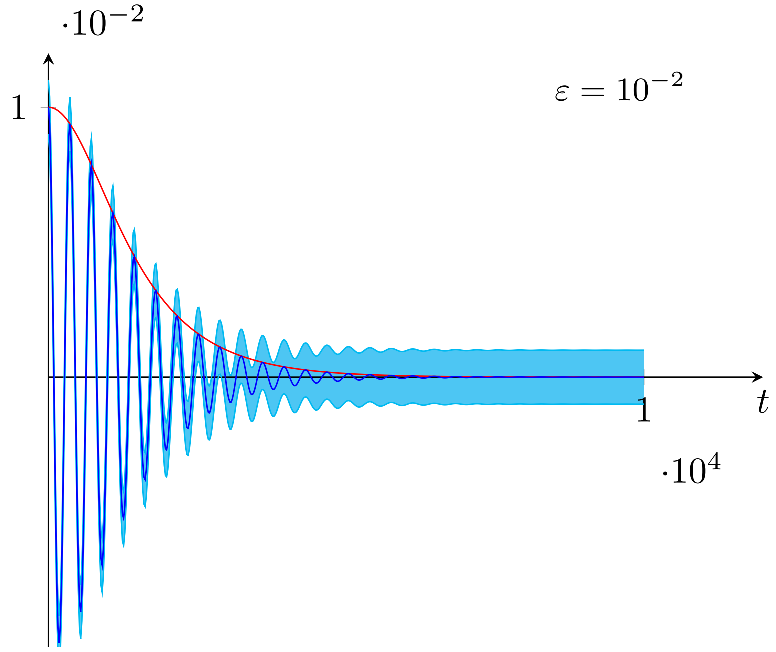


Figure 1.3: Illustration of theorem 1.1.1. The solution of (1.2) cannot leave the $\mathcal{O}(\varepsilon^{3/2})$-tube around the NLS approximation ψ_{NLS} on the $\mathcal{O}(\varepsilon^{-2})$ timescale and the amplitude of ψ_{NLS} is determined by the NLS equation (1.7).

The error the approximation makes is of order $\mathcal{O}(\varepsilon^{3/2})$, which is small compared to the solution u and the approximation $\varepsilon\psi_{NLS}$ that are both of order $\mathcal{O}(\varepsilon)$ in L^∞. Thus, since our estimate holds on the natural time scale of the NLS equation, the dynamics of the NLS equation can be found in (1.2) too. The construction of ψ_{NLS} is always possible since the NLS equation is a completely integrable Hamiltonian system that can be solved explicitly with the help of some inverse scattering scheme, see for example [AS81].

There are various counterexamples, where approximation equations derived by reasonable formal arguments make wrong predictions about the dynamics of the original systems, see for example [SSZ15]. An approximation theorem like theorem 1.1.1 should therefore never be taken for granted.

The smoothness in our error bound is equal to the assumed smoothness of the amplitude. We achieve this by using a modified approximation that has compact support in Fourier space but differs only slightly from $\varepsilon\psi_{NLS}$. Such an approximation can be constructed because the Fourier transform of $\varepsilon\psi_{NLS}$ is sufficiently strongly concentrated around the wave numbers $\pm k_0$, see section 2.1.

Our NLS approximation (1.8) describes wave packets moving to the right with the group velocity c_g. By simply replacing $-\omega_0$ by ω_0 and $-c_g$ by c_g in (1.8), one could describe wave packets that are moving to the left with the group velocity

c_g. Implicitly, such a NLS approximation is also rigorously derived and justified here since the system (1.2) with $(\omega, \rho) = (\tilde{\omega}, \tilde{\rho})$ and the system (1.2) with $(\omega, \rho) = (-\tilde{\omega}, -\tilde{\rho})$ are equivalent to each other, as one can directly see by looking at (1.3).

1.1.1 Difficulties and method of proof

In order to prove theorem (1.1.1), we first off derive the NLS equation (1.7) in section 2.1. We do this by showing that the residual of the NLS approximation is small, i.e. that the terms that remain after plugging in the approximation into the equations of system (1.2) are small. The intuition behind this is that a residual close to zero should be a good indication for that an approximation could work, since the residual of a true solution to the system is zero.
We transform the system (1.2) into an equivalent system of the form

$$\partial_t V = \Lambda V + B(V, V) \tag{1.18}$$

where $V(x,t) \in \mathbb{R}^2$, $\Lambda = \operatorname{diag}(-i\omega, i\omega)$ and B is a symmetric bilinear operator. Then we make the ansatz

$$\begin{aligned} V = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} &= \varepsilon \left(A(X,T)\mathbf{E} + \overline{A}(X,T)\mathbf{E}^{-1} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\quad + \varepsilon^2 \begin{pmatrix} A_0(X,T) + A_2(X,T)\mathbf{E}^2 + \overline{A_2}(X,T)\mathbf{E}^{-2} \\ D_0(X,T) + D_2(X,T)\mathbf{E}^2 + \overline{D_2}(X,T)\mathbf{E}^{-2} \end{pmatrix}, \end{aligned} \tag{1.19}$$

where $X := \varepsilon(x - c_g t)$, $T := \varepsilon^2 t$ and $\mathbf{E} = e^{i(k_0 x - \omega_0 t)}$. Exploiting Taylor's theorem to expand all expressions, like for example

$$i\omega\big[A\mathbf{E}\big] = i\omega(k_0)A\mathbf{E} + \varepsilon\omega'(k_0)\partial_X A\mathbf{E} - \varepsilon^2 i\omega''(k_0)\partial_X^2 A\mathbf{E} + \mathcal{O}(\varepsilon^3)\,,$$

and equating the coefficients in front of $\varepsilon^m \mathbf{E}^j$ for $m \in \{1,2,3\}$ and $j \in \{0,1,2\}$, we obtain the NLS equation (1.7) and a residual of the formal order $\mathcal{O}(\varepsilon^2)$.
Due to (1.6), we can modify our ansatz (1.19) further to obtain an even smaller residual. We finally prove in section 2.1 that there even exists some analytic function Ψ, for which we have a residual $\operatorname{Res}(\varepsilon\Psi)$ with

$$\|\operatorname{Res}(\varepsilon\Psi)\|_{H^s} = \mathcal{O}(\varepsilon^{11/2})$$

for all $s \geq 0$, while

$$\|\varepsilon\Psi - (1,\,0)^T\,\varepsilon\psi_{NLS}\|_{H^{s_A}} = \mathcal{O}(\varepsilon^{3/2}). \tag{1.20}$$

Although a small residual is a good indication for a working approximation, an approximation with a small residual still can fail, see [SSZ15]. For this reason,

we prove in section 2.2 via a priori estimates that the error between the NLS approximation $\varepsilon\Psi$ and an original solution of the system (1.2) stays small on the natural timescale of the modulation.
The following properties of the system (1.2) make this difficult

- a quadratic nonlinearity in the presence of a nontrivial resonance,
- a nonlinear term that can lose regularity and is on top of that quadratic,
- a nonlinearity that can lose an arbitrary amount of regularity and is on top of that quadratic.

We write the error as

$$\varepsilon^{\beta}\vartheta R = V - \varepsilon\Psi \tag{1.21}$$

where $\beta > 1$ and ϑ is an invertible operator on $L^2(\mathbb{R})$ that is given by some weight function $\hat{\vartheta}$ in Fourier space. The constant β and the operator ϑ will be chosen fix later. We now find the rescaled error R to satisfy the evolution equation

$$\partial_t R = \Lambda R + 2\varepsilon\vartheta^{-1}B(\Psi, \vartheta R) + \varepsilon^{\beta}\vartheta^{-1}B(\vartheta R, \vartheta R) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}(\varepsilon\Psi)\,,$$

where $\vartheta^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the inverse of the operator ϑ.
If, by assuming $\|R\|_{H^s} \leq C$ for some $C \geq 0$, we could obtain an estimate of the form

$$\partial_t \|R\|_{H^s}^2 \leq \varepsilon^2\, \mathcal{O}\big(\|R\|_{H^s}^2 + 1\big)$$

then we could exploit Gronwall's inequality to obtain the boundedness of the rescaled error R on the $\mathcal{O}(\varepsilon^{-2})$-timescale.
Choosing $\beta = 5/2$, we obtain

$$\partial_t R = \Lambda R + 2\varepsilon\vartheta^{-1}B(\Psi, \vartheta R) + \mathcal{O}(\varepsilon^2)\,.$$

The term ΛR is fine, since we have $\Lambda = \mathrm{diag}(-i\omega, i\omega)$, $\partial_t\|R\|_{L^2}^2 = \int_{\mathbb{R}} R\,\partial_t R\,dx$ and $\int_{\mathbb{R}} i\omega f\, f\,dx = 0$ for $f \in H^{\deg(\omega)}(\mathbb{R})$.
The term $\varepsilon\vartheta^{-1}B(\Psi, \vartheta R)$, however, has not the right ε-power.
In subsection 2.2.1, we try to eliminate this term by preforming a normal form transformation

$$R \to \tilde{R} = R + \varepsilon\vartheta^{-1}N(\Psi, R)$$

such that

$$\partial_t \tilde{R} = \Lambda\tilde{R} + \mathcal{O}(\varepsilon^2).$$

The idea of using a bilinear mapping N to eliminate a $\mathcal{O}(\varepsilon)$-term was first used by Kalyakin in [K88]. It is well known in literature that due to the strong concentration of the NLS approximation ψ_{NLS} around the wavenumbers $\pm k_0$, a well-defined normal form transformation can be found if the equations (1.12), i.e.

$$\omega(k) - j_1 j_2 \omega(k \mp k_0) + j_1 \omega(\pm k_0) = 0\,,$$

have no solutions for $j_1, j_2 \in \{-1, 1\}$. Solutions to (1.12) are also called resonances, since a normal form transformation $N(\Psi, R)$ could potentially grow unlimitedly in Fourier space for arguments near theses solutions.
We here allow resonances in $k = 0$ and $k = \pm k_0$. Our resonance in $k = 0$ is trivial, i.e. $N(\Psi, R)$ does not grow unlimitedly in Fourier space for arguments near $k = 0$. The resonances in $k = \pm k_0$ however are nontrivial.
Just like in [DS06, DH18], we are still able to find a well-defined normal form transformation by suitably choosing the operator ϑ. However this comes at a price. Due to this, the operator ϑ^{-1} can lose us a ε-power and we only obtain

$$\partial_t \tilde{R} = \Lambda \tilde{R} + \varepsilon^2 \vartheta^{-1} N\big(\Psi, B(\Psi, \vartheta R)\big) + \mathcal{O}(\varepsilon^2).$$

We have to preform a second normal form transformation

$$\tilde{R} \to \check{R} = \tilde{R} + \varepsilon^2 \vartheta^{-1} T(\Psi, \Psi, R)$$

with a trilinear mapping T, before we finally obtain

$$\partial_t \check{R} = \Lambda \check{R} + \mathcal{O}(\varepsilon^2).$$

We can show that this second normal form transformation $T(\Psi, \Psi, R)$ only has trivial resonances, especially due to our additional conditions (1.16), (1.17) for resonances in $k = 0^{\pm}$.
At this point, due to the amount of regularity we allow the nonlinear term of (1.2) to lose, we are now confronted with the following challenges:

- The normal form transformation $R \to \tilde{R} = R + \varepsilon \vartheta^{-1} N(\Psi, R)$ is not invertible since in general $N(\Psi, \cdot)$ does not map L^2 onto L^2. The operator $N(\Psi, \cdot)$ only maps $H^r(\mathbb{R})$ onto $L^2(\mathbb{R})$, when $r \geq \min\big\{\deg^*(\rho),\, 1 + \deg^*(\rho) - \deg(\omega)\big\}$. In particular, this prevents us from estimating the H^s-norm of $\check{R}$ against the H^s-norm of R (the other way around can still be handled).

- A more grave problem is that in general, even by assuming $\|\check{R}\|_{H^s} \leq C$ for some $C \geq 0$, the Gronwall estimates for $\check{R}$ cannot be closed. This is since we have

$$\partial_t \check{R} = \Lambda \check{R} + \varepsilon^2 h(\check{R})\,,$$

where, when $\deg^*(\rho) > 0$, the function h only maps H^{s+r} onto H^s for

$$r \geq \deg^*(\rho) + \min\big\{\deg^*(\rho),\, 1 + \deg^*(\rho) - \deg(\omega)\big\}.$$

In subsection 2.2.2, in order to address the above issues, we proceed similarly as in [D17, DH18] and use a modified energy

$$\begin{aligned}\mathcal{E}_s &:= \|\check{R}\|_{L^2} + E_s\,, \\ E_s &:= \frac{1}{2}\|\partial_x^s R\|_{L^2}^2 + \varepsilon \int_{\mathbb{R}} \partial_x^s R\, \partial_x^s \vartheta^{-1} N(\Psi, R)\, dx\,.\end{aligned} \tag{1.22}$$

Since $\|\partial_x^s \check{R}\|_{L^2}^2$ and $2E_s$ only differ by terms of order $\mathcal{O}(\varepsilon^2)$ we maintain

$$\partial_t \mathcal{E}_s = \mathcal{O}(\varepsilon^2).$$

We gain that:

- The modified energy $\mathcal{E}_s$ is equivalent to the H^s-energy, i.e.

  $$\|R\|_{H^s}^2 \le C_1 \mathcal{E}_s \le C_2 \|R\|_{H^s}^2$$

  for some $C_1, C_2 \ge 0$ (and ε small enough).
- The evolution $\partial_t \mathcal{E}_s$ contains less derivatives falling on R than the evolution of $\|\check{R}\|_{H^s}^2$.

However, due to the amount of regularity we allow the nonlinear term of (1.2) to lose, we still in general do not get

$$\partial_t \mathcal{E}_s = \varepsilon^2\, \mathcal{O}(\mathcal{E}_s + 1)\,,$$

but only

$$\partial_t \mathcal{E}_s = \varepsilon^2\, \mathcal{O}(\mathcal{E}_s + 1) + \varepsilon^2\, g(R)\,,$$

where the function g only maps H^{s+r} onto H^s for $r \ge \deg^*(\rho)$.
The occurrence of this problem is not directly linked to the normal form transformation. In fact, this problem also occurs when one skips the normal form transformation and tries to prove error estimates on a $\mathcal{O}(\varepsilon^{-1})$-timescale.
In subsection 2.2.3, we solve this problem by showing that an expression $\varepsilon^2 \mathcal{D}(R)$ can be constructed such that

$$\varepsilon^2\, \partial_t \mathcal{D}(R) = \varepsilon^2\, g(R) + \varepsilon^2\, \mathcal{O}(\mathcal{E}_s + 1)$$

while at the same time

$$\varepsilon^2\, \mathcal{D}(R) = \varepsilon\, \mathcal{O}(\mathcal{E}_s).$$

This basic idea may go back to [C87] and has already been used in [D17] and [CW17], where two systems of the form (1.2) with $\deg^*(\rho) \le 1$ are considered.

However, for systems with $\deg^*(\rho) > 1$ the expression $\varepsilon^2\mathcal{D}(R)$ is much more difficult to find since a new class of problematic terms is occurring.
We here will present a construction of an expression $\varepsilon^2\mathcal{D}(R)$ that works for arbitrary $\deg^*(\rho)$. This construction heavily relies on (1.9), i.e. it exploits that the linear part of our system is at least as strong as the nonlinear part. A fact that enables us to replace troubling spatial derivatives by time derivatives.
After $\varepsilon^2\mathcal{D}(R)$ is constructed, the final modified energy

$$\tilde{\mathcal{E}}_s := \mathcal{E}_s - \varepsilon^2\mathcal{D}(R)\,, \tag{1.23}$$

fulfills

$$\partial_t\tilde{\mathcal{E}}_s = \varepsilon^2\mathcal{O}(\tilde{\mathcal{E}}_s + 1)\,,$$

while simultaneously being equivalent to the H^s-energy of the error.
After an application of Gronwall's inequality theorem 1.1.1 then follows with (1.20).

1.1.2 Related literature

The first time the NLS equation was derived, it was derived as an amplitude equation for the water wave problem by Zakharov in [Z68]. The first one to actually prove a NLS approximation theorem was Kalyakin in [K88]. Later, Kirrmann, Schneider and Mielke developed a simple method to justify the NLS approximation for systems without quadratic nonlinear terms in [KSM92]. Quadratic nonlinear terms are considered more problematic than other nonlinear terms due to the cubic lifespan the NLS-approximation requires. To illustrate this, let us look at the ordinary differential equation

$$\partial_t u = u^3, \qquad \text{with} \quad u(0) = u_0 \in \mathbb{R}, \quad u_0 = \mathcal{O}(\varepsilon).$$

The solution to this equation with a cubic nonlinearity

$$u(t) = \frac{u_0}{\sqrt{1 - 2u_0^2 t}}$$

has a cubic lifespan, the $\mathcal{O}(\varepsilon^{-2})$-time-interval $[0, \frac{1}{2u_0^2}[$.
However, for the equation with a quadratic nonlinearity

$$\partial_t u = u^2, \qquad \text{with} \quad u(0) = u_0 \in \mathbb{R}, \quad u_0 = \mathcal{O}(\varepsilon)\,,$$

the solution

$$u(t) = \frac{u_0}{1 - u_0 t}$$

only exists on the $\mathcal{O}(\varepsilon^{-1})$-time-interval $[0, \frac{1}{u_0}[$.
As one might guess from this example, the key to a cubic lifespan of solutions to systems like (1.2) has to lie in the linear term. And so, the basic approach for handling quadratic terms is to search for a normal form transformation that, by taking advantage of the linear part of the system, transforms it into a system without quadratic terms. Using normal form transformations Schneider further developed the method of [KSM92] in [S98a] such that quadratic nonlinear terms could be handled, if some non-resonance conditions are fulfilled. There followed some papers, like e.g. [DS06], where these non-resonance conditions were weakened such that more difficult systems with quadratic terms could be considered.
This however excluded systems with quasilinear terms. These are especially problematic since quasilinear terms make it much harder to close error estimates. Quasilinear quadratic terms can also cause normal form transformations to be non-invertible.
Schneider and Wayne were the first ones to prove the validity of the NLS approximation for a system with a quasilinear quadratic term on the qualitatively correct timescale in [SW11]. Thanks to the techniques developed in [SW11] it was then possible to justify the NLS-approximation for the 2-D water wave problem in case of zero surface tension and finite depth in [DSW16]. In [SW11] and [DSW16] quasilinear quadratic terms that lose half a derivative, i.e. $\deg^*(\rho) = 1/2$, were handled with the help of a Cauchy-Kowalevskaya argument. However, the obtained result was still not optimal in the sense that they could not justify the NLS approximation on the whole interval of modulation $[0, T_0/\varepsilon^2]$ but only some smaller $\mathcal{O}(\varepsilon^{-2})$-interval. Another problem of their method is that the Cauchy-Kowalevskaya argument does not work for quasilinear terms with $\deg^*(\rho) > 1/2$, i.e. for quasilinear terms that lose more than half a derivative.
In [HITW15], Hunter, Ifrim, Tataru and Wong proved the existence of solutions with a cubic lifespan for a non-dispersive equation with a quasilinear quadratic term that loses one derivative. They further developed the idea behind normal form transforms by using a modified energy in order to circumvent the non-invertibility of their normal form transformation.
Motivated by this, we showed the existence of long time solutions for a quasilinear dispersive equation with resonances in [DH18]. Further, we proved a NLS approximation theorem for this quasilinear dispersive equation in [DH18] by using a similar modified energy. In [D17] the NLS approximation was justified for a quasilinear dispersive system with $\deg^*(\rho) = 1$ by using a modified energy. Cummings and Wayne also improved the result of [SW11] in [CW17] by using a modified energy. The two systems looked at in [D17] and [CW17] are systems of the form (1.2) with $\deg^*(\rho) \leq 1$ and $\deg(\omega) = \deg^*(\rho)$ that directly fall into the class of systems that we consider in this thesis, however this work is much closer in spirit to [D17, DH18].

In this thesis the NLS approximation is now also justified for quasilinear dispersive systems with arbitrarily large $\deg^*(\rho)$, i.e. for dispersive systems with a quadratic term that loses an arbitrary amount of derivatives. This is particularly the first time a NLS approximation theorem is proven for $\deg^*(\rho) > 1$. The case $\deg^*(\rho) > 1$ is more difficult than the case $\deg^*(\rho) \leq 1$ due to the fact that a new additional class of problematic terms arises in the error estimates. The works [SW11, D17, CW17] consider the situation $\deg^*(\rho) \leq 1$ with $\omega = \rho$ or $\deg(\omega) = \deg^*(\rho)$ while this thesis makes do with the lighter restriction $\deg(\omega) \geq \deg^*(\rho)$. This thesis is further distinguished from the above mentioned works in that our NLS approximation theorem does not only hold true for one particular quasilinear dispersive system but for a whole class of dispersive systems. Due to the generality of the obtained result, our framework and techniques should be easily extendable to systems with more complicated nonlinear terms. What we in particular also will show is that our techniques can be useful for the justification of the NLS-approximation for the 2-D water wave problem with finite depth and surface tension.

1.1.3 Example systems

An important example of a system that suffices our conditions, i.e. for that theorem 1.1.1 applies, is the system

$$\begin{aligned} \partial_t u &= -\, i\omega v\,, \\ \partial_t v &= -\, i\omega u - i\rho u^2 \end{aligned} \tag{1.24}$$

where ω is given in Fourier space by the function

$$\omega(k) = \operatorname{sign}(k)\sqrt{(k + bk^3)\tanh(k)}$$

and ρ either by the function

$$\rho(k) = \rho_1(k) = \operatorname{sign}(k)\sqrt{k\tanh(k)} + bk|k|^{1/2}$$

or

$$\rho(k) = \rho_2(k) = \operatorname{sign}(k)\sqrt{k\tanh(k)} + bk.$$

These are model problems for the 2D water wave problem with finite depth and a surface tension proportional to $b \geq 0$. The systems have the same linear dispersion relation as the 2D water wave problem, i.e.

$$\big(\omega(k)\big)^2 - (k + bk^3)\tanh(k) = 0.$$

Therefore they also share the difficulty of a trivial resonance at $k = 0$ and a non-trivial resonance at $k = k_0$ with the water wave problem. The pseudo differential operator ρ was chosen such that the quasilinear quadratic terms of (1.24) pose similar difficulties as the ones of the 2D water wave problem.

It has to be mentioned that for some combinations of $0 < b < 1/3$ and $k_0 > 0$ there are additional resonances happening. So depending on b our theorem cannot always be applied for all wavenumbers $k_0 > 0$.
As a model problem for the 2D water wave problem in the case of no surface tension, i.e. for $b = 0$, this system was already successfully studied in [SW11] and in [CW17].

Another interesting equation for that the validity of the NLS approximation can be shown with theorem 1.1.1 is the nonlinear beam equation

$$\partial_t^2 u = -\partial_x^4 u - \partial_x^4 u^2 \,, \tag{1.25}$$

which is equivalent to the first order system

$$\begin{aligned} \partial_t u &= -\,\partial_x^2 \mathcal{H} v \,, \\ \partial_t v &= -\,\partial_x^2 \mathcal{H} u - \partial_x^2 \mathcal{H} u^2 \,, \end{aligned}$$

where $\mathcal{H}$ is the so-called Hilbert transformation that is given in Fourier space by the symbol $\hat{\mathcal{H}}(k) = -i\,\mathrm{sign}(k)$. One could may also call this equation a double dispersion equation. Beam equations usually model the deformations of an elastic beam, while double dispersion equations can appear for surface waves in shallow water, in the dislocation theory of crystals or the interaction between waves guides and some external medium. There exists various results for both kind of equations, see for example [LG19, KV19, WC06]. However, the above fully quasilinear case seems so far to be avoided due to the difficulties arising from such a nonlinearity, cf. introduction of [LG19]. We could also not find an article, where the NLS equation is justified. This could have something to do with the fact that there are always nontrivial resonances occurring in $\pm k_0$. The above equation may also be of relevance for models with water under a thick ice cover, where a similar dispersion relation can occur, cf. [I15].
Using theorem 1.1.1 the NLS approximation is now justified for the above equation for all $k_0 > 0$. One can easily directly check all conditions, only for (1.12) a case analysis and the quadratic formula are needed.

Only as an example to underline the fact that theorem 1.1.1 allows arbitrarily large $\deg^*(\rho)$, i.e. an arbitrary amount of derivatives falling on the quadratic term, we give the equation

$$\partial_t^2 u = -\omega^2 u - \partial_x^{100} u^2 \,, \tag{1.26}$$

with some suitable ω satisfying $\deg(\omega) \geq 50$.
When an k_0 satisfying the conditions of theorem 1.1.1 can be found, the NLS approximation is valid for such an equation.

1.2 A reduced system for the water wave problem

While the result of this chapter also stands for itself, it at the same time works as an example of how the techniques acquired from proving theorem 1.1.1 can be applied for more general systems.
A goal of this thesis was to develop techniques that can be used for the justification of the NLS approximation for the water wave problem. The (2-D) water wave problem is the problem of finding the irrational flow of an incompressible fluid in an infinitely long canal with flat bottom and a free surface under the influence of gravity. For more information about the water wave problem we refer to [D18] and the references therein.

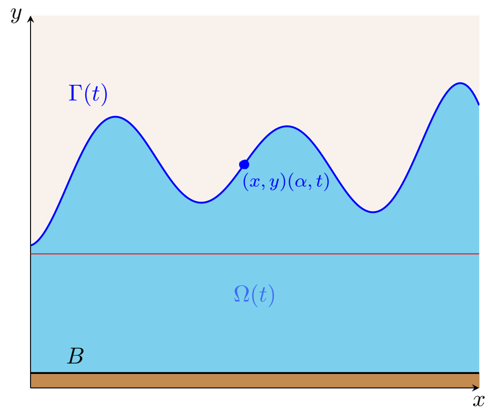


Figure 1.4: 2-D water wave problem with finite depth. $\Gamma(t)$ is the free surface, B is the bottom.

Zakharov non-rigorously derived the NLS equation as an amplitude equation for the water wave problem in [Z68]. Quite some time passed until the NLS-approximation for the 2-D water wave problem was rigorously justified on the right time scale by Totz and Wu in the case of zero surface tension and infinite depth in [TW12] and by Düll, Schneider and Wayne in the case of zero surface tension and finite depth in [DSW16]. In this thesis, we will now present techniques that can also be used to justify the NLS-approximation for the 2-D water wave problem in case of surface tension and finite depth. Without neglecting surface tension the water wave problem seemed to get way more complicated, so the case of nonzero surface tension was until recently still a open problem. In [SSZ15] it was shown that for weak surface tension the NLS approximation can even fail in some scenarios. Very recently, the NLS-approximation was justified for the 2-D water wave problem in case of finite depth and possibly of surface tension in [D19] (as long as there are no additional nontrivial resonances or k_0 is stable). The error estimates in [D19] are even uniform with respect to the strength of surface tension as the height of the wave packet and the surface tension tend to zero.

We here heuristically derive a system from the arc length formulation of the 2-D water wave problem with finite depth and possibly of surface tension. This reduced system is the system

$$\begin{aligned}\partial_t u_{-1} = -i\omega u_{-1} + \partial_\alpha\Big(&- D_\alpha^{-2}(u_{-1}+u_1)u_{-1} \\ &- \frac{1}{2}[\sigma, D_\alpha^{-2}(u_{-1}+u_1)]\sigma^{-1}(u_{-1}-u_1) \\ &+ \frac{1}{2}K_0 D_\alpha^{-1}\sigma^{-1}(u_{-1}-u_1)\sigma^{-1}(u_{-1}-u_1) \\ &- \frac{1}{2}b\sigma^{-1}(u_{-1}-u_1)K_0\sigma^{-1}\partial_\alpha(u_{-1}-u_1) \\ &- \frac{1}{2}(D_\alpha^{-1}(u_{-1}+u_1))^2 + \frac{1}{2}(K_0 D_\alpha^{-1}(u_{-1}+u_1))^2\Big),\end{aligned} \tag{1.27}$$

$$\begin{aligned}\partial_t u_1 = i\omega u_1 + \partial_\alpha\Big(&- D_\alpha^{-2}(u_{-1}+u_1)u_1 \\ &+ \frac{1}{2}[\sigma, D_\alpha^{-2}(u_{-1}+u_1)]\sigma^{-1}(u_{-1}-u_1) \\ &+ \frac{1}{2}K_0 D_\alpha^{-1}\sigma^{-1}(u_{-1}-u_1)\sigma^{-1}(u_{-1}-u_1) \\ &- \frac{1}{2}b\sigma^{-1}(u_{-1}-u_1)K_0\sigma^{-1}\partial_\alpha(u_{-1}-u_1) \\ &- \frac{1}{2}(D_\alpha^{-1}(u_{-1}+u_1))^2 + \frac{1}{2}(K_0 D_\alpha^{-1}(u_{-1}+u_1))^2\Big),\end{aligned}$$

where the linear operator $i\omega$ is given in Fourier space by its symbol

$$\omega(k) = \omega(k; b) = \mathrm{sign}(k)\sqrt{(k + bk^3)\tanh(k)} \tag{1.28}$$

and $b \geq 0$ is the Bond number that is proportional to the strength of surface tension, i.e. $b > 0$ would mean nonzero surface tension.
The operator σ is defined in Fourier space by its symbol

$$\sigma(k) = \sigma(k; b) = \sqrt{\frac{k + bk^3}{\tanh(k)}}\,, \tag{1.29}$$

K_0 by its symbol $K_0(k) = -i\tanh(k)$ and the operator σ^{-1} by its symbol $\sigma^{-1}(k)$. The operator $D_\alpha^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is given in Fourier space by some fixed function $\hat{D}_\alpha^{-1}$, which is smooth, odd and fulfills $\hat{D}_\alpha^{-1}(k) = \mathcal{O}(-ik^{-1})$ for $|k| \to \infty$. In order to avoid resonances being caused by the nonlinear terms, we did not use $D_\alpha^{-1} = \partial_\alpha^{-1}$.

This system has the same dispersion relation as the full water wave problem. Furthermore the quadratic terms of this system can be compared to the ones of

the diagonalized water wave problem in the arc length formulation. We call it a reduced system since we can heuristically derive it by simplifying a system that describes the full water wave problem, see section 3.1.
The nonlinear terms of the reduced system have certain properties, which we will call key properties in the following. These key properties can also be partially found in the full water wave problem. We will prove our error estimates by only relying on these key properties such that our result extends to every dispersive system that shares them.
These key properties are as follows.

The reduced system is an abstract system of the form

$$\partial_t u_{-1} = -i\omega u_{-1} + \mathcal{A}_{-1}(u_{-1}, u_{-1}) + \mathcal{B}_{-1}(u_1, u_1) + \mathcal{C}_{-1}(u_{-1}, u_1) \tag{1.30}$$

$$\partial_t u_1 = i\omega u_1 + \mathcal{A}_1(u_{-1}, u_{-1}) + \mathcal{B}_1(u_1, u_1) + \mathcal{C}_1(u_{-1}, u_1)\,, \tag{1.31}$$

where the linear operator $i\omega$ is given exactly like before. The quadratic terms are given in Fourier space by

$$\widehat{\mathcal{A}}_j(u_{-1}, u_{-1})(k) := \int_{\mathbb{R}} a_j(k, k-m, m)\, \widehat{u}_{-1}(k-m)\widehat{u}_{-1}(m)\, dm\,, \tag{1.32}$$

$$\widehat{\mathcal{B}}_j(u_1, u_1)(k) := \int_{\mathbb{R}} b_j(k, k-m, m)\, \widehat{u}_1(k-m)\widehat{u}_1(m)\, dm\,, \tag{1.33}$$

$$\widehat{\mathcal{C}}_j(u_{-1}, u_1)(k) := \int_{\mathbb{R}} c_j(k, k-m, m)\, \widehat{u}_{-1}(k-m)\widehat{u}_1(m)\, dm\,, \tag{1.34}$$

where $j \in \{\pm 1\}$ and the functions a_j, b_j and c_j are sufficiently smooth.
For $\mathcal{Z} \in \{\mathcal{A}_{-1}, \mathcal{A}_1, \mathcal{B}_{-1}, \mathcal{B}_1, \mathcal{C}_{-1}, \mathcal{C}_1\}$ and accordingly chosen $z \in \{a_{-1}, a_1, b_{-1}, b_1, c_{-1}, c_1\}$, we have

$$z(k, k-m, m) = \mathcal{O}(k) \qquad\qquad \text{for } |k| \to 0\,, \tag{1.35}$$

as long as $|k-m|$ gets uniformly bounded.
The $z(k, k-m, m)$ suffice the conditions of lemma 3.3.10.
Moreover, the operators $\mathcal{Z}$ always map a pair of real-valued functions on a real-valued function and satisfy a priori estimates of the form

$$\|\mathcal{Z}(u, v)\|_{L^2} \lesssim \begin{cases} \|u\|_{H^2}\|v\|_{H^1}\,, \\ \|u\|_{H^1}\|v\|_{H^2}\,, \\ \|\widehat{u}\|_{L^1(4)}\|v\|_{H^1}, \\ \|u\|_{H^1}\|\widehat{v}\|_{L^1(4)} \end{cases}. \tag{1.36}$$

On top of that, we have the a priori estimates

$$
\begin{aligned}
\|\mathcal{A}_{-1,s}(f,g)+\mathcal{A}^*_{-1,s}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\
\|\mathcal{B}_{1,s}(f,g)+\mathcal{B}^*_{1,s}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\
\|\mathcal{C}_{-1}(g,f)+\mathcal{C}_{-1,*}(g,f)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\
\|\mathcal{C}_{1}(f,g)+\mathcal{C}^*_{1}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,,
\end{aligned}
\tag{1.37}
$$

and

$$
\begin{aligned}
\|\mathcal{A}^*_{1,s}(f,g)+\mathcal{C}_{-1}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{H^{1/2}}\,, \\
\|\mathcal{B}_{-1,s}(f,g)+\mathcal{C}_{1,*}(g,f)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{H^{1/2}}\,.
\end{aligned}
\tag{1.38}
$$

Here, we are using the notations

$$\mathcal{Z}_s(f,\cdot) := \mathcal{Z}(f,\cdot)+\mathcal{Z}(\cdot,f), \tag{1.39}$$

$$
\begin{aligned}
\int_{\mathbb{R}} \mathcal{Z}^*(g,f)\,h\,dx &:= \int_{\mathbb{R}} f\,\mathcal{Z}(g,h)\,dx\,, \\
\int_{\mathbb{R}} \mathcal{Z}_*(g,f)\,h\,dx &:= \int_{\mathbb{R}} g\,\mathcal{Z}(h,f)\,dx\,.
\end{aligned}
\tag{1.40}
$$

In subsection 3.2.1, we give some more detailed information on these key properties and also explain how the conditions (1.37) and (1.38) can be understood.

We assume the local existence of real-valued solutions to our system (1.2) in H^{s_A} with s_A as in theorem 1.2.1. In [A03], well-posedness of water waves with surface tension has been shown.
We chose $k_0>0$ such that (1.4), i.e.

$$\omega''(k_0)\neq 0, \tag{1.41}$$

(1.5), i.e.

$$\omega'(k_0)\neq \pm\omega'(0) \tag{1.42}$$

and (1.6), i.e.

$$m\omega(k_0)\neq \pm\omega(mk_0) \qquad\qquad \text{for } m=\pm 2,\dots,\pm 5\,, \tag{1.43}$$

(1.14), i.e.

$$\omega(k_0)\neq 0 \qquad\qquad \text{when } \deg(\omega)<1\,, \tag{1.44}$$

are true. Moreover, we chose $k_0>0$ such that for $j_1,j_2\in\{\pm 1\}$ the only possible solutions of the equations (1.12), i.e.

$$\omega(k)-j_1j_2\omega(k\mp k_0)+j_1\omega(\pm k_0)=0\,, \tag{1.45}$$

are $k = \pm k_0$ and $k = 0$.
Solutions of (1.45) will correspond to resonances in our normal form transforms. In literature it has been shown that when $b = 0$ or $b > 1/3$ there can always only occur resonances in $k = \pm k_0$ and $k = 0$ for all $k_0 > 0$. When $b \in]0, 1/3[$ there can occur more than three resonances for some $k_0 > 0$.
We will assume that we can rigorously derive the Nonlinear Schrödinger equation

$$\partial_T A = i\, \frac{\omega''(k_0)}{2}\, \partial_X^2 A + i\nu_2(k_0) A|A|^2\,, \tag{1.46}$$

with some $\nu_2(k_0) \in \mathbb{R}$, via an ansatz of the form

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \varepsilon \psi_{NLS} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2)\,.$$

Here

$$\varepsilon \psi_{NLS}(x,t) = \varepsilon A\big(\varepsilon(x - c_g t), \varepsilon^2 t\big) e^{i(k_0 x - \omega_0 t)} + c.c. \tag{1.47}$$

is the Nonlinear Schrödinger approximation for solutions of (1.27).
The complex-valued amplitude A is the solution of the NLS equation (1.46) and $0 < \varepsilon \ll 1$ is a small perturbation parameter. The basic temporal wave number $\omega_0 := \omega(k_0)$ of the underlying carrier wave $e^{i(k_0 x - \omega_0 t)}$ is associated to the basic spatial wave number $k_0 > 0$. The group velocity of the wave packet is $c_g := \omega'(k_0)$ and *c.c.* simply denotes the complex conjugate.
Additionally, we assume that we can derive the Nonlinear Schrödinger equation in a similar way as for (1.2), we further specify this in section 3.2. We disregarding the derivation of the NLS equation here, since it is already known that it can be derived for the water wave problem. The only reason we are looking at the reduced system is to get an idea how the NLS equation can be justified for the water wave problem with surface tension.

Under these assumptions, we obtain the following result.

Theorem 1.2.1. *Let $b > 0$. Fix $k_0 > 0$ as above and $s_A \geq 7$. For all $C_1, T_0 > 0$ there exists $\varepsilon_0 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ of the NLS equation* (1.46) *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

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} \in C\big([0, T_0/\varepsilon^2], H^{s_A}(\mathbb{R}, \mathbb{R})\big)$$

of (1.27) *which satisfy*

$$\sup_{t \in [0, T_0/\varepsilon^2]} \Big\| \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix}(\cdot, t) - \varepsilon \psi_{NLS}(\cdot, t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Big\|_{H^{s_A}(\mathbb{R}, \mathbb{R})} \lesssim \varepsilon^{3/2}.$$

More interesting than the theorem itself is the fact that the above key properties suffice to prove it, leaving aside the other assumptions we made. In other words the result extends to a whole class of systems whose nonlinearities have this certain form. The structure provided by these key properties can also be found in the arc length formulation of the water wave problem with finite depth and surface tension, although one cannot directly embed the full water wave problem into our setting since it is more complicated. Nevertheless we get a good idea of how the loss of regularity occurring in the error estimates of the full water wave problem could be approached.

We follow [D19] to first present the Eulerian formulation and then the arc length formulation of the 2D water wave problem in the case of finite depth and possibly surface tension. Then we heuristically derive the reduced system (1.27) from the arc length formulation of the water wave problem and proceed to talk about the key properties of the reduced system. In section 3.3, we then make error estimates by applying the techniques from section 2.2. We first construct the normal form transformations and a modified energy. Then we show the equivalence of this energy $\mathcal{E}_{s_A}$ to the H^{s_A}-norm of the error and

$$\partial_t \mathcal{E}_{s_A} = \mathcal{O}(\varepsilon^2)\,.$$

Finally, we improve the energy a little to obtain

$$\partial_t \tilde{\mathcal{E}}_{s_A} = \varepsilon^2\, \mathcal{O}(\tilde{\mathcal{E}}_{s_A} + 1)$$

such that theorem 1.2.1 follows.

1.3 General Notation

For functions $f : \mathbb{R} \to \mathbb{K}$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we use the norms

$$\|f\|_{L^2} := \sqrt{\int_{\mathbb{R}} |f(x)|^2 \, dx} \, ,$$

$$\|f\|_{\infty} := \operatorname*{ess\,sup}_{x \in \mathbb{R}} |f(x)| \, ,$$

$$\|f\|_{C^n} := \sum_{j=0}^{n} \max_{x \in \mathbb{R}} |\partial_x^j f(x)| \, ,$$

$$\|f\|_{H^s} := \|(1 + |\cdot|^2)^{s/2} \widehat{f}(\cdot)\|_{L^2} = \sqrt{\int_{\mathbb{R}} (1 + |k|^2)^s |\widehat{f}(k)|^2 \, dk} \, ,$$

$$\|\widehat{f}\|_{L^1(s)} := \int_{\mathbb{R}} (1 + |k|^2)^{s/2} |\widehat{f}(k)| \, dk \, .$$

Here $\widehat{f}$ denotes the Fourier transformation of f.
We choose the Fourier transformation $\mathcal{F}$ that is defined by

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

for suitable functions f.
We write $f \in L^2(\mathbb{R}, \mathbb{K})$ when $\|f\|_{L^2} < \infty$, $f \in H^s(\mathbb{R}, \mathbb{K})$ when $\|f\|_{H^s} < \infty$, $\widehat{f} \in L^1(s)(\mathbb{R}, \mathbb{K})$ when $\|\widehat{f}\|_{L^1(s)} < \infty$ and $f \in C_b^n(\mathbb{R}, \mathbb{K})$ when $\|f\|_{C^n} < \infty$ for $f : \mathbb{R} \to \mathbb{K}$. We write $f \in C_c^\infty(\mathbb{R}, \mathbb{K})$, when f has a compact support and $f \in C_b^n$ for all $n \in \mathbb{N}$. Sometimes, we write for example $f \in H^s(\mathbb{R})$ instead of $f \in H^s(\mathbb{R}, \mathbb{C})$.
For $f_1, f_2 \in H^s(\mathbb{R})$, we a few times use

$$\| \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \|_{H^s} = \|f_1\|_{H^s} + \|f_2\|_{H^s} \, .$$

We have the estimate $\|f\|_{C^n} \leq C \|\widehat{f}\|_{L^1(n)}$ for some $C \geq 0$.
For $\psi(x) = A(\varepsilon x) e^{ik_0 x}$, we have $\widehat{\psi}(k) = \varepsilon^{-1} \widehat{A}\big((k - k_0)\varepsilon^{-1}\big)$. Thus, we only have $\|\psi\|_{H^s} \leq \varepsilon^{-1/2} C \|A\|_{H^s}$ but $\|\psi\|_{C^n} \leq C \|A\|_{C^n}$ and $\|\widehat{\psi}\|_{L^1(s)} \leq C \|\widehat{A}\|_{L^1(s)}$.

We write $\deg^*(\gamma) \leq s$ for a function $\gamma : \mathbb{R} \to \mathbb{R}$ when there are some constants C, M such that

$$|\gamma(k)| \leq C(1 + |k|)^s \quad \text{for} \quad |k| \geq M \, .$$

We write $\deg(\gamma) = s$ when there on top of that is some $c > 0$ such that

$$c(1 + |k|)^s \leq |\gamma(k)| \leq C(1 + |k|)^s \quad \text{for} \quad |k| \geq M.$$

We write $\deg^*(\gamma) = s$, when s is the minimal s for that $\deg^*(\gamma) \leq s$ is true.

For expressions I and E, we often write

$$I \leq \mathcal{O}(E)\,,$$

when we want to express that there exists some constant $C > 0$ such that

$$I \leq C\,E\,.$$

The constant C can then always be chosen independently of E and the small perturbation parameter ε.
A few times, we write $I \lesssim E$ instead of $I \leq \mathcal{O}(E)$. We sometimes write $I = \mathcal{O}(E)$, when we want to express that $I \leq \mathcal{O}(E)$ and $-I \leq \mathcal{O}(E)$.

For an operator γ and some functions g, f, we denote the commutator $[\gamma, f]g$ by

$$[\gamma, f]g := \gamma(fg) - f\,\gamma g\,.$$

For convenience, we often call the operators defining a normal form transformation also normal form transformations.

Chapter 2

Quasilinear dispersive systems

2.1 The Derivation of the NLS approximation

In this section, we will first show how the NLS equation (1.7) can be derived for the dispersive system (1.2). Then, we will prove residual estimates for a improved NLS approximation Ψ, which only differs slightly from ψ_{NLS}. For the derivation of the NLS equation and for all estimates in this section, we only need the conditions (1.4), (1.5) and (1.6) to be fulfilled and that the function ω or $\operatorname{sign}(\cdot)\omega(\cdot)$, and, the function ρ or $\operatorname{sign}(\cdot)\rho(\cdot)$, lie in $C^5(\mathbb{R})$.

Before we derive the NLS-equation, we diagonalize our dispersive system

$$\begin{aligned} \partial_t u &= -\, i\omega v\,, \\ \partial_t v &= -\, i\omega u - i\rho u^2 \end{aligned}$$

via the transformation

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}, \tag{2.1}$$

which we later could invert again by the transformation

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix}. \tag{2.2}$$

We obtain the diagonalized system

$$\begin{aligned} \partial_t u_{-1}(x,t) &= -i\omega u_{-1}(x,t) - \frac{1}{2} i\rho\big(u_{-1}+u_1\big)^2(x,t)\,, \\ \partial_t u_1(x,t) &= i\omega u_1(x,t) + \frac{1}{2} i\rho\big(u_{-1}+u_1\big)^2(x,t)\,, \end{aligned} \tag{2.3}$$

where $t, x \in \mathbb{R}$ and $u_{-1}(x,t), u_1(x,t) \in \mathbb{R}$.

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

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \varepsilon\Psi_S := \varepsilon \left(A_1(X,T)\mathbf{E} + \overline{A_1}(X,T)\mathbf{E}^{-1} \right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.4}$$
$$+ \varepsilon^2 \begin{pmatrix} A_0(X,T) + A_2(X,T)\mathbf{E}^2 + \overline{A_2}(X,T)\mathbf{E}^{-2} \\ D_0(X,T) + D_2(X,T)\mathbf{E}^2 + \overline{D_2}(X,T)\mathbf{E}^{-2} \end{pmatrix},$$

where $X := \varepsilon(x - c_g t)$, $T := \varepsilon^2 t$, $\mathbf{E} = e^{i(k_0 x - \omega_0 t)}$, $\omega_0 = \omega(k_0)$ and $c_g = \omega'(k_0)$. This is an ansatz that leads to an approximation that describes waves moving to the right with the group velocity c_g.

Remark 2.1.1. *In order to obtain an approximation that describes waves moving to the left with the group velocity c_g, one could replace in the above ansatz the vector $(1,0)^T$ by $(0,1)^T$ as well as $-\omega_0$ by ω_0 and c_g by $-c_g$ (cf. [SW11]), or just replace the operators ω and ρ by the operators $\tilde{\omega} = -\omega$ and $\tilde{\rho} = -\rho$.*

We insert the ansatz $\varepsilon\Psi_S$ into the diagonalized system (2.3).
In order to be able to directly compare powers in ε, we use Taylors theorem to expand all resulting terms of the forms $\omega\big[A_j\mathbf{E}^j\big]$, $\omega\big[D_j\mathbf{E}^j\big]$, $\rho\big[A_{j_1}A_{j_2}\mathbf{E}^{j_1+j_2}\big]$, $\rho\big[D_{j_1}A_{j_2}\mathbf{E}^{j_1+j_2}\big]$ and $\rho\big[D_{j_1}D_{j_2}\mathbf{E}^{j_1+j_2}\big]$ (cf. Lemma 25 of [SW11]).
When $\omega \in C^5(\mathbb{R})$, i.e. in the case $\omega(0) = 0$, Taylor's theorem yields

$$\omega(k) = \omega(jk_0) + \omega'(jk_0)\,(k - jk_0) + \sum_{n=2}^{4} \frac{\omega^{(n)}(jk_0)}{n!}(k - jk_0)^n + \mathcal{O}\big((k - jk_0)^5\big),$$

and we therefore have the expansion

$$i\omega\big[A_j\mathbf{E}^j\big] = i\omega(jk_0)\,A_j\mathbf{E}^j + \varepsilon\,\omega'(jk_0)\,\partial_X A_j\mathbf{E}^j + \sum_{n=2}^{4} \varepsilon^n \frac{i(-i)^n\omega^{(n)}(jk_0)}{n!}\partial_X^n A_j\mathbf{E}^j + \mathcal{O}(\varepsilon^5).$$

Analogously, we expand expressions involving the operator ρ when $\rho(0) = 0$, i.e. when $\rho \in C^5(\mathbb{R})$.
When $\lim_{k\to\pm 0}\omega(k) \neq 0$ or $\lim_{k\to\pm 0}\rho(k) \neq 0$, we have to expand more carefully since the function ω or ρ has a jump in $k = 0$.
In the case $\lim_{k\to\pm 0}\omega(k) \neq 0$, we write

$$\omega(k) = i\widehat{\mathcal{H}}(k)\,v(k) = \mathrm{sign}(k)v(k),$$

and then expand the function $v := \mathrm{sign}(\cdot)\omega(\cdot) \in C^5(\mathbb{R})$ with Taylor's theorem in order to get

$$\omega(k) = \mathrm{sign}(k)\Big(\mathrm{sign}(jk_0)\omega(jk_0) + \mathrm{sign}(jk_0)\omega'(jk_0)\,(k - jk_0) + \sum_{n=2}^{4} \frac{\mathrm{sign}(jk_0)\omega^{(n)}(jk_0)}{n!}(k - jk_0)^n + \mathcal{O}\big((k - jk_0)^5\big)\Big).$$

If the support $\mathcal{S}_j$ of $A_j\mathbf{E}^j$ in Fourier space is strictly restricted to a small enough neighborhood of jk_0, we have

$$\operatorname{sign}(k)\operatorname{sign}(jk_0) = 1 \qquad \text{for } k \in \mathcal{S}_j$$

and thus can still use the expansion

$$\begin{aligned} i\omega\big[A_j\mathbf{E}^j\big] &= i\omega(jk_0)\, A_j\mathbf{E}^j + \varepsilon\, \omega'(jk_0)\, \partial_X A_j\mathbf{E}^j \\ &\quad + \sum_{n=2}^{4} \varepsilon^n \frac{i(-i)^n\omega^{(n)}(jk_0)}{n!} \partial_X^n A_j\mathbf{E}^j + \mathcal{O}(\varepsilon^5). \end{aligned}$$

for $j \neq 0$. However, for $j = 0$, we obtain

$$i\omega A_j = -\omega(0^+)\, \mathcal{H}A_j + \varepsilon^2 \frac{\omega^{(2)}(0^+)}{2}\, \mathcal{H}\partial_X^2 A_j + \varepsilon^4 \frac{\omega^{(4)}(0^+)}{4!}\, \mathcal{H}\partial_X^4 A_j + \mathcal{O}(\varepsilon^5).$$

Note that $\omega^{(n)}(0^+) = 0$ for odd numbers n, what simply is reflecting the fact that the function $v = \operatorname{sign}(\cdot)\omega(\cdot) \in C^5(\mathbb{R})$ is even.
Analogously, we expand expressions involving ρ in the case $\lim_{k\to\pm 0} \rho(k) \neq 0$.
After having expanded all expression like this, we now equate the coefficients in front of $\varepsilon^m\mathbf{E}^j$ to zero. Since ε is really small, we expect the terms with higher ε-powers to be smaller and start by looking at the terms with lowest ε-power.
Due to the different expansions being valid, we here have to distinguish between the two cases $\omega(0) = 0$ and $\lim_{k\to\pm 0} \omega(k) \neq 0$.

2.1.1 Derivation in the case where $\omega(0) = 0$

In the case where $\omega(0) = 0$, due to (1.5), we have

$$\omega'(k_0) \neq \pm\omega'(0) \quad \text{and} \quad \rho(0) = 0$$

on top of having (1.4) and (1.6):

$$\omega''(k_0) \neq 0,$$

$$m\omega(k_0) \neq \pm\omega(mk_0) \qquad\qquad \text{for } m = \pm 2, \ldots, \pm 5\,.$$

We obtain the following equations for the coefficients in front of $\varepsilon\mathbf{E}^j$:

$$\begin{aligned} &\varepsilon\mathbf{E}\begin{pmatrix}1\\0\end{pmatrix}: &&\omega(k_0) - \omega_0 = 0\,, \\ &\varepsilon\mathbf{E}^{-1}\begin{pmatrix}1\\0\end{pmatrix}: &&\omega(k_0) - \omega_0 = 0\,. \end{aligned}$$

Thus, since we have chosen $\omega_0 = \omega(k_0)$, all terms of order ε cancel.
In front of $\varepsilon^2 \mathbf{E}^j$, we have the following equations for the coefficients:

$$
\begin{aligned}
&\varepsilon^2 \mathbf{E}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} : && \frac{1}{2}\rho(2k_0)A_1^2 + \big(\omega(2k_0) - 2\omega_0\big)A_2 = 0\,, \\
&\varepsilon^2 \mathbf{E} \begin{pmatrix} 1 \\ 0 \end{pmatrix} : && \omega'(k_0) - c_g = 0\,, \\
&\varepsilon^2 \mathbf{E}^0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} : && \rho(0) = 0\,, \\
&\varepsilon^2 \mathbf{E}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} : && \omega'(-k_0) - c_g = 0\,, \\
&\varepsilon^2 \mathbf{E}^{-2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} : && \frac{1}{2}\rho(-2k_0)A_1^2 + \big(\omega(-2k_0) + 2\omega_0\big)A_2 = 0\,, \\
\\
&\varepsilon^2 \mathbf{E}^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} : && \frac{1}{2}\rho(2k_0)A_1^2 + \big(\omega(2k_0) + 2\omega_0\big)D_2 = 0\,, \\
&\varepsilon^2 \mathbf{E}^0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} : && \rho(0) = 0\,, \\
&\varepsilon^2 \mathbf{E}^{-2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} : && \frac{1}{2}\rho(-2k_0)A_1^2 + \big(\omega(-2k_0) - 2\omega_0\big)D_2 = 0\,.
\end{aligned}
$$

Due to (1.6), we can choose

$$
\begin{aligned}
A_2 &= -\frac{\rho(2k_0)}{2\big(\omega(2k_0) - 2\omega_0\big)}A_1^2, \\
D_2 &= -\frac{\rho(2k_0)}{2\big(\omega(2k_0) + 2\omega_0\big)}A_1^2.
\end{aligned}
$$

By this choice, all terms of order ε^2 cancel, since the functions ω and ρ are odd, $c_g = \omega'(k_0)$ and $\rho(0) = 0$.

In front of $\varepsilon^3\mathbf{E}^j$, we get the following equations:

$$\varepsilon^3\mathbf{E}\begin{pmatrix}1\\0\end{pmatrix}:\ \partial_T A_1 - i\frac{\omega''(k_0)}{2}\partial_X^2 A_1 + i\rho(k_0)\Big(A_1\big(A_0+D_0\big)+\overline{A_1}\big(A_2+D_2\big)\Big)=0\,,$$

$$\varepsilon^3\mathbf{E}^0\begin{pmatrix}1\\0\end{pmatrix}:\ \rho'(0)\partial_X|A_1|^2+(\omega'(0)-c_g)\partial_X A_0=0\,,$$

$$\varepsilon^3\mathbf{E}^{-1}\begin{pmatrix}1\\0\end{pmatrix}:$$

$$\partial_T\overline{A_1} - i\frac{\omega''(-k_0)}{2}\partial_X^2\overline{A_1} + i\rho(-k_0)\Big(\overline{A_1}\big(A_0+D_0\big)+A_1\big(\overline{A_2}+\overline{D_2}\big)\Big)=0\,,$$

$$\varepsilon^3\mathbf{E}^0\begin{pmatrix}0\\1\end{pmatrix}:\ \rho'(0)\partial_X|A_1|^2+(\omega'(0)+c_g)\partial_X D_0=0\,.$$

We also have coefficients in front of $\varepsilon^3\mathbf{E}^j$ for $j\in\{-3,-2,2,3\}$ and in front of $\varepsilon^3\mathbf{E}\,(0,\,1)^T$ and $\varepsilon^3\mathbf{E}^{-1}\,(0,\,1)^T$, but, unlike the coefficients above, we can always get rid of these coefficients by simply extending our ansatz by some $\mathcal{O}(\varepsilon^2)$-terms and exploiting (1.6).
Due to (1.5), we can choose

$$A_0=-\frac{\rho'(0)}{\omega'(0)-c_g}|A_1|^2,$$
$$D_0=-\frac{\rho'(0)}{\omega'(0)+c_g}|A_1|^2,$$

such that the coefficients in front of $\varepsilon^3\mathbf{E}^0\,(1,\,0)^T$ and $\varepsilon^3\mathbf{E}^0\,(1,\,0)^T$ vanish.
When we now plug in our choices for A_2, D_2, A_0 and D_0, we obtain, in the coefficients in front of $\varepsilon^3\mathbf{E}\,(1,\,0)^T$ or $\varepsilon^3\mathbf{E}^{-1}\,(1,\,0)^T$, the NLS-equation

$$\partial_T A_1 = i\,\frac{\omega''(k_0)}{2}\,\partial_X^2 A_1 + i\nu_2(k_0)A_1|A_1|^2\,,$$

with

$$\nu_2(k_0)=-\rho(k_0)\left(\frac{2\rho'(0)\,\omega'(0)}{c_g^2-\big(\omega'(0)\big)^2}+\frac{\rho(2k_0)\,\omega(2k_0)}{4\omega_0^2-\big(\omega(2k_0)\big)^2}\right).\tag{2.5}$$

This is how the NLS equation is derived in the case where $\omega(0)=0$.

2.1.2 Derivation in the case where $\lim_{k\to\pm 0}\omega(k)\neq 0$

We now derive the NLS-equation in the case where $\lim_{k\to\pm 0}\omega(k)\neq 0$, that obviously is the case where we have

$$\lim_{k\to\pm 0}\omega(k)\neq 0$$

on top of having (1.4) and (1.6):

$$\omega''(k_0) \neq 0,$$

$$m\omega(k_0) \neq \pm\omega(mk_0) \qquad\qquad \text{for } m = \pm 2, \dots, \pm 5\,.$$

As hinted before, in order to be able to compare the coefficients in front of $\varepsilon^m \mathbf{E}^j$ in a similar way as we did in the case where $\omega(0) = 0$, we here will assume that for $j \neq 0$, the support of $A_j \mathbf{E}^j$ and $D_j \mathbf{E}^j$ in Fourier space is such strictly concentrated around the wavenumbers jk_0, that we can replace the expression $\mathrm{sign}(k)\mathrm{sign}(jk_0)$ by $1 + \mathcal{O}(\varepsilon^6)$ in our Taylor expansions. This assumption is automatically fulfilled when $A_j, D_j \in H^s$ for some large enough $s \geq 0$, due to the estimate

$$\|\chi_{[-\delta,\delta]}\varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot) - \varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot)\|_{L^2(m)} \leq C(\delta)\,\varepsilon^{m+M-1/2}\|f\|_{H^{m+M}} \tag{2.6}$$

for $f \in H^{m+M}$ and for all $M, m \geq 0$, where $\chi_{[-\delta,\delta]}$ is the characteristic function on $[-\delta, \delta]$ (see (24) in [S98b]).

We again obtain the following equations for the coefficients in front of $\varepsilon \mathbf{E}^j$:

$$\varepsilon\mathbf{E}\begin{pmatrix}1\\0\end{pmatrix}: \qquad \omega(k_0) - \omega_0 = 0\,,$$

$$\varepsilon\mathbf{E}^{-1}\begin{pmatrix}1\\0\end{pmatrix}: \qquad \omega(k_0) - \omega_0 = 0\,.$$

So, since we have chosen $\omega_0 = \omega(k_0)$, all terms of order ε cancel.
In front of $\varepsilon^2 \mathbf{E}^j$, we now have the following equations for the coefficients:

$$\varepsilon^2\mathbf{E}^2\begin{pmatrix}1\\0\end{pmatrix}: \qquad \frac{1}{2}\rho(2k_0)A_1^2 + \big(\omega(2k_0) - 2\omega_0\big)A_2 = 0\,,$$

$$\varepsilon^2\mathbf{E}\begin{pmatrix}1\\0\end{pmatrix}: \qquad \omega'(k_0) - c_g = 0\,,$$

$$\varepsilon^2\mathbf{E}^0\begin{pmatrix}1\\0\end{pmatrix}: \qquad \rho(0^+)\mathcal{H}|A_1|^2 + \omega(0^+)\mathcal{H}A_0 = 0\,,$$

$$\varepsilon^2\mathbf{E}^{-1}\begin{pmatrix}1\\0\end{pmatrix}: \qquad \omega'(-k_0) - c_g = 0\,,$$

$$\varepsilon^2\mathbf{E}^{-2}\begin{pmatrix}1\\0\end{pmatrix}: \qquad \frac{1}{2}\rho(-2k_0)A_1^2 + \big(\omega(-2k_0) + 2\omega_0\big)A_2 = 0\,,$$

$$\varepsilon^2\mathbf{E}^2\begin{pmatrix}0\\1\end{pmatrix}: \qquad \frac{1}{2}\rho(2k_0)A_1^2 + \big(\omega(2k_0) + 2\omega_0\big)D_2 = 0\,,$$

$$\varepsilon^2\mathbf{E}^0\begin{pmatrix}0\\1\end{pmatrix}: \qquad \rho(0^+)\mathcal{H}|A_1|^2 + \omega(0^+)\mathcal{H}D_0 = 0\,,$$

$$\varepsilon^2\mathbf{E}^{-2}\begin{pmatrix}0\\1\end{pmatrix}: \qquad \frac{1}{2}\rho(-2k_0)A_1^2 + \big(\omega(-2k_0) - 2\omega_0\big)D_2 = 0\,.$$

Due to (1.6), we can choose

$$
\begin{aligned}
A_2 &= -\frac{\rho(2k_0)}{2\big(\omega(2k_0)-2\omega_0\big)}A_1^2,\\
D_2 &= -\frac{\rho(2k_0)}{2\big(\omega(2k_0)+2\omega_0\big)}A_1^2,
\end{aligned}
$$

and, since $\omega(0^+)=\lim_{k\to 0^+}\omega(k)\neq 0$, we can choose

$$
\begin{aligned}
A_0 &= -\frac{\rho(0^+)}{\omega(0^+)}|A_1|^2,\\
D_0 &= -\frac{\rho(0^+)}{\omega(0^+)}|A_1|^2.
\end{aligned}
$$

By this choice all terms of order ε^2 cancel, since the functions ω and ρ are odd and $c_g=\omega'(k_0)$.
In front of $\varepsilon^3\mathbf{E}^j$, we get the equations:

$$
\varepsilon^3\mathbf{E}\begin{pmatrix}1\\0\end{pmatrix}:\ \partial_T A_1 - i\frac{\omega''(k_0)}{2}\partial_X^2 A_1 + i\rho(k_0)\Big(A_1\big(A_0+D_0\big)+\overline{A_1}\big(A_2+D_2\big)\Big)=0\,,
$$

$$
\varepsilon^3\mathbf{E}^{-1}\begin{pmatrix}1\\0\end{pmatrix}:
$$

$$
\partial_T\overline{A_1} - i\frac{\omega''(-k_0)}{2}\partial_X^2\overline{A_1} + i\rho(-k_0)\Big(\overline{A_1}\big(A_0+D_0\big)+A_1\big(\overline{A_2}+\overline{D_2}\big)\Big)=0\,.
$$

We also have coefficients in front of $\varepsilon^3\mathbf{E}^j$ for $j\in\{-3,-2,0,2,3\}$, in front of $\varepsilon^3\mathbf{E}\,(\,0,\,1)^T$ and in front of $\varepsilon^3\mathbf{E}^{-1}\,(\,0,\,1)^T$, however we can always get rid of these coefficients by adding some $\mathcal{O}(\varepsilon^2)$-terms to our ansatz and exploiting (1.6) and $\omega(0^+)\neq 0$.
When we now plug in our choices for A_2, D_2, A_0 and D_0, we obtain, in the coefficients in front of $\varepsilon^3\mathbf{E}\,(\,1,\,0)^T$ or $\varepsilon^3\mathbf{E}^{-1}\,(\,1,\,0)^T$, the NLS-equation

$$
\partial_T A_1 = i\,\frac{\omega''(k_0)}{2}\,\partial_X^2 A_1 + i\nu_2(k_0)A_1|A_1|^2\,,
$$

with

$$
\nu_2(k_0) = -\rho(k_0)\left(\frac{\rho(2k_0)\,\omega(2k_0)}{4\omega_0^2-\big(\omega(2k_0)\big)^2} - 2\,\frac{\rho(\,0^+)}{\omega(\,0^+)}\right). \tag{2.7}
$$

Remark 2.1.2. *Looking at (2.5) and (2.7), we see that the NLS equation (1.7) can be either defocusing or focusing depending on ω, ρ and $k_0>0$.*

2.1.3 Residual estimates for an improved NLS approximation

When we just derived the NLS-equation by using the ansatz $(u_{-1}, u_1)^T = \varepsilon\Psi_S$, we also showed that the residual

$$\mathrm{Res}_u(\varepsilon\Psi_S) = \begin{pmatrix} \mathrm{Res}_{u_{-1}}(\varepsilon\Psi_S) \\ \mathrm{Res}_{u_1}(\varepsilon\Psi_S) \end{pmatrix},$$

which contains all terms that do not cancel after inserting ansatz (2.4) into system (2.3), is formally $\mathcal{O}(\varepsilon^3)$. However, for our error estimates, we need a much smaller residual and have to control its norm in high Sobolev spaces.
In order to get a smaller residual, we extent the ansatz $\varepsilon\Psi_S$ by some $\mathcal{O}(\varepsilon^2)$-terms to a approximation $\varepsilon\tilde{\Psi}$. Then, we exploit that the approximation $\varepsilon\tilde{\Psi}$ is strongly concentrated around a finite number of integer multiples of the basic wave number $k_0 > 0$ such that we can use some cut-off function to restrict the support of the approximation $\varepsilon\tilde{\Psi}$ in Fourier space to small neighborhoods of these wave numbers jk_0, with $j \in \{-5, \ldots, 5\}$, without changing the size of the residual. This way, we obtain a approximation $\varepsilon\Psi$ that is an analytic function and has a residual of the formal order $\mathcal{O}(\varepsilon^6)$. For more details on this construction we refer to Section 2 of [DSW16].
The final approximation that we use, is

$$\varepsilon\Psi = \varepsilon\Psi_c + \varepsilon^2\Psi_q\,, \tag{2.8}$$

where

$$\begin{aligned}
\varepsilon\Psi_c &= \varepsilon\psi_c \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \varepsilon(\psi_1 + \psi_{-1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \varepsilon\left(A_1(\varepsilon(x - c_g t), \varepsilon^2 t)\,\mathbf{E} + c.c.\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
\varepsilon^2\Psi_q &= \varepsilon^2 \begin{pmatrix} \psi_{q_{-1}} \\ \psi_{q_1} \end{pmatrix} = \varepsilon^2\Psi_0 + \varepsilon^2\Psi_2 + \varepsilon^2\Psi_h\,,
\end{aligned}$$

$$\begin{aligned}
\varepsilon^2\Psi_0 &= \varepsilon^2 \begin{pmatrix} A_0(\varepsilon(x - c_g t), \varepsilon^2 t) \\ D_0(\varepsilon(x - c_g t), \varepsilon^2 t) \end{pmatrix}, \\
\varepsilon^2\Psi_2 &= \varepsilon^2 \begin{pmatrix} A_2(\varepsilon(x - c_g t), \varepsilon^2 t)\,\mathbf{E}^2 + c.c. \\ D_2(\varepsilon(x - c_g t), \varepsilon^2 t)\,\mathbf{E}^2 + c.c. \end{pmatrix},
\end{aligned}$$

$$
\begin{aligned}
\varepsilon^2\Psi_h \;=\; & \sum_{n=1,2,3,4} \varepsilon^{1+n} \begin{pmatrix} A_1^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}+c.c. \\ D_1^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}+c.c. \end{pmatrix} \\
& + \sum_{n=1,2,3} \varepsilon^{2+n} \begin{pmatrix} A_0^n(\varepsilon(x-c_g t),\varepsilon^2 t) \\ D_0^n(\varepsilon(x-c_g t),\varepsilon^2 t) \end{pmatrix} \\
& + \sum_{n=1,2,3} \varepsilon^{2+n} \begin{pmatrix} A_2^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^2+c.c. \\ D_2^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^2+c.c. \end{pmatrix} \\
& + \sum_{n=0,1,2} \varepsilon^{3+n} \begin{pmatrix} A_3^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^3+c.c. \\ D_3^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^3+c.c. \end{pmatrix} \\
& + \sum_{n=0,1} \varepsilon^{4+n} \begin{pmatrix} A_4^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^4+c.c. \\ D_4^n(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^4+c.c. \end{pmatrix}, \\
& + \varepsilon^5 \begin{pmatrix} A_5^0(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^5+c.c. \\ D_5^0(\varepsilon(x-c_g t),\varepsilon^2 t)\mathbf{E}^5+c.c. \end{pmatrix},
\end{aligned}
$$

$\mathbf{E} = e^{i(k_0 x - \omega_0 t)}$, $\omega_0 = \omega(k_0)$ and $c_g = \omega'(k_0)$.
Here, $A_1\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)$ is the restriction of $A\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)$ in Fourier space to the interval $\{k \in \mathbb{R} : |k| \le \delta < k_0/20\}$ by some cut-off function, where A is the solution of the NLS-equation (1.7) and $\delta > 0$. More precisely

$$
\begin{aligned}
A_1\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big) &= P_{0,\delta}\big[A\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)\big] \\
&:= \mathcal{F}^{-1}\Big[\chi_{[-\delta,\delta]}(\cdot)\mathcal{F}\big[A\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)\big](\cdot)\Big]
\end{aligned}
$$

where $\chi_{[-\delta,\delta]}$ is the characteristic function on the interval $[-\delta,\delta]$, i.e. $\chi_{[-\delta,\delta]}(k) = 1$ for $[-\delta,\delta]$ and $\chi_{[-\delta,\delta]}(k) = 0$ for $k \notin [-\delta,\delta]$.
The A_j^n and D_j^n are chosen suitably depending on A_1 such that the support of $A_j^n\mathbf{E}^j$ and $D_j^n\mathbf{E}^j$ in Fourier space lies in a small neighborhood of the wave number jk_0.
One can think of $\varepsilon\psi_c$ as $\varepsilon\psi_{NLS}$, just with a support in Fourier space which is restricted to small neighborhoods of the wave numbers $\pm k_0$.
Similarly as in [SW11, DSW16], we obtain:

Lemma 2.1.3. *Let $s_A \ge 7$ and $A \in C([0,T_0], H^{s_A}(\mathbb{R},\mathbb{C}))$ be a solution of the NLS equation* (1.7) *with*

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

Then for all $s \ge 0$ there exist $C_{Res}, C_\Psi, \varepsilon_0 > 0$ depending on C_A such that for all

$\varepsilon \in (0, \varepsilon_0)$ *the approximation* $\varepsilon\Psi = \varepsilon\Psi_c + \varepsilon^2\Psi_q$ *satisfies*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\mathrm{Res}_u(\varepsilon\Psi)\|_{H^s} \leq C_{\mathrm{Res}}\, \varepsilon^{11/2}, \tag{2.9}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} \Big\|\varepsilon\Psi - \varepsilon\psi_{NLS} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Big\|_{H^{s_A}} \leq C_\Psi\, \varepsilon^{3/2}, \tag{2.10}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} (\|\widehat{\Psi}_c\|_{L^1(s+1)(\mathbb{R},\mathbb{C})} + \|\widehat{\Psi}_q\|_{L^1(s+1)(\mathbb{R},\mathbb{C})}) \leq C_\Psi\,, \tag{2.11}$$

$$\|\partial_t\widehat{\psi}_{\pm1} + i\widehat{\omega\psi}_{\pm1}\|_{L^1(s)} \leq C_\Psi\, \varepsilon^2\,. \tag{2.12}$$

Remark 2.1.4. *When* (1.6) *is also true for* $m \in \{\pm5, \pm6, \dots\}$, *by choosing* s_A *higher and expanding* $\varepsilon\Psi$ *further, one could make* $\mathrm{Res}_u(\varepsilon\Psi)$ *arbitrary small (if* ω *or* $\mathrm{sign}(\cdot)\omega(\cdot)$, *and,* ρ *or* $\mathrm{sign}(\cdot)\rho(\cdot)$, *are smooth enough).*

The bound (2.10) *is the reason why we can work with* $\varepsilon\Psi$ *and will still obtain a result for* $\varepsilon\psi_{NLS}$.

We need the bound (2.11) *in order to make estimates like*

$$\|\psi_c f\|_{H^s} \leq C\|\psi_c\|_{C_b^s}\|f\|_{H^s} \leq C\|\widehat{\psi}_c\|_{L^1(s)}\|f\|_{H^s}$$

without losing powers in ε *as we would with* $\|\psi_c\|_{H^s} = \|\widehat{\psi}_c\|_{L^2(s)}$, *where the problem is that* $\|g(\varepsilon\cdot)\|_{L^2} = \varepsilon^{-1/2}\|g(\cdot)\|_{L^2}$.

The bound (2.12) *will be used to approximate* $\partial_t\psi_{\pm1}$.

Proof of lemma 2.1.3. We only give a short proof, for more details compare section 2.4 of [DSW16].
By proceeding for Ψ exactly as for Ψ_S above, we obtain the NLS-equation (1.7) in $\varepsilon^3\mathbf{E}$ and $\varepsilon^3\mathbf{E}^{-1}$. Due to the estimate (see (24) in [S98b])

$$\|\chi_{[-\delta,\delta]}\varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot) - \varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot)\|_{L^2(m)} \leq C(\delta)\, \varepsilon^{m+M-1/2}\|f\|_{H^{m+M}} \tag{2.13}$$

and the fact that A solves the NLS-equation (1.7), we have

$$\Big\|\partial_T A_1 - i\,\frac{\omega''(k_0)}{2}\,\partial_X^2 A_1 - i\nu_2(k_0)A_1|A_1|^2\Big\|_{L^2} = \mathcal{O}(\varepsilon^{3-1/2}).$$

By now looking at the terms $\varepsilon^4\mathbf{E}^j$, $\varepsilon^5\mathbf{E}^j$ and suitably choosing A_j^n, D_j^n, we obtain

$$|\mathrm{Res}_u(\varepsilon\Psi)\|_{L^2} = \mathcal{O}(\varepsilon^{11/2}).$$

As before, for $j \neq 1$, the A_j^n, D_j^n are chosen depending on A_1 by exploiting (1.5). The A_1^1, A_1^2, sometimes even the D_1^1, D_1^2, are determined by solving linear, but inhomogeneous, Schrödinger equations, in which the inhomogeneous terms are determined by A_1. Since $A_1\big((\cdot - c_g t), \varepsilon^2 t\big)$ has a compact support in Fourier space the $A_j^n\big((\cdot - c_g t), \varepsilon^2 t\big)\mathbf{E}^j$, $D_j^n\big((\cdot - c_g t), \varepsilon^2 t\big)\mathbf{E}^j$ can also be chosen with a support in

Fourier space that lies in a small neighborhood of jk_0. Thus, (2.9) is true for all $s \geq 0$ due to the compact support of Ψ in Fourier space.

The estimate (2.10) is a consequence of

$$\varepsilon\Psi = \varepsilon\left(A_1(\varepsilon(x - c_g t), \varepsilon^2 t)\,\mathbf{E} + c.c.\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \varepsilon^2\Psi_q$$

and (2.13).

We obtain (2.11), since $\widehat{\Psi}_c$ and $\widehat{\Psi}_q$ have compact support and

$$\|\mathcal{F}\big[f(\varepsilon\cdot)\big]\|_{L^1} = \|\varepsilon^{-1}\widehat{f}(\varepsilon^{-1}\cdot)\|_{L^1} = \|\widehat{f}(\cdot)\|_{L^1}\,,$$

for $f \in L^1$.

The estimate (2.12) can be seen by expanding the expressions $i\omega A_1 E$ and $i\omega\overline{A}_1 E^{-1}$ in the same way as we did before.

□

For the derivation of the NLS equation and the above lemma, we do not need the assumption (1.9), i.e. that the linear part of our system is at least as strong as the nonlinear part. In other words, we can rigorously derive the NLS equation for a much wider class of systems than the class for which we justify the NLS equation for in this thesis. While this does not necessarily mean that the NLS equation can also be justified for systems where the nonlinear part is stronger than the linear part (see [SSZ15]), it still lets one hope for that.
In [DH18], the NLS equation was justified for an equation with a nonlinear part stronger than the linear one, however the strategy used there only works for equations of the form $\partial_t v = i\omega v + i\rho(v^2)$ with $v(x,t) \in \mathbb{R}$ and $\deg^*(\rho) \leq 1$.

2.2 The error estimates

In this section, we ultimately will prove theorem 1.1.1. That means we will show that there is a solution u to (1.2) such that the H^{s_A}-norm of the error

$$\mathcal{R}_{err} = u - \varepsilon\psi_{NLS} \tag{2.14}$$

remains bounded on the $\mathcal{O}(\varepsilon^{-2})$-time interval $[0, T_0/\varepsilon]$ and we have the estimate

$$\|u - \varepsilon\psi_{NLS}\|_{H^{s_A}} \leq \mathcal{O}(\varepsilon^{3/2})\,. \tag{2.15}$$

To do so, we consider the diagonalized system (2.3) from section 2.1:

$$\begin{aligned}\partial_t u_{-1}(x,t) &= -i\omega u_{-1}(x,t) - \frac{1}{2}i\rho\big(u_{-1} + u_1\big)^2(x,t)\,,\\ \partial_t u_1(x,t) &= i\omega u_1(x,t) + \frac{1}{2}i\rho\big(u_{-1} + u_1\big)^2(x,t)\,,\end{aligned}$$

with $t, x \in \mathbb{R}$ and $u_{-1}(x,t), u_1(x,t) \in \mathbb{R}$, that emerges from the system (1.2) by the transformation

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}.$$

We then proceed to estimate the error

$$\begin{pmatrix} \mathcal{R}_{-1} \\ \mathcal{R}_1 \end{pmatrix} = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi\,, \tag{2.16}$$

which the improved approximation $\varepsilon\Psi$ from section 2.1 makes on the $\mathcal{O}(\varepsilon^{-2})$-time interval $[0, T_0/\varepsilon^2]$. This will make the proof much simpler and, due to the estimate (2.10), we directly obtain an estimate for (2.14) from an estimate for (2.16).

Let

$$\Psi = \begin{pmatrix} \psi_c + \varepsilon\psi_{q_{-1}} \\ \varepsilon\psi_{q_1} \end{pmatrix},$$

be the approximation (2.8) from section 2.1. We write the error (2.16) as

$$\varepsilon^\beta \begin{pmatrix} \vartheta R_{-1} \\ \vartheta R_1 \end{pmatrix} = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi \tag{2.17}$$

where $\beta = 5/2$ and ϑ is an invertible operator on $L^2(\mathbb{R})$ that later will be given by some weight function $\hat{\vartheta}$ in Fourier space. Throughout this section, we will now work with the rescaled error

$$\begin{pmatrix} R_{-1} \\ R_1 \end{pmatrix} = \varepsilon^{-\beta}\vartheta^{-1}\Big(\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi\Big)\,, \tag{2.18}$$

where $\vartheta^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the inverse of the operator ϑ.
The dynamics we obtain for this rescaled error by plugging in the above definition into the diagonalized system are given by

$$\begin{aligned} \partial_t R_{-1} &= -i\omega R_{-1} - \varepsilon i\rho\vartheta^{-1}(R_\psi(\vartheta R_{-1} + \vartheta R_1)) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{-1}}(\varepsilon\Psi)\,, \\ \partial_t R_1 &= i\omega R_1 + \varepsilon i\rho\vartheta^{-1}(R_\psi(\vartheta R_{-1} + \vartheta R_1)) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_1}(\varepsilon\Psi)\,, \end{aligned} \tag{2.19}$$

where

$$R_\psi := \psi + \frac{1}{2}\varepsilon^{\beta-1}(\vartheta R_{-1} + \vartheta R_1), \tag{2.20}$$

$$\psi := \psi_c + \varepsilon\psi_Q := \psi_c + \varepsilon(\psi_{q_{-1}} + \psi_{q_1})\,. \tag{2.21}$$

2.2.1 $\mathcal{O}(\varepsilon^{-2})$- time scale via normal form transformations

If we could obtain an estimate of the form

$$\partial_t\big(\|R_{-1}\|_{H^s}^2+\|R_1\|_{H^s}^2\big)\le\varepsilon^2\,\mathcal{O}\big(\|R_{-1}\|_{H^s}^2+\|R_1\|_{H^s}^2+1\big)$$

for $t\in[0,T_0]$, an application of Gronwall's inequality would give us

$$\sup_{t\in[0,T_0/\varepsilon^2]}\|R_{-1}(t)\|_{H^s}+\|R_1(t)\|_{H^s}\le C$$

for some $C\ge 0$ and we would have

$$\sup_{t\in[0,T_0/\varepsilon^2]}\Big\|\begin{pmatrix}u_{-1}\\u_1\end{pmatrix}-\varepsilon\Psi\Big\|_{H^s}=\varepsilon^\beta\sup_{t\in[0,T_0/\varepsilon^2]}\Big\|\begin{pmatrix}\vartheta R_{-1}\\\vartheta R_1\end{pmatrix}\Big\|_{H^s}\le\varepsilon^\beta\|\vartheta\|_{L^2\to L^2}\,C.$$

We have chosen $\beta=5/2$ large enough and constructed the approximation Ψ in such a way that we formally have

$$\begin{aligned}\partial_t R_j&=ji\omega R_j+j\varepsilon i\rho\vartheta^{-1}(R_\psi(\vartheta R_{-1}+\vartheta R_1))+\varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi)\\&=ji\omega R_j+j\varepsilon i\rho\vartheta^{-1}\big(\psi(\vartheta R_{-1}+\vartheta R_1)\big)\\&\quad+j\frac{1}{2}\varepsilon^\beta i\rho\vartheta^{-1}(\vartheta R_{-1}+\vartheta R_1)^2+\varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi)\\&=ji\omega R_j+j\varepsilon i\rho\vartheta^{-1}\big(\psi(\vartheta R_{-1}+\vartheta R_1)\big)+\mathcal{O}(\varepsilon^2)\,,\end{aligned}$$

for $j=\pm1$.
We will choose the operator ϑ such that we can be sure that

$$\|i\rho\vartheta^{-1}f\|_{L^2}\le\mathcal{O}(\|f\|_{H^{\deg^*\rho}})\,,\tag{2.22}$$

$$\|\vartheta f\|_{L^2}\le\mathcal{O}(\|f\|_{L^2})\,,\tag{2.23}$$

i.e. that the operators $i\rho\vartheta^{-1}$ and ϑ do not cause a loss of ε-powers.
For this reason, we can take advantage of the fact that

$$\psi=\psi_c+\varepsilon\psi_Q\tag{2.24}$$

in order to get

$$\partial_t R_j=ji\omega R_j+j\varepsilon i\rho\vartheta^{-1}\big(\psi_c(\vartheta R_{-1}+\vartheta R_1)\big)+\mathcal{O}(\varepsilon^2)\,.$$

The term $ji\omega R_j$ is harmless, since, due to the fact that the function ω is odd, it will vanish when we multiply the equation with $2R_j$ and then integrate in space to obtain $\partial_t\|R_j\|_{L^2}$ on the left hand side.
The $\mathcal{O}(\varepsilon)$-term however prevents us from obtaining an estimate on a $\mathcal{O}(\varepsilon^{-2})$-timescale. The basic approach to get rid of this term is preforming a normal form transformation

$$R_j\to R_j+\varepsilon\sum_{j_2\in\{\pm1\}}\vartheta^{-1}N_{jj_2}(\psi_c,R_{j_2}),\tag{2.25}$$

where N is a suitably chosen bilinear mapping.

Remark 2.2.1. *We treat the operator ϑ^{-1} similarly as the small perturbation parameter ε since ϑ^{-1} can cause a loss of ε-powers.*

By defining

$$\tilde{R}_j = R_j + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}), \tag{2.26}$$

we formally obtain

$$\begin{aligned}
\partial_t \tilde{R}_j &= \partial_t R_j + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(\partial_t \psi_c, R_{j_2}) + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(\psi_c, \partial_t R_{j_2}) \\
&= j i \omega \tilde{R}_j - j\varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} i\omega N_{jj_2}(\psi_c, R_{j_2}) \\
&\quad + j\varepsilon i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big) \\
&\quad + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(-i\omega\psi_c, R_{j_2}) + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(\psi_c, j_2 i \omega R_{j_2}) \\
&\quad + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(i\omega\psi_c + \partial_t \psi_c, R_{j_2}) \\
&\quad + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}\Big(\psi_c, j_2 \varepsilon i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big) + \mathcal{O}(\varepsilon^2)\Big) + \mathcal{O}(\varepsilon^2)\,.
\end{aligned}$$

The first idea would be to choose normal form transformations $N_{jj_2} = \tilde{N}_{jj_2}$, where $\tilde{N}_{jj_2}$ is defined by

$$-ji\omega \tilde{N}_{jj_2}(\psi_c, R_{j_2}) + ji\rho\big(\psi_c \vartheta R_{j_2}\big) + \tilde{N}_{jj_2}(-i\omega\psi_c, R_{j_2}) + \tilde{N}_{jj_2}(\psi_c, j_2 i \omega R_{j_2}) = 0\,,$$

i.e.

$$\widehat{\tilde{N}}_{j_1 j_2}(\psi_c, R_{j_2})(k) = \int_{\mathbb{R}} \tilde{n}_{j_1 j_2}(k, k-m, m) \widehat{\psi}_c(k-m) \widehat{R}_{j_2}(m)\, dm\,,$$

$$\tilde{n}_{j_1 j_2}(k, k-m, m) = \frac{\rho(k)\, \hat{\vartheta}(m)\, \chi_c(k-m)}{\omega(k) - j_1 j_2 \omega(m) + j_1 \omega(k-m)}\,,$$

where χ_c is the characteristic function on $\operatorname{supp} \widehat{\psi}_c$.
We choose the approximation Ψ from section 2.1 such that ψ_c in Fourier space has the compact support

$$\operatorname{supp} \widehat{\psi}_c = \{k \in \mathbb{R} : |k \pm k_0| \leq \delta < k_0/20\}\,, \tag{2.27}$$

where the parameter δ will later be chosen suitably small, but independent of ε.
Due to our conditions concerning the solution of the equations

$$\omega(k) - j_1 j_2 \omega(k \mp k_0) + j_1 \omega(\pm k_0) = 0 \tag{2.28}$$

for $j_1, j_2 \in \{\pm 1\}$, which we have formulated in (1.12)-(1.17), we are able to show that when (1.16) is true, i.e. when

$$0 \neq \pm\omega(0^+) \neq 2\omega(k_0)\,, \tag{2.29}$$

the mappings $\tilde{N}_{jj_2}(\psi_c, \cdot)$ are indeed well-defined and map $H^1(\mathbb{R})$ on $L^2(\mathbb{R})$.

However, when we do not have (2.29), the mappings $\tilde{N}_{jj_2}$ are in general not well-defined for functions of $H^1(\mathbb{R})$, or even for functions of $C_c^\infty(\mathbb{R})$. This is due to the resonance happening in $m = 0$ that corresponds to the solutions $k = \pm k_0$ of (2.28), and is in general nontrivial. In other words, we have the problem that $\tilde{n}_{j_1 j_2}(k, k-m, m)$ can grow unlimitedly near $m = 0$.
To address this issue, we define the operator ϑ in Fourier space via the weight function

$$\hat{\vartheta}(k) = \begin{cases} \quad 1 & \text{, when } 0 \neq \pm\omega(0^+) \neq 2\omega(k_0), \\ \begin{cases} \varepsilon + (1-\varepsilon)\frac{|k|}{\delta} & \text{for } |k| \leq \delta \\ 1 & \text{for } |k| > \delta \end{cases} & \text{, else.} \end{cases} \tag{2.30}$$

Herby the parameter δ is the same δ as the one in (2.27). The idea to handle a nontrivial resonance in $m = 0$, i.e. $k = \pm k_0$, by rescaling the error with such a weight function has already been used in many papers (e.g. [DS06, DSW16, CW17, DH18]).
Let $\hat{P}_{a,b}$ denote the characteristic function on the set $\{k : a \leq |k| \leq b\}$ and $P_{a,b}$ be the operator defined by the symbol $\hat{P}_{a,b}$ in Fourier space.
Due to our choice of ϑ, we formally have in the case that (2.29) is hurt

$$\begin{aligned} & j\varepsilon i\rho\vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big) \\ &\quad = j\varepsilon i\rho\vartheta^{-1}\big(\psi_c P_{\varepsilon,\infty}\vartheta(R_{-1} + R_1)\big) + j\varepsilon i\rho\vartheta^{-1}\big(\psi_c P_{0,\varepsilon}\vartheta(R_{-1} + R_1)\big) \\ &\quad = j\varepsilon i\rho\vartheta^{-1}\big(\psi_c P_{\varepsilon,\infty}\vartheta(R_{-1} + R_1)\big) + \mathcal{O}(\varepsilon^2)\,, \end{aligned}$$

since $\hat{\vartheta}(k) \leq \mathcal{O}(\varepsilon)$ for $|k| \leq \varepsilon$.
Therefore, we now only have to eliminate the term

$$j\varepsilon i\rho\vartheta^{-1}\big(\psi_c P_{\varepsilon,\infty}\vartheta(R_{-1} + R_1)\big)$$

when (2.29) is hurt.
So, we define the normal form transformation by

$$\widehat{N}_{j_1 j_2}(\psi_c, R_{j_2})(k) = \int_{\mathbb{R}} n_{j_1 j_2}(k, k-m, m)\widehat{\psi}_c(k-m)\widehat{R}_{j_2}(m)\,dm\,, \tag{2.31}$$

$$n_{j_1 j_2}(k, k-m, m) = \frac{\rho(k)\,\hat{\vartheta}_{\varepsilon,\infty}(m)\,\chi_c(k-m)}{\omega(k) - j_1 j_2\omega(m) + j_1\omega(k-m)}\,,$$

where $\hat{\vartheta}_{\varepsilon,\infty}(m) = 1$ when $0 \neq \pm\omega(0^+) \neq 2\omega(k_0)$ is true, otherwise

$$\hat{\vartheta}_{\varepsilon,\infty}(m) = \begin{cases} 0 & \text{for } |m| \leq \varepsilon\,, \\ \varepsilon + (1-\varepsilon)\frac{|m|}{\delta} & \text{for } \varepsilon < |m| \leq \delta\,, \\ 1 & \text{for } |m| > \delta\,. \end{cases}$$

In the case of no resonances happening, i.e. where (2.29) is true, the normal form transformation $N_{j_1 j_2}$ looks exactly like $\tilde{N}_{j_1 j_2}$ from above. In the case where (2.29) is hurt however, the resonance in $m = 0$ is cut out.
Due to this (and the fact that $\rho(0) = 0$ when (2.29) is hurt due to (1.16)-(1.17)) we can show that the normal form transformations $N_{j_1 j_2}$ are always well-defined and the $N_{j_1 j_2}(\psi_c, \cdot)$ map $H^1(\mathbb{R})$ on $L^2(\mathbb{R})$.

Defining ϑ by (2.30) allowed us to obtain a well-defined normal form transformation $N_{j_1 j_2}$, but it has the direct consequence that when (2.29) is hurt, we have

$$\|\vartheta^{-1}\|_{L^2 \to L^2} = \mathcal{O}(\varepsilon^{-1}). \tag{2.32}$$

So, while we successfully obtain

$$\partial_t \tilde{R}_j = j i \omega \tilde{R}_j + \mathcal{O}(\varepsilon^2)$$

in the case that (2.29) is true, we only obtain

$$\partial_t \tilde{R}_j = j i \omega \tilde{R}_j + \varepsilon^2 \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{j j_2}\Big(\psi_c, j_2 i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big)\Big) + \mathcal{O}(\varepsilon^2)\,.$$

when (2.29) is not true.
We first observe that

$$\begin{aligned} &\varepsilon^2 \vartheta^{-1} N_{j j_2}\Big(\psi_c, j_2 i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big)\Big) \\ &\quad = \varepsilon^2 P_{0,\delta} \vartheta^{-1} N_{j j_2}\Big(\psi_c, j_2 i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big)\Big) + \mathcal{O}(\varepsilon^2)\,, \end{aligned}$$

due to the definition (2.30) of the operator ϑ.
Analyzing this term further in Fourier space, we will see that

$$\begin{aligned} &\varepsilon^2 \vartheta^{-1} N_{j j_2}\Big(\psi_c, j_2 i \rho \vartheta^{-1}\big(\psi_c(\vartheta R_{-1} + \vartheta R_1)\big)\Big) \\ &\quad = \varepsilon^2 P_{0,\delta} \vartheta^{-1} \sum_{j_4 \in \{\pm 1\}} N_{j j_2}\Big(\psi_{j_4}, j_2 i \rho \vartheta^{-1}\big(\psi_{j_4}(\vartheta R_{-1} + \vartheta R_1)\big)\Big) + \mathcal{O}(\varepsilon^2)\,, \end{aligned}$$

where

$$\psi_1(x,t) = A_1\big(\varepsilon(x - c_g t), \varepsilon^2 t\big) e^{i(k_0 x - \omega_0 t)}\,, \qquad \psi_{-1} = \overline{\psi_1}\,, \qquad \psi_c = \psi_1 + \psi_{-1}\,,$$

just as in section 2.1.
For this reason, we have to preform a second normal form transformation

$$\tilde{R}_j \to \check{R}_j := \tilde{R}_j + \varepsilon^2 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} \vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) \tag{2.33}$$

where $\mathcal{T}_{jj_2j_3j_4}$ is a multi-linear mapping such that

$$\begin{aligned} & P_{0,\delta}\vartheta^{-1} N_{j_1j_2}(\psi_{j_4}, j_2\vartheta^{-1}\, i\rho(\psi_{j_4}\vartheta R_{j_3})) \\ & - j_1\, i\omega\vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(-i\omega\psi_{j_4}, \psi_{j_4}, R_{j_3}) \\ & + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, -i\omega\psi_{j_4}, R_{j_3}) + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, j_3\, i\omega R_{j_3}) \quad = \mathcal{O}(1)\,. \end{aligned} \tag{2.34}$$

We finally obtain this by defining

$$\widehat{\mathcal{T}}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3})(k) = \int_{\mathbb{R}} t_{j_1,j_2,j_3,j_4}(k)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R}_{j_3}(n)\, dn\, dm\,, \tag{2.35}$$

$$t_{j_1,j_2,j_3,j_4}(k) = \begin{cases} 0 & \text{if (2.29) is true,} \\ \dfrac{-j_2\, \hat{P}_{0,\delta}(k)\, n_{j_1j_2}(k, j_4k_0, k-j_4k_0)\, \rho(k-j_4k_0)}{\big(-j_1\omega(k) - 2\omega(j_4k_0) + j_3\omega(k-2j_4k_0)\big)} & \text{else.} \end{cases}$$

Due to our definition of $n_{j_1j_2}$ and thanks to (1.16)-(1.17), we can show that this normal form transformations $\mathcal{T}_{j_1j_2j_3j_4}$ are free of nontrivial resonances and that the $\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, \cdot)$ map $L^2(\mathbb{R}, \mathbb{C})$ onto $L^2(\mathbb{R}, \mathbb{C})$.
We now can obtain

$$\partial_t \check{R}_j = ji\omega\check{R}_j + \mathcal{O}(\varepsilon^2)$$

and therefore show

$$\partial_t\big(\|\check{R}_{-1}\|_{L^2}^2 + \|\check{R}_1\|_{L^2}^2\big) = \mathcal{O}(\varepsilon^2).$$

Definition 2.2.2. *We define the energy*

$$E_0(R) = \|\check{R}_{-1}\|_{L^2}^2 + \|\check{R}_1\|_{L^2}^2\,, \tag{2.36}$$

with

$$\check{R}_j = R_j + \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) + \varepsilon^2 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3})$$

for $j \in \{-1, 1\}$, where N_{jj_2} is as in (2.31)*, $\mathcal{T}_{jj_2j_3j_4}$ is as in* (2.35)*, ψ_{j_4} is as in* (2.8) *and $\vartheta^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the inverse of the operator ϑ that was defined in* (2.30).

Remark 2.2.3. *Since we sometimes have to use the estimate* (2.32)*, we have placed the operator* ϑ^{-1} *outside of our normal-form transforms. This way our estimates for the normal-form transforms* $N_{j_1 j_2}$ *and* $\mathcal{T}_{j_1 j_2 j_3 j_4}$ *will be independent of* ε*.*

We will now prove the statements we have made so far.

Lemma 2.2.4. *We always have*

$$|k\hat{\vartheta}^{-1}(k)| \leq 1 + |k| \,, \tag{2.37}$$

that means in particular that

$$\|i\rho\vartheta^{-1} f\|_{L^2} \leq \mathcal{O}(\|f\|_{H^{\deg^*\rho}}). \tag{2.38}$$

When (2.29) *is not true, i.e.* $\vartheta \neq id_{L^2}$*, there is a constant* $C = C(\delta)$ *such that for all* $k \in \mathbb{R}$*:*

$$|k^{-1}\,\hat{\vartheta}_{\varepsilon,\infty}(k)| \leq C. \tag{2.39}$$

Proof. Obviously (2.37) is true for $\hat{\vartheta}(k) = 1$, otherwise we have

$$|k\,\hat{\vartheta}^{-1}(k)| = \begin{cases} |k| & \text{for } |k| > \delta \,, \\ \dfrac{|k|}{\varepsilon + (1-\varepsilon)\frac{|k|}{\delta}} & \text{for } |k| \leq \delta \,. \end{cases}$$

For $0 < |k| \leq \delta$, we have

$$\frac{|k|}{\varepsilon + (1-\varepsilon)\frac{|k|}{\delta}} \leq \frac{1}{\frac{\varepsilon}{|k|} + \frac{1-\varepsilon}{\delta}} \leq \delta$$

such that (2.37) is true.

Estimate (2.38) now follows due to the fact that $\vartheta^{-1} \neq id_{L^2}$ implicates $\rho(0) = 0$, i.e. implicates $\rho(k) = \mathcal{O}(k)$ for $|k| \to 0$.

When (2.29) is not true, we have

$$|k^{-1}\,\hat{\vartheta}_{\varepsilon,\infty}(k)| = \begin{cases} 0 & \text{for } 0 < |k| \leq \varepsilon \,, \\ \dfrac{\varepsilon}{|k|} + \dfrac{(1-\varepsilon)}{\delta} & \text{for } \varepsilon \leq |k| \leq \delta \,, \\ \dfrac{1}{|k|} & \text{for } |k| \geq \delta \,. \end{cases}$$

Thus we get (2.39).

□

Lemma 2.2.5. *The normal-form transforms $N_{j_1j_2}$ were constructed such that for all $f \in H^{\deg^*(\rho)+1}(\mathbb{R})$:*

$$-j_1 i\omega N_{j_1j_2}(\psi_c, f) - N_{j_1j_2}(i\omega\psi_c, f) + j_2 N_{j_1j_2}(\psi_c, i\omega f) = -j_1 i\rho(\psi_c \vartheta_{\varepsilon,\infty} f)\,, \quad (2.40)$$

where

$$\|j_1 i\rho(\psi\vartheta f) - j_1 i\rho(\psi_c\vartheta_{\varepsilon,\infty} f)\|_{L^2} = \mathcal{O}(\varepsilon)\, \|f\|_{H^{\deg^*(\rho)}}. \quad (2.41)$$

Moreover, for every fix $h \in L^2(\mathbb{R},\mathbb{R})$ the operators $N_{j_1j_2}(h,\cdot)$ are continuous linear operators which map $H^1(\mathbb{R},\mathbb{R})$ into $L^2(\mathbb{R},\mathbb{R})$. In particular, there is a $C = C(\|\widehat{h}(\cdot)\chi_c(\cdot)\|_{L^1})$ such that for all $g \in H^1(\mathbb{R})$ we have

$$\|N_{jj}(h,g)\|_{L^2} \le C\|g\|_{H^1}\,, \quad (2.42)$$

$$\|N_{j-j}(h,g)\|_{L^2} \le C\|g\|_{L^2}. \quad (2.43)$$

Remark 2.2.6. *More precisely, we have*

$$\|N_{jj}(h,g)\|_{L^2} \le C\|g\|_{H^q}\,, \quad (2.44)$$

$$\|N_{j-j}(h,g)\|_{H^r} \le C\|g\|_{L^2}\,, \quad (2.45)$$

for

$$q \ge \min\{\deg^*(\rho), \deg^*(\rho) - \deg(\omega) + 1\} \quad \textit{and} \quad r \le \deg(\omega) - \deg^*(\rho).$$

Proof.

In order to find possible resonances for $N_{j_1j_2}$, we have to look at the zeros of the denominator of $n_{j_1j_2}$, i.e. of

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m),$$

for $\chi_c(k-m) \neq 0$, i.e. for $|k - m \mp k_0| \le \delta$.
Due to (1.12), we can chose δ such small that for $|k - m \mp k_0| \le \delta$ the equation

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m) = 0, \quad (2.46)$$

can have no other solutions than $k = 0$ or $m = 0$.

We first check $k = 0$ and therefore assume $|k| \le \delta$.
For $|k| \le \delta$, we also have $|-m \mp k_0| \le 2\delta$ since $|k - m \mp k_0| \le \delta$. Using Taylor in order to expand $\omega(k)$ in the point $\mathrm{sign}(k)\cdot 0^+$ and $\omega(k-m)$ in the point $-m$, we obtain

$$\begin{aligned}
&\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m) \\
&\qquad = \omega(\mathrm{sign}(k)\cdot 0^+) - j_1j_2\omega(m) + j_1\omega(-m) \\
&\qquad\quad + \omega'(\mathrm{sign}(k)\cdot 0^+)\, k + j_1\omega'(-m)\, k + \mathcal{O}(k^2) \\
&\qquad = \omega(\mathrm{sign}(k)\cdot 0^+) - j_1(j_2+1)\,\omega(m) + \big(\omega'(0) + j_1\omega'(m)\big)\, k + \mathcal{O}(k^2).
\end{aligned}$$

Thus, if

$$\omega(0^{\mp}) \neq (j_2+1)\,\omega(k_0), \tag{2.47}$$

and we choose δ small enough, $N_{j_1j_2}$ has no resonance in $k=0$.
If (2.47) is hurt but

$$\pm\omega'(0) \neq \omega'(k_0), \tag{2.48}$$

we can choose δ small enough such that

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m) = \mathcal{O}(k) \qquad \text{for } k \to 0.$$

When (1.16) is true, we have (2.47) and thus $N_{j_1j_2}$ has no resonance in $k=0$. When instead (1.17) is true, we always have (2.48) and $\rho(k) = \mathcal{O}(k)$ for $k \to 0$, thus $N_{j_1j_2}$ can at worst have a trivial resonance in $k=0$.

Now, we check $m=0$ and assume $|m| \leq \delta$.
For $|m| \leq \delta$, we also have $|k \mp k_0| \leq 2\delta$ since $|k-m\mp k_0| \leq \delta$. Using Taylor in order to expand $\omega(m)$ in the point $\operatorname{sign}(m)\cdot 0^+$ and $\omega(k-m)$ in the point k, we get

$$\begin{aligned}
&\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m)\\
&\quad = \omega(k) - j_1j_2\omega(\operatorname{sign}(m)\cdot 0^+) + j_1\omega(k)\\
&\qquad - j_1j_2\omega'(\operatorname{sign}(m)\cdot 0^+)m - j_1\omega'(k)m + \mathcal{O}(m^2)\\
&\quad = -j_1j_2\omega(\operatorname{sign}(m)\cdot 0^+) + (1+j_1)\omega(k) + \big(-j_1j_2\omega'(0) - j_1\omega'(k)\big)m + \mathcal{O}(m^2).
\end{aligned}$$

If we have

$$\omega(0^{\mp}) \neq (j_1+1)\,\omega(k_0), \tag{2.49}$$

and choose δ small enough, $N_{j_1j_2}$ has no resonance in $m=0$.
If (2.49) is hurt but

$$\pm\omega'(0) \neq \omega'(k_0), \tag{2.50}$$

we can choose δ small enough such that

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m) = \mathcal{O}(m) \qquad \text{for } m \to 0.$$

When (1.16) holds, we have (2.49) and thus $N_{j_1j_2}$ has no resonance in $m=0$.

When instead (1.17) holds, we have (2.50). Due to (2.39), this means $N_{j_1j_2}$ can at worst have a trivial resonance in $m=0$.

Last but not least, we have to check $|k|, |m| \to \infty$.
Note that $|m| \to \infty$ always implies $|k| \to \infty$, since $|k-m\mp k_0| \leq \delta$. Using Taylor, we get

$$\begin{aligned}
&\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m)\\
&\quad = (1-j_1j_2)\omega(k) + j_1\omega(k-m) - j_1j_2\omega'(k)(m-k) + \mathcal{O}\big(\omega''(k)\big).
\end{aligned}$$

Due to (1.11), we now only have to look at the following four cases.
When $\deg(\omega) > 1$, we can choose δ such small that there are no resonances, especially since $k_0 \neq 0$.
When $\deg(\omega) = 1$, we can choose δ such small that there are no resonances due to (1.13).
When $0 < \deg(\omega) < 1$, we can choose δ such small that there are no resonances due to (1.14).
When $\deg(\omega) \leq 0$, we can choose δ such small that there are no resonances due to (1.14) and (1.15).

After we have now shown that our normal-form transform $N_{j_1 j_2}$ has no non-trivial resonances, we can show the rest of the lemma.

The property (2.40) can be easily checked in Fourier space.
For the estimate (2.41), we have

$$\begin{aligned} \|j_1 i\rho(\psi\vartheta f) - j_1 i\rho(\psi_c \vartheta_{\varepsilon,\infty} f)\|_{L^2} &= \|i\rho(\varepsilon\psi_Q \vartheta f) + i\rho(\psi_c \vartheta_{0,\varepsilon} f)\|_{L^2} \\ &= \mathcal{O}(\varepsilon)\, \|f\|_{H^{\deg^*(\rho)}}\,, \end{aligned}$$

especially, since $\hat{\vartheta}_{0,\varepsilon}(k) \leq \mathcal{O}(\varepsilon)$.

We now will show that the $N_{j_1 j_2}(h, \cdot)$ are continuous linear operators.
For later purposes, we will especially focus on writing the bilinear operators $N_{j_1 j_2}(\cdot,\cdot)$ as a sum of products of linear operators, plus some smoothing bilinear operator.

First we look at N_{jj}.
For $|k| \to \infty$, we have

$$n_{jj}(k, k-m, m) = \frac{\rho(k)\, \chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)}.$$

We want a form for $n_{jj}(k, k-m, m)$, for $|k| \to \infty$, which only consists of terms that are products of functions in one variable, plus some smoothing term. In order to obtain this, we only have to look at

$$\frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)}.$$

Using Taylor, we get

$$\omega(k) - \omega(m) = \omega'(m)\,(k-m) + T(k, k-m, m),$$

where

$$T(k, k-m, m)\chi_c(k-m) = \Big(\sum_{l=2}^{p} \frac{1}{l!}\, \omega^{(l)}(m)\,(k-m)^l + \mathcal{O}\big(\omega^{(p+1)}(m)\big) \Big)\chi_c(k-m),$$

for some sufficiently large chosen $p \geq \lceil \deg^*(\rho) \rceil$.
We then use the expansion

$$\frac{a}{b+c} = \sum_{l=0}^{n} (-1)^l \frac{ac^l}{b^{l+1}} + (-1)^{n+1} \frac{ac^{n+1}}{b^{n+1}(b+c)} \qquad (b+c \neq 0,\, b \neq 0) \tag{2.51}$$

in order to obtain a form for

$$\frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)} \quad (\text{for } |k| \to \infty)$$

which only consists of terms that are products of functions in one variable plus some $\mathcal{O}(|m|^{-\deg^*(\rho) - \deg^*(\rho')})$-term.
We distinguish the three cases $\deg(\omega) > 1$, $\deg(\omega) = 1$ and $\deg(\omega) < 1$.

If $\deg(\omega) > 1$ (i.e. $\deg(\omega') > 0$), we have for $|k| \to \infty$:

$$\begin{aligned}
&\frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)} \qquad (2.52)\\
&\quad = \frac{\chi_c(k-m)}{\omega'(m)\,(k-m) + T(k, k-m, m) + j\,\omega(k-m)}\\
&\quad = \Big(\frac{1}{\omega'(m)\,(k-m)} - \frac{T(k, k-m, m) + j\,\omega(k-m)}{\omega'(m)^2\,(k-m)^2}\\
&\qquad + \frac{\big(T(k, k-m, m) + j\,\omega(k-m)\big)^2}{\omega'(m)^3\,(k-m)^3} - \frac{\big(T(k, k-m, m) + j\,\omega(k-m)\big)^3}{\omega'(m)^4\,(k-m)^4}\\
&\qquad \pm \cdots + \mathcal{O}(|m|^{-\deg^*(\rho) - \deg^*(\rho')}) \Big) \chi_c(k-m).
\end{aligned}$$

If $\deg(\omega) = 1$ (i.e. $\deg(\omega') = 0$), we have for $|k| \to \infty$:

$$\begin{aligned}
&\frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)} \qquad (2.53)\\
&\qquad = \Big(\frac{1}{\omega'(m)\,(k-m) + j\omega(k-m)} + \mathcal{O}(|m|^{-1)}) \Big) \chi_c(k-m).
\end{aligned}$$

If $\deg(\omega) < 1$ (i.e. $\deg(\omega') < 0$), we have for $|k| \to \infty$:

$$\begin{aligned}
&\frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j\,\omega(k-m)} \qquad (2.54)\\
&\qquad = \sum_{n=0}^{n(\omega')} (-1)^n j^{n+1} \frac{(\omega'(m))^n\,(k-m)^n}{((\omega(k-m))^{n+1}} + \mathcal{O}(|m|^{-1}) \Big) \chi_c(k-m).
\end{aligned}$$

Due to (1.9), (1.11) and (1.10), by exploiting (2.52), (2.53) and (2.54), we can now see that the $N_{jj}(h, \cdot)$ map $H^1(\mathbb{R})$ on $L^2(\mathbb{R})$ by taking advantage of Young's inequality for convolutions

$$
\begin{aligned}
\|N_{jj}(h,g)\|_{L^2} &\lesssim \|\widehat{N}_{jj}(h,g)\|_{L^2} = \Big\| \int_{\mathbb{R}} n_{jj}(\cdot,\cdot-m,m)\widehat{h}(\cdot-m)\widehat{g}(m)\,dm\Big\|_{L^2} \\
&\leq \mathcal{O}\Big(\sup_{k,m\in\mathbb{R}} \frac{|n_{jj}(k,k-m,m)|}{(|m|^2+1)^{1/2}}\Big)\Big\| \int_{\mathbb{R}} |\widehat{h}(\cdot-m)\chi_c(\cdot-m)\,(|m|^2+1)^{1/2}\,\widehat{g}(m)|\,dm\Big\|_{L^2} \\
&\leq \mathcal{O}\big(\|\widehat{h}(\cdot)\chi_c(\cdot)\|_{L^1}\big)\,\|g\|_{H^1}\,.
\end{aligned}
$$

Now, we look at $N_{j,-j}$.
Using Taylor, we get for $|k|\to\infty$:

$$
\begin{aligned}
n_{j,-j}(k,k-m,m) &= \frac{\rho(k)\,\chi_c(k-m)}{\omega(k)+\omega(m)+j\,\omega(k-m)} \\
&= \frac{\rho(k)\,\chi_c(k-m)}{2\omega(k)+T(k,k-m)+j\,\omega(k-m)},
\end{aligned}
$$

where $T(k,k-m)$ is now given by

$$
T(k,k-m,m) = \sum_{l=1}^{p} \frac{(-1)^l}{l!}\,\omega^{(l)}(k)\,(k-m)^l + \mathcal{O}\big(\omega^{(p+1)}(k)\big).
$$

for some sufficiently large chosen $p \geq \lceil \deg^*(\rho)\rceil$.

As before, we use the expansion (2.51) in order to obtain a form for $n_{j,-j}(k,k-m,m)$ for $|k|\to\infty$ which consists of terms that are products of functions in one variable plus terms which are harmless:

$$
\begin{aligned}
n_{j,-j}(k,k-m,m) = \Big(&\frac{\rho(k)}{2\omega(k)} - \frac{\rho(k)\,\big(T(k,k-m,m)+j\,\omega(k-m)\big)}{4\omega(k)^2} \\
&+ \frac{\rho(k)\,\big(T(k,k-m,m)+j\,\omega(k-m)\big)^2}{8\omega(k)^3} \\
&\mp \cdots + \mathcal{O}(|k|^{-\deg^*(\rho)})\Big)\chi_c(k-m).
\end{aligned}
\tag{2.55}
$$

We can now see that due to (1.9) and (1.11) the $N_{j-j}(h,\cdot)$ map $L^2(\mathbb{R})$ on $L^2(\mathbb{R})$ by exploiting Young's inequality for convolutions..

Finally, since

$$
n_{j_1j_2}(-k,-(k-m),-m) = n_{j_1j_2}(k,k-m,m) \in \mathbb{R},
$$

the $N_{j_1j_2}(h,\cdot)$ map real-valued functions on real-valued functions.

□

Lemma 2.2.7. (cf. Lemma 3.5 in [DSW16]). *Fix $p \in \mathbb{R}$. Assume that $\kappa \in C(\mathbb{R}^3, \mathbb{C})$, that $g \in C^2(\mathbb{R}, \mathbb{C})$ has a finitely supported Fourier transform and that $f \in H^s(\mathbb{R}, \mathbb{C})$ for $s \geq 0$.*

a) *If κ is Lipschitz continuous with respect to its second argument in some neighborhood of p, then there exist $C_{g,\kappa,p} > 0$, $\varepsilon_0 > 0$ such that*

$$\Big\| \int \big(\kappa(\cdot, \cdot - \ell, \ell) - \kappa(\cdot, p, \ell)\big)\, \varepsilon^{-1} \widehat{g}\Big(\frac{\cdot - \ell - p}{\varepsilon}\Big) \widehat{f}(\ell)\, d\ell \Big\|_{L^2(s)} \leq C_{g,\kappa,p}\, \varepsilon \|f\|_{H^s} \quad (2.56)$$

for all $\varepsilon \in (0, \varepsilon_0)$.

b) *If κ is globally Lipschitz continuous with respect to its third argument, then there exist $D_{g,\kappa} > 0$, $\varepsilon_0 > 0$ such that*

$$\Big\| \int \big(\kappa(\cdot, \cdot - \ell, \ell) - \kappa(\cdot, \cdot - \ell, \cdot - p)\big)\, \varepsilon^{-1} \widehat{g}\Big(\frac{\cdot - \ell - p}{\varepsilon}\Big) \widehat{f}(\ell)\, d\ell \Big\|_{L^2(s)} \leq D_{g,\kappa}\, \varepsilon \|f\|_{H^s} \quad (2.57)$$

for all $\varepsilon \in (0, \varepsilon_0)$.

Proof. The Lemma is a special case of Lemma 3.5 in [DSW16]. □

Lemma 2.2.8. *The normal-form transforms $\mathcal{T}_{j_1 j_2 j_3 j_4}$ were constructed such that for all $j_1, j_2, j_3, j_4 \in \{\pm 1\}$, we have*

$$\|\vartheta^{-1} Y_{j_1,j_2,j_3}\|_{L^2} \leq \mathcal{O}\big(\|R_{j_3}\|_{H^{\deg^*(\rho)+1}}\big). \quad (2.58)$$

where

$$\begin{aligned} Y_{j_1,j_2,j_3} = {} & N_{j_1 j_2}(\psi_c, j_2 \vartheta^{-1}\, i\rho(\psi \vartheta R_{j_3})) \\ & + \sum_{j_4=\pm 1} \Big(- j_1\, i\omega \mathcal{T}_{j_1 j_2 j_3 j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1 j_2 j_3 j_4}(-i\omega\psi_{j_4}, \psi_{j_4}, R_{j_3}) \\ & \qquad + \mathcal{T}_{j_1 j_2 j_3 j_4}(\psi_{j_4}, -i\omega\psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1 j_2 j_3 j_4}(\psi_{j_4}, \psi_{j_4}, j_3\, i\omega R_{j_3})\Big). \end{aligned} \quad (2.59)$$

Furthermore, for every fix functions g, h with $\widehat{g}, \widehat{h} \in L^1(\mathbb{R}, \mathbb{C})$, the mapping $f \mapsto \mathcal{T}_{jj_3}(g, h, f)$ defines a continuous linear map from $L^2(\mathbb{R}, \mathbb{C})$ into $L^2(\mathbb{R}, \mathbb{C})$ and there exists a constant $C = C\big(\|\widehat{g}\|_{L^1}\|\widehat{h}\|_{L^1}\big)$ such that for all $f \in L^2(\mathbb{R}, \mathbb{C})$, we have

$$\|\mathcal{T}_{j_1 j_2 j_3 j_4}(g, h, f)\|_{L^2} \leq C\|f\|_{L^2}\,. \quad (2.60)$$

Proof. When we are in the case that $0 \neq \pm\omega(0^+) \neq 2\omega(k_0)$ and therefore have $\mathcal{T}_{j_1 j_2 j_3 j_4} = 0$, (2.60) is trivially true and (2.58) is true due to lemma 2.2.5 and the fact that we have $\vartheta^{-1} = 1$ in this case.
We now prove the case, where we do not have $0 \neq \pm\omega(0^+) \neq 2\omega(k_0)$.

We first show that the normal-form transform $\mathcal{T}_{j_1j_2j_3j_4}$ is well-defined. Therefore, we look at the zeros of the denominator of $t_{j_1,j_2,j_3,j_4}(k)$, i.e. the zeros of

$$\big(\omega(k) - j_1j_2\omega(k-j_4k_0) + j_1\omega(j_4k_0)\big)\big(-j_1\omega(k) - 2\omega(j_4k_0) + j_3\omega(k-2j_4k_0)\big)$$

for $|k| \leq \delta$.
For the first factor, we have (1.12), so we know that the only possible zero of the first factor is $k = 0$.
For the second factor, we get by expanding the expression $\omega(k)$ in the point $\mathrm{sign}(k)\cdot 0^+$ and $\omega(k-2j_4k_0)$ in the point $-2j_4k_0$:

$$\begin{aligned} &-j_1\omega(k) - 2\omega(j_4k_0) + j_3\omega(k-2j_4k_0) \\ &\qquad = -j_1\omega(\mathrm{sign}(k)\cdot 0^+) - 2\omega(j_4k_0) + j_3\omega(-2j_4k_0) + \mathcal{O}(k). \end{aligned}$$

When $\omega(0) = 0$, we can choose δ such small that the second factor has no zeros due to (1.6).
Otherwise, we can choose δ such small that the second factor has no zeros due to (1.17).
To sum up, there can only occur a resonance in $k = 0$. Exactly as in the proof of lemma 2.2.5 we can see that the normal-form transforms $\mathcal{T}_{j_1j_2j_3j_4}$ can have, at worst, a trivial resonance in $k = 0$.

We now obtain (2.60) by using Young's inequality for convolutions and the fact that $\|\widehat{t}_{j_1,j_2,j_3,j_4}\|_{L^\infty}$ can be uniformly bounded

$$\|\mathcal{T}_{j_1j_2j_3j_4}(g,h,R_{j_3})\|_{L^2} \leq \|t_{j_1,j_2,j_3,j_4}\|_{L^\infty}\|\widehat{g}\|_{L^1}\|\widehat{h}\|_{L^1}\|R_{j_3}\|_{L^2} \leq C\|R_{j_3}\|_{L^2}.$$

We will now show (2.58).
As a first step, we will prove

$$\Big\|\vartheta^{-1}\Big(N_{j_1j_2}\big(\psi_c, j_2\vartheta^{-1}i\rho(\psi\vartheta R_{j_3})\big) - \sum_{j_4=\pm1} P_{0,\delta}N_{j_1j_2}\big(\psi_{j_4}, j_2\vartheta^{-1}i\rho(\psi_{j_4}\vartheta R_{j_3})\big)\Big)\Big\|_{L^2} \tag{2.61}$$

$$= \mathcal{O}\big(\,\|R_{j_3}\|_{H^{\deg^*(\rho)+1}}\big).$$

By exploiting the fact that $\vartheta^{-1} = P_{0,\delta}\vartheta^{-1} + P_{\delta,\infty}$, $\psi = \psi_c + \varepsilon\psi_Q$ and $\psi_c = \psi_1 + \psi_{-1}$, we get

$$\begin{aligned} &\vartheta^{-1}N_{j_1j_2}(\psi_c, j_2\vartheta^{-1}\,i\rho(\psi\vartheta R_{j_3})) \\ &\quad = \sum_{j_4=\pm1}\Big(P_{0,\delta}\vartheta^{-1}N_{j_1j_2}\big(\psi_{j_4}, j_2\vartheta^{-1}\,i\rho(\psi_{j_4}\vartheta R_{j_3})\big) \\ &\qquad\qquad + P_{0,\delta}\vartheta^{-1}N_{j_1j_2}\big(\psi_{j_4}, j_2\vartheta^{-1}\,i\rho(\psi_{-j_4}\vartheta R_{j_3})\big)\Big) \\ &\qquad + \varepsilon P_{0,\delta}\vartheta^{-1}N_{j_1j_2}\big(\psi_c, j_2\vartheta^{-1}\,i\rho(\psi_Q\vartheta R_{j_3})\big) + P_{\delta,\infty}N_{j_1j_2}\big(\psi_c, j_2\vartheta^{-1}\,i\rho(\psi\vartheta R_{j_3})\big). \end{aligned}$$

Using (2.32), (2.42) and (2.43), and (2.38), we see that the L^2-norm of the last two summands can be estimated against $\mathcal{O}(\|R_{j_3}\|_{H^{\deg^*(\rho)+1)}})$.
For the remaining summands, we have in Fourier space

$$\begin{aligned}
&\mathcal{F}\big[P_{0,\delta}\vartheta^{-1}N_{j_1j_2}(\psi_{j_4}, j_2\vartheta^{-1}\, i\rho(\psi_\ell\vartheta R_{j_3}))\big](k) \\
&\quad= \hat{P}_{0,\delta}(k)\int_{\mathbb{R}}\int_{\mathbb{R}} K_\varepsilon(k, k-m, m, n)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_\ell(m-n)\widehat{R_{j_3}}(n)\, dndm
\end{aligned}$$

where

$$K_\varepsilon(k, k-m, m, n) = j_2\,\frac{i\rho(k)\,\rho(m)\,\hat{\vartheta}(n)}{\hat{\vartheta}(k)\,\big(\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m)\big)}\,.$$

Please note that we could replace the term $\hat{\vartheta}_{\varepsilon,\infty}(m)\hat{\vartheta}^{-1}(m)$ by 1, since $|k| \le \delta$ and $|k - m - j_4k_0| \le \delta$ implies $|m| > k_0/2 > \varepsilon$.

For $\ell = -j_4$, we can apply Fubini's theorem, Young's inequality for convolutions and Lemma 2.2.7 to obtain

$$\begin{aligned}
&\|\mathcal{F}\big[P_{0,\delta}\vartheta^{-1}N_{j_1j_2}(\psi_{j_4}, j_2\vartheta^{-1}\, i\rho(\psi_{-j_4}\vartheta R_{j_3}))\big]\|_{L^2} \\
&\quad= \Big\|\hat{P}_{0,\delta}(\cdot)\int_{\mathbb{R}}\int_{\mathbb{R}} K_\varepsilon(\cdot, j_4k_0, \cdot - j_4k_0, \cdot)\widehat{\psi}_{j_4}(\cdot - m)\widehat{\psi}_{-j_4}(m-n)\widehat{R_{j_3}}(n)\, dndm\Big\|_{L^2} \\
&\qquad+ \mathcal{O}(\|R_{j_3}\|_{L^2})\,.
\end{aligned}$$

We could especially apply Lemma 2.2.7 since the Lipschitz continuity of the function K in some neighborhood of $\pm k_0$ with respect to its second argument, respectively its third argument, was sufficient due to the finite support of the integrand. Moreover, the fact that K_ε may has a jump discontinuity in $k = 0$ does not pose a problem since we could split $\hat{P}_{0,\delta}(k)$ accordingly into two characteristic functions. Now, since

$$\begin{aligned}
K_\varepsilon(k, j_4k_0, k - j_4k_0, k) &= j_2\,\frac{i\rho(k)\,\rho(k-j_4k_0)\,\hat{\vartheta}(k)}{\hat{\vartheta}(k)\,\big(\omega(k) - j_1j_2\omega(k-j_4k_0) + j_1\omega(j_4k_0)\big)} \\
&= j_2\,\frac{i\rho(k)\,\rho(k-j_4k_0)}{\big(\omega(k) - j_1j_2\omega(k-j_4k_0) + j_1\omega(j_4k_0)\big)}\,,
\end{aligned}$$

the term $K_\varepsilon(k, jk_0, k - jk_0, k)$ contains no factors which could be of order $\mathcal{O}(\varepsilon^{-1})$ such that

$$\hat{P}_{0,\delta}(k)\, K_\varepsilon(k, j_4k_0, k - j_4k_0, k) = \mathcal{O}(1)\,.$$

Hence, by using (2.11) and Young's inequality for convolutions, we obtain

$$\|\sum_{j_4=\pm1} P_{0,\delta}\vartheta^{-1}N_{j_1j_2}(\psi_{j_4}, j_2\vartheta^{-1}\, i\rho(\psi_{-j_4}\vartheta R_{j_3}))\|_{L^2} = \mathcal{O}(\|R_{j_3}\|_{L^2})$$

and thus have verified (2.61).

Due to (2.61), we now have

$$\|\vartheta^{-1}Y_{j_1,j_2,j_3}\|_{L^2} \leq \|\vartheta^{-1}\tilde{Y}_{j_1,j_2,j_3}\|_{L^2} + \mathcal{O}\big(\, \|R_{j_3}\|_{H^{\deg^*(\rho)+1}}\big)\,,$$

where

$$\begin{aligned}\vartheta^{-1}\tilde{Y}_{j_1,j_2,j_3} = \sum_{j_4=\pm 1} \Big(& P_{0,\delta}\vartheta^{-1}N_{j_1j_2}(\psi_{j_4}, j_2\vartheta^{-1}\, i\rho(\psi_{j_4}\vartheta R_{j_3})) \\ & - j_1\, i\omega\vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3}) + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(-i\omega\psi_{j_4},\psi_{j_4},R_{j_3}) \\ & + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4},-i\omega\psi_{j_4},R_{j_3}) + \vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4},\psi_{j_4}, j_3\, i\omega R_{j_3})\Big)\,.\end{aligned}$$

In Fourier space, we have

$$\begin{aligned}&\hat{\vartheta}^{-1}(k)\widehat{\tilde{Y}}_{j_1,j_2,j_3}(k) \\ &= \sum_{j_4=\pm 1} \hat{P}_{0,\delta}(k) \int_{\mathbb{R}}\int_{\mathbb{R}} K_\varepsilon(k,k-m,m,n)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R_{j_3}}(n)\,dndm \\ &\quad + \sum_{j_4=\pm 1} \hat{P}_{0,\delta}(k) \int_{\mathbb{R}}\int_{\mathbb{R}} \check{K}_\varepsilon(k,k-m,m-n,n)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R_{j_3}}(n)\,dndm\,,\end{aligned}$$

where

$$\begin{aligned}&\check{K}_\varepsilon(k,k-m,m-n,n) \\ &\qquad = \hat{\vartheta}^{-1}(k)\,\widehat{t}_{j_1,j_2,j_3,j_4}(k)\,\big(-j_1\, i\omega(k) - i\omega(k-m) - i\omega(m-n) + j_3\, i\omega(n)\big)\end{aligned}$$

and K_ε is as above.
We can exploit Lemma 2.2.7 together with Fubini's theorem and Young's inequality for convolutions in order to obtain

$$\begin{aligned}&\hat{\vartheta}^{-1}(k)\widehat{\tilde{Y}}_{j_1,j_2,j_3}(k) \\ &= \sum_{j_4=\pm 1} \int_{\mathbb{R}}\int_{\mathbb{R}} \hat{P}_{0,\delta}(k)\, K_\varepsilon(k, j_4k_0, k-j_4k_0, k-2j_4k_0)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R_{j_3}}(n)\,dndm \\ &\quad + \sum_{j_4=\pm 1} \int_{\mathbb{R}}\int_{\mathbb{R}} \hat{P}_{0,\delta}(k)\, \check{K}_\varepsilon(k, j_4k_0, j_4k_0, k-2j_4k_0)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R_{j_3}}(n)\,dndm \\ &\qquad + \mathcal{O}(\|R_{j_3}\|_{L^2})\,.\end{aligned}$$

Here, we again only needed the Lipschitz continuity of the function $\check{K}_\varepsilon$ with respect to the respective arguments on some bounded sets, for the application of Lemma 2.2.7 due to the finite support of $\widehat{\psi}_{j_4}$ and the presence of $\hat{P}_{0,\delta}$.

Since $\widehat{t}_{j_1,j_2,j_3,j_4}(k)$ was constructed such that

$$\hat{P}_{0,\delta}(k)\check{K}_\varepsilon(k, j_4k_0, j_4k_0, k-2j_4k_0) = -\,\hat{P}_{0,\delta}(k)K_\varepsilon(k, j_4k_0, k-j_4k_0, k-2j_4k_0)\,,$$

the two integral kernels, which could be both of order $\mathcal{O}(\varepsilon^{-1})$, cancel each other out such that we get

$$\|\vartheta^{-1}\tilde{Y}_{j_1,j_2,j_3}\|_{L^2} = \mathcal{O}(\|R_{j_3}\|_{L^2})\,.$$

□

Lemma 2.2.9. *For all $m \geq \deg^*(\rho)+1$, we have*

$$\partial_t E_0 \leq \varepsilon^2\,\mathcal{O}(\varepsilon^{1/2}\,\mathcal{G}_m^{3/2} + \mathcal{G}_m + 1)\,, \tag{2.62}$$

where $\mathcal{G}_m := \|R_{-1}\|_{H^m}^2 + \|R_1\|_{H^m}^2$ and E_0 is as in (2.36).

Proof. Exploiting the skew symmetry of $i\omega$, Cauchy-Schwarz and the estimates (2.42), (2.43) and (2.60), we obtain

$$\begin{aligned}
\partial_t E_0 &= \sum_{j=\pm1}\int_{\mathbb{R}} \overline{\check{R}_j}\,\partial_t\check{R}_j + \check{R}_j\,\partial_t\overline{\check{R}_j}\,dx \\
&= \sum_{j=\pm1}\int_{\mathbb{R}} \overline{\check{R}_j}\,\partial_t\check{R}_j + \check{R}_j\,\partial_t\overline{\check{R}_j} - \overline{\check{R}_j}\,ji\omega\check{R}_j - \check{R}_j\,ji\omega\overline{\check{R}_j}\,dx \\
&\leq 2\sum_{j=\pm1}\|\check{R}_j\|_{L^2}\|\partial_t\check{R}_j - ji\omega\check{R}_j\|_{L^2} \\
&\leq \mathcal{O}(\sqrt{\mathcal{G}_1})\sum_{j=\pm1}\|\partial_t\check{R}_j - ji\omega\check{R}_j\|_{L^2},
\end{aligned}$$

where

$$\partial_t\check{R}_j = \partial_t R_j + \varepsilon\sum_{j_2\in\{\pm1\}}\vartheta^{-1}\partial_t N_{jj_2}(\psi_c, R_{j_2}) + \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}}\vartheta^{-1}\partial_t\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3}).$$

Due to

$$\begin{aligned}
j\,i\omega R_j = j\,i\omega\check{R}_j &- \varepsilon\sum_{j_2\in\{\pm1\}} j\,i\omega\vartheta^{-1}N_{jj_2}(\psi_c, R_{j_2}) \\
&- \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}} j\,i\omega\vartheta^{-1}\mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3})\,,
\end{aligned}$$

we get

$$
\begin{aligned}
\partial_t \check{R}_j =& j\, i\omega \check{R}_j \\
&+ \varepsilon \vartheta^{-1}\Big(\sum_{j_2 \in \{\pm 1\}} j\, i\rho(\psi\, \vartheta R_{j_2}) \\
&\qquad + \sum_{j_2 \in \{\pm 1\}} \Big(-j\, i\omega N_{jj_2}(\psi_c, R_{j_2}) + N_{jj_2}(\partial_t \psi_c, R_{j_2}) + N_{jj_2}(\psi_c, j_2\, i\omega R_{j_2})\Big)\Big) \\
&+ \varepsilon^2 \vartheta^{-1}\Big(\sum_{j_2, j_3 \in \{\pm 1\}} N_{jj_2}(\psi_c, j_2\, i\rho\vartheta^{-1}(\psi\, \vartheta R_{j_3})) \\
&\qquad + \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \Big(-j\, i\omega \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) + \mathcal{T}_{jj_2j_3j_4}(\partial_t \psi_{j_4}, \psi_{j_4}, R_{j_3}) \\
&\qquad\qquad + \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \partial_t \psi_{j_4}, R_{j_3}) + \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, j_3\, i\omega R_{j_3})\Big)\Big) \\
&+ \frac{\varepsilon^\beta}{2}\, j\, i\rho\vartheta^{-1}(\vartheta R_{-1} + \vartheta R_1)^2 + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi) \\
&+ \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}\big(\psi_c, \frac{\varepsilon^\beta}{2} j_2\, i\rho\vartheta^{-1}(\vartheta R_{-1} + \vartheta R_1)^2 + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_2}}(\varepsilon\Psi)\big) \\
&+ \varepsilon^3 \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}\Big(\psi_{j_4}, \psi_{j_4}, j_3\, i\rho\vartheta^{-1}\big(R_\psi\, (\vartheta R_{-1} + \vartheta R_1)\big)\Big) \\
&+ \varepsilon^2 \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4}, \psi_{j_4}, \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_3}}(\varepsilon\Psi)\big).
\end{aligned}
$$

By construction of our normal-form transforms, i.e. due to (2.40) and (2.41), and, (2.58) and (2.59), we obtain

$$
\begin{aligned}
\partial_t \check{R}_j =& j\, i\omega \check{R}_j \\
&+ \varepsilon \vartheta^{-1} \sum_{j_2 \in \{\pm 1\}} N_{jj_2}(\partial_t \psi_c + i\omega\psi_c, R_{j_2}) \\
&+ \varepsilon^2 \vartheta^{-1} \Big(\sum_{j_2, j_3, j_4 \in \{\pm 1\}} \mathcal{T}_{jj_2j_3j_4}(\partial_t \psi_{j_4} + i\omega \psi_{j_4}, \psi_{j_4}, R_{j_3}) \\
&\qquad + \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \partial_t \psi_{j_4} + i\omega\psi_{j_4}, R_{j_3}) \Big) \\
&+ \frac{\varepsilon^\beta}{2} j\, i\rho \vartheta^{-1} (\vartheta R_{-1} + \vartheta R_1)^2 + \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_j}(\varepsilon \Psi) \\
&+ \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}\big(\psi_c, \frac{\varepsilon^\beta}{2} j_2\, i\rho\vartheta^{-1}(\vartheta R_{-1} + \vartheta R_1)^2 + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_2}}(\varepsilon\Psi)\big) \\
&+ \varepsilon^3 \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}\Big(\psi_{j_4}, \psi_{j_4}, j_3\, i\rho\vartheta^{-1}\big(R_\psi\,(\vartheta R_{-1} + \vartheta R_1)\big)\Big) \\
&+ \varepsilon^2 \sum_{j_2, j_3, j_4 \in \{\pm 1\}} \vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4}, \psi_{j_4}, \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_3}}(\varepsilon\Psi)\big) \\
&+ \mathcal{O}(\varepsilon^2)\sqrt{\mathcal{G}_m}.
\end{aligned}
$$

Due to the bound (2.12) for $\partial_t \psi_{\pm 1} + i\omega\psi_{\pm 1}$, we obtain that the L^2-Norms of the second, third and forth term are $\mathcal{O}(\varepsilon^2)\sqrt{\mathcal{G}_1}$ by using the estimates (2.32), and (2.42), (2.43) and (2.60).
Due to our choice of $\beta = 5/2$ and Ψ, i.e. due to (2.38), and, (2.32) and (2.9), the L^2-Norm of the fifth and sixth term are bounded by $\mathcal{O}(\varepsilon^2)\big(\varepsilon^{1/2}\,\mathcal{G}_m + 1\big)$.
Now, we also see, by using the estimates (2.32), (2.42), (2.43) and (2.60) that the L^2-Norms of the last three terms are bounded by $\mathcal{O}(\varepsilon^2)\big(\varepsilon^{1/2}\mathcal{G}_m + 1\big)$.

We now obtain

$$
\begin{aligned}
\partial_t E_0 &\leq \mathcal{O}(\sqrt{\mathcal{G}_1}) \sum_{j=\pm 1} \|\partial_t \check{R}_j - j i\omega \check{R}_j\|_{L^2} \\
&\leq \varepsilon^2\, \mathcal{O}(\varepsilon^{1/2}\, \mathcal{G}_m^{3/2} + \mathcal{G}_m + \sqrt{\mathcal{G}_1} + 1) \\
&\leq \varepsilon^2\, \mathcal{O}(\varepsilon^{1/2}\, \mathcal{G}_m^{3/2} + \mathcal{G}_m + 1).
\end{aligned}
$$

□

2.2.2 Preserving regularity via a modified energy method

In the last subsection, we successfully obtained

$$\partial_t\big(\|\check{R}_{-1}\|_{L^2}^2 + \|\check{R}_1\|_{L^2}^2\big) \leq \varepsilon^2\, \mathcal{O}(\varepsilon^{1/2}\, \mathcal{G}_m^{3/2} + \mathcal{G}_m + 1)\,,$$

where $\mathcal{G}_m := \|R_{-1}\|_{H^m}^2 + \|R_1\|_{H^m}^2$ and $m \geq \deg^*(\rho) + 1$.
That means, when we assume

$$\|R_{-1}\|_{H^m}^2 + \|R_1\|_{H^m}^2 = \mathcal{O}(1)\,,$$

we have

$$\partial_t\big(\|\check{R}_{-1}\|_{L^2}^2 + \|\check{R}_1\|_{L^2}^2\big) \leq \varepsilon^2\, \mathcal{O}(\|R_{-1}\|_{H^m}^2 + \|R_1\|_{H^m}^2 + 1)\,.$$

If we could also obtain the estimates

$$\partial_t\big(\|\check{R}_{-1}\|_{H^s}^2 + \|\check{R}_1\|_{H^s}^2\big) \leq \varepsilon^2\, \mathcal{O}(\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2 + 1) \tag{2.63}$$

and

$$\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2 \leq c\,\big(\|\check{R}_{-1}\|_{H^s}^2 + \|\check{R}_1\|_{H^s}^2\big) \leq C\,\big(\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2\big) \tag{2.64}$$

for some $c, C > 0$, an application of Gronwall's inequality would yield

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|R_{-1}(t)\|_{H^s} + \|R_1(t)\|_{H^s} \leq \check{C}.$$

However, we have the following issues.

- The energies $\big(\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2\big)$ and $\big(\|\check{R}_{-1}\|_{H^s}^2 + \|\check{R}_1\|_{H^s}^2\big)$ are in general not equivalent, i.e. (2.64) is in general not true. This is since the normal form transformations $N_{j_1j_2}$ can lose regularity. To be more precise, we can in general not obtain the estimate

  $$\big(\|\check{R}_{-1}\|_{H^s}^2 + \|\check{R}_1\|_{H^s}^2\big) \leq C\,\big(\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2\big)$$

  since

  $$\|\check{R}_j\|_{H^s}^2 = \varepsilon^2\,\Big\|\sum_{j_2\in\{\pm1\}} \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2})\Big\|_{H^s}^2 + \mathcal{O}\big(\|R_{-1}\|_{H^s}^2 + \|R_1\|_{H^s}^2\big)$$

  and $\|N_{jj}(\psi_c, R_j)\|_{H^s}$ can only be estimated against the $H^{s'}$-norm of R_j when

  $$s' \geq s + 1 + \deg^*(\rho) - \deg(\omega).$$

- As we have already seen in the proof of (2.62), we can in general not obtain (2.63). This is since

  $$\partial_t \check{R}_j = ji\omega\check{R}_j + \varepsilon^2\, h_j(R_{-1}, R_1)\,,$$

  where the mapping h_j only maps $H^{s+r}(\mathbb{R}) \times H^{s+r}(\mathbb{R})$ onto $H^s(\mathbb{R})$ when

  $$r \geq 2\deg^*(\rho) - \max\{0, \deg(\omega) - 1\}.$$

We will now explain the key idea for addressing these issues.
Looking at

$$\|\check{R}_j\|_{L^2}^2 = \Big\| R_j + \varepsilon \sum_{j_2\in\{\pm1\}} \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) + \varepsilon^2 \sum_{j_2,j_3,j_4\in\{\pm1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) \Big\|_{L^2}^2,$$

we see that the most problematic term for our estimates is

$$\varepsilon^2 \Big\| \sum_{j_2\in\{\pm1\}} \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) \Big\|_{L^2}^2.$$

It is the reason, we can in general not obtain (2.64) and on top of that produces the terms with the most derivatives falling on R_{-1} or R_1 in the evolution of $\|\check{R}_j\|_{L^2}^2$.
However, when we look at

$$\begin{aligned}&\|\partial_x \check{R}_j\|_{L^2}^2 \\ &= \Big\| \partial_x R_j + \varepsilon \sum_{j_2\in\{\pm1\}} \partial_x\vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) + \varepsilon^2 \sum_{j_2,j_3,j_4\in\{\pm1\}} \partial_x\vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) \Big\|_{L^2}^2,\end{aligned}$$

for $\ell \geq 1$, we formally have

$$\|\partial_x \check{R}_j\|_{L^2}^2 = \big\|\partial_x^\ell R_{j_1}\big\|_{L^2}^2 + 2\varepsilon \sum_{j_2\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2})\, dx + \mathcal{O}(\varepsilon^2)$$

due to (2.37).
The temporal derivative of the terms of the order $\mathcal{O}(\varepsilon^2)$ will also be of the order $\mathcal{O}(\varepsilon^2)$, such that they can only influence terms of the order $\mathcal{O}(\varepsilon^2)$ in the evolution of $\|\partial_x \check{R}_j\|_{L^2}^2$. That means that these terms of the formal order $\mathcal{O}(\varepsilon^2)$, which include the problematic term

$$\varepsilon^2 \Big\| \sum_{j_2\in\{\pm1\}} \partial_x^\ell \vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) \Big\|_{L^2}^2,$$

are not required to obtain a $\mathcal{O}(\varepsilon^{-2})$-timescale and therefore redundant.
For this very reason, instead of working with the energy $\big(\|\check{R}_{-1}\|_{H^\ell}^2 + \|\check{R}_1\|_{H^\ell}^2\big)$, we use the modified energy:

Definition 2.2.10.

$$\mathcal{E}_\ell = E_0 + E_\ell\,, \tag{2.65}$$

$$E_\ell = \sum_{j_1\in\{\pm1\}} \Big(\frac{1}{2}\big\|\partial_x^\ell R_{j_1}\big\|_{L^2}^2 + \varepsilon \sum_{j_2\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \partial_x^\ell \vartheta^{-1} N_{j_1j_2}(\psi_c, R_{j_2})\, dx \Big),$$

where $\ell \geq 1$, and E_0, $N_{j_1j_2}$ and ϑ^{-1} are exactly as in the last subsection.
I.e. E_0 is as in (2.36), $N_{j_1j_2}$ is as in (2.31) and ϑ^{-1} is the inverse of the operator ϑ defined by (2.30).

We can now show that this modified energy is equivalent to the energy

$$\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2$$

and its evolution only contains $\mathcal{O}(\varepsilon^2)$-terms. In other words, we solve the first issue while preserving the $\mathcal{O}(\varepsilon^{-2})$-timescale.
We also address the second issue since the terms, which can potentially have the most derivatives in the evolution of $\left(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2\right)$, do no longer occur in the evolution of $\mathcal{E}_\ell$. However the second issue is not completely solved by using the above energy, such that we will have to further modify this energy in the next subsection.

We will now prove the statements we made so far.

We need the following lemma, which can be understood as some generalization of integration by parts.

Lemma 2.2.11. *Let $f, g, h \in L^2(\mathbb{R}, \mathbb{R})$ be real-valued functions and $K : \mathbb{R}^3 \to \mathbb{C}$. If*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left| K(k, k-m, m)\, \overline{\widehat{f}(k)}\, \widehat{h}(k-m)\, \widehat{g}(m) \right| dm\, dk < \infty,$$

then we have

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\mathbb{R}} K(k, k-m, m)\, \overline{\widehat{f}(k)}\, \widehat{h}(k-m)\, \widehat{g}(m)\, dm\, dk \\ &\qquad = \int_{\mathbb{R}} \int_{\mathbb{R}} K(-m, k-m, -k)\, \overline{\widehat{g}(k)}\, \widehat{h}(k-m)\, \widehat{f}(m)\, dm\, dk. \end{aligned} \tag{2.66}$$

Proof. The result is obtained by first exploiting the fact that $\overline{\widehat{f}(k)} = \widehat{f}(-k)$, $\overline{\widehat{g}(k)} = \widehat{g}(-k)$ then making a change of variables and using Fubini's theorem

$$\begin{aligned} &\int_{\mathbb{R}} \int_{\mathbb{R}} K(k, k-m, m)\, \overline{\widehat{f}(k)}\, \widehat{h}(k-m)\, \widehat{g}(m)\, dm\, dk \\ &\qquad = \int_{\mathbb{R}} \int_{\mathbb{R}} K(k, k-m, m)\, \widehat{f}(-k)\, \widehat{h}(k-m)\, \overline{\widehat{g}(-m)}\, dm\, dk \\ &\qquad = \int_{\mathbb{R}} \int_{\mathbb{R}} K(-m, k-m, -k)\, \widehat{f}(m)\, \widehat{h}(k-m)\, \overline{\widehat{g}(k)}\, dk\, dm \\ &\qquad = \int_{\mathbb{R}} \int_{\mathbb{R}} K(-m, k-m, -k)\, \widehat{f}(m)\, \widehat{h}(k-m)\, \overline{\widehat{g}(k)}\, dm\, dk. \end{aligned}$$

□

Remark 2.2.12. *By introducing the notation*

$$\widehat{N}^*_{j_1 j_2}(h, f)(k) := \int_{\mathbb{R}} n_{j_1 j_2}(-m, k-m, -k) \widehat{h}(k-m) \widehat{f}(m)\, dm\,, \tag{2.67}$$

we now have

$$\int_{\mathbb{R}} N_{j_1j_2}(h,g)\, f\, dx = \int_{\mathbb{R}} N^*_{j_1j_2}(h,f)\, g\, dx \tag{2.68}$$

for $h, f, g \in H^1(\mathbb{R},\mathbb{R})$.

Lemma 2.2.13. *We have*

$$\int_{\mathbb{R}} f\, N_{jj}(h,f)\, dx \leq \mathcal{O}\big(\|\widehat{h}_c\|_{L^1}\big)\, \|f\|^2_{L^2}\,, \tag{2.69}$$

where $\widehat{h}_c(k) := \chi_c(k)\widehat{h}(k)$.

Proof. Exploiting lemma 2.2.11, i.e. using (2.67) and (2.68), we have

$$\int_{\mathbb{R}} f\, N_{jj}(h,f)\, dx = \frac{1}{2}\int_{\mathbb{R}} f\, \big(N_{jj}(h,f) + N^*_{jj}(h,f)\big)\, dx\,.$$

Due to the skew-symmetry of ρ and ω, we have for $|k| \to \infty$:

$$n_{jj}(k, k-m, m) + n_{jj}(-m, k-m, -k) = \frac{\rho(k) - \rho(m)}{\omega(k) - \omega(m) + j\,\omega(k-m)}\, \chi_c(k-m).$$

After using Taylor to expand $\rho(k)$ in the point m and exploiting (2.52), (2.53) and (2.54), we get

$$\begin{aligned} &n_{jj}(k, k-m, m) + n_{jj}(-m, k-m, -k) \\ &\qquad = \frac{\rho'(m)(k-m) + \mathcal{O}\big(\rho''(m)\big)}{\omega(k) - \omega(m) + j\,\omega(k-m)}\, \chi_c(k-m) \qquad &&\text{for } |k| \to \infty \\ &\qquad = \mathcal{O}\big(\,\chi_c(k-m)\big) &&\text{for } |k| \to \infty\,, \end{aligned}$$

due to (1.11), (1.10) and most importantly (1.9).
We can now get the estimate (2.69) by using Cauchy-Schwarz together with the Plancherel theorem and Young's inequality.

□

Lemma 2.2.14. *There are constants* $C_0, \check{C}_0$ *such that the following estimates hold*

$$\sqrt{E_0} \leq C_0\, \big(\|R_1\|_{H^1} + \|R_{-1}\|_{H^1}\big), \tag{2.70}$$

$$\|R_1\|_{L^2} + \|R_{-1}\|_{L^2} \leq \check{C}_0\, \sqrt{E_0} + \varepsilon\, \mathcal{O}(\|R_{-1}\|_{L^2} + \|R_1\|_{L^2}). \tag{2.71}$$

Proof. Estimate (2.70) is a direct consequence of the triangle inequality and the estimates (2.32), (2.42), (2.43) and (2.60).

Let the operator $P_{a,b}$ be defined for all $f \in L^2(\mathbb{R})$ by

$$\widehat{P_{a,b}f}(k) = \begin{cases} \widehat{f}(k) & \text{when } a \leq |k| \leq b \\ 0 & \text{else.} \end{cases}$$

In order to prove (2.71), we define

$$R_j^0 := P_{0,\delta} R_j, \qquad \check{R}_j^0 := P_{0,\delta} \check{R}_j, \qquad R_j^1 := P_{\delta,\infty} R_j, \qquad \check{R}_j^1 := P_{\delta,\infty} \check{R}_j$$

and split $R_j = R_j^0 + R_j^1$ and $\check{R}_j = \check{R}_j^0 + \check{R}_j^1$, here $\check{R}_j$ is as in (2.36).

We first look at R_j^0.
Since $|k - m \pm k_0| \leq \delta$ and $|k| \leq \delta$ implies

$$|m \mp k_0| \leq 2\delta\,,$$

we have

$$P_{0,\delta}\vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) = P_{0,\delta}\vartheta^{-1} N_{jj_2}(\psi_c, P_{0,2k_0} R_{j_2}^1)$$

due to the nature of the compact support from $\widehat{\psi}_c$.
Exploiting this fact and using the triangle inequality, (2.32), (2.42), (2.43), and (2.60), we obtain

$$\begin{aligned} &\|R_j^0\|_{L^2} \\ &= \Big\|\check{R}_j^0 - \varepsilon \sum_{j_2 \in \{\pm 1\}} P_{0,\delta}\vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) - \varepsilon^2 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} P_{0,\delta}\vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}(\psi_c, \psi_c, R_{j_3})\Big\|_{L^2} \\ &\leq \|\check{R}_j^0\|_{L^2} + \mathcal{O}(1)\big(\|R_{-1}^1\|_{L^2} + \|R_1^1\|_{L^2}\big) + \mathcal{O}(\varepsilon)\big(\|R_{-1}\|_{L^2} + \|R_1\|_{L^2}\big). \end{aligned}$$

Now, we look at R_j^1.
Due to the definitions of $\mathcal{T}_{jj_2j_3j_4}$ and ϑ, we see

$$\begin{aligned} R_j^1 &= \check{R}_j^1 - \varepsilon \sum_{j_2 \in \{\pm 1\}} P_{\delta,\infty}\vartheta^{-1} N_{jj_2}(\psi_c, R_{j_2}) - \varepsilon^2 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} P_{\delta,\infty}\vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}(\psi_c, \psi_c, R_{j_3}) \\ &= \check{R}_j^1 - \varepsilon \sum_{j_2 \in \{\pm 1\}} P_{\delta,\infty} N_{jj_2}(\psi_c, R_{j_2}). \end{aligned}$$

Multiplying the equation with ${R_j}^1$ and then integrating over $\mathbb{R}$ in x, we get

$$\|{R_j}^1\|_{L^2}^2 = \int_{\mathbb{R}} {R_j}^1\, \check{R}_j^1\, dx + \varepsilon \sum_{j_2 \in \{\pm 1\}} \int_{\mathbb{R}} {R_j}^1\, P_{\delta,\infty} N_{jj_2}(\psi_c, R_{j_2})\, dx.$$

For $j_2 = -j$, we have

$$\varepsilon \int_{\mathbb{R}} R_j{}^1 \, P_{\delta,\infty} N_{j-j}(\psi_c, R_{j_2}) \, dx = \varepsilon \, \mathcal{O}(\|R_j^1\|_{L^2} \|R_{-j}\|_{L^2})$$

by using Cauchy-Schwarz and (2.43).
For $j_2 = j$, we have

$$\begin{aligned} \varepsilon \int_{\mathbb{R}} R_j{}^1 \, P_{\delta,\infty} N_{jj}(\psi_c, R_{j_2}) dx &= \varepsilon \int_{\mathbb{R}} R_j{}^1 \, N_{jj}(\psi_c, R_j) \, dx \\ &= \varepsilon \int_{\mathbb{R}} R_j{}^1 \, N_{jj}(\psi_c, R_j^1) \, dx + \varepsilon \, \mathcal{O}(\|R_j^1\|_{L^2} \|R_j^0\|_{L^2}) \end{aligned}$$

due to the definition of $P_{a,b}$, $R_j{}^1$ and R_j^0, the bilinearity of N_{jj} and estimate (2.42). With estimate (2.69), we then get

$$\varepsilon \int_{\mathbb{R}} R_j{}^1 \, P_{\delta,\infty} N_{jj}(\psi_c, R_{j_2}) dx = \varepsilon \, \mathcal{O}(\|R_j^1\|_{L^2} \|R_j\|_{L^2}) \, .$$

Now, we have

$$\|R_j{}^1\|_{L^2}^2 = \int_{\mathbb{R}} R_j{}^1 \, \check{R}_j^1 \, dx + \varepsilon \|R_j{}^1\|_{L^2} \, \mathcal{O}(\|R_{-1}\|_{L^2} + \|R_1\|_{L^2})$$

such that, with the help of Cauchy-Schwarz, we can obtain

$$\|R_j{}^1\|_{L^2} \leq \|\check{R}_j^1\|_{L^2} + \varepsilon \, \mathcal{O}(\|R_{-1}\|_{L^2} + \|R_1\|_{L^2}).$$

Combining the two inequalities for $\|R_j{}^0\|_{L^2}$ and $\|R_j{}^1\|_{L^2}$ finally proves the estimate (2.71). □

Lemma 2.2.15. (see lemma 4.4. in [DH18].) *Let $f \in H^\ell(\mathbb{R}, \mathbb{R})$ and $g \in H^{m+1}(\mathbb{R}, \mathbb{R})$ with $\ell, m \geq 0$. Then we have*

$$\int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^m \vartheta g \, dx = \int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^m g \, dx + \mathcal{O}(\|f\|_{L^2} \|g\|_{L^2}) \, , \tag{2.72}$$

$$\int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^{m+1} \vartheta^{-1} g \, dx = \int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^{m+1} g \, dx + \mathcal{O}(\|f\|_{L^2} \|g\|_{L^2}) \, . \tag{2.73}$$

Proof. Using the definition of ϑ and integration by parts, we get

$$\int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^m \vartheta g \, dx = \int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^m g \, dx + (-1)^\ell \int_{\mathbb{R}} f \, \partial_x^{\ell+m} P_{0,\delta}(\vartheta - 1) g \, dx \, ,$$

$$\int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^{m+1} \vartheta^{-1} g \, dx = \int_{\mathbb{R}} \partial_x^\ell f \, \partial_x^{m+1} g \, dx + (-1)^\ell \int_{\mathbb{R}} f \, \partial_x^{\ell+m+1} P_{0,\delta}(\vartheta^{-1} - 1) g \, dx \, ,$$

which yields (2.72) and, due to (2.37), also (2.73). □

Corollary 2.2.16. *Let $\varepsilon < \varepsilon_0$ and ε_0 be sufficiently small. For $\ell \geq 1$, the energy $\mathcal{E}_\ell$ is equivalent to $\big(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\big)^2$, i.e. there are constants $C_1, C_2 > 0$ such that*

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

Proof. We examine E_ℓ.
Thanks to (2.73), and, (2.42) and (2.43), we have

$$\begin{aligned}
E_\ell &= \sum_{j_1 \in \{\pm 1\}} \Big(\frac{1}{2}\big\|\partial_x^\ell R_{j_1})\big\|_{L^2}^2 + \varepsilon \sum_{j_2 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2})\, dx\Big) \\
&= \sum_{j_1 \in \{\pm 1\}} \Big(\frac{1}{2}\big\|\partial_x^\ell R_{j_1})\big\|_{L^2}^2 + \varepsilon \sum_{j_2 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell N_{j_1 j_2}(\psi_c, R_{j_2})\, dx\Big) \\
&\quad + \varepsilon\,\mathcal{O}\Big(\big(\|R_{-1}\|_{H^1} + \|R_1\|_{H^1}\big)^2\Big)
\end{aligned}$$

such that we only have to look at the regularity of the terms

$$\int_{\mathbb{R}} \partial_x^\ell R_{j_1} \partial_x^\ell N_{j_1 j_2}(\psi_c, R_{j_2})\, dx.$$

For $(j_1, j_2) = (j, -j)$, we can see by using Cauchy-Schwarz and (2.43) that

$$\int_{\mathbb{R}} \partial_x^\ell R_j\, \partial_x^\ell N_{j-j}(\psi_c, R_{-j})\, dx = \mathcal{O}(\|R_{-1}\|_{H^\ell}\|R_1\|_{H^\ell}).$$

For $(j_1, j_2) = (j, j)$ however, we can have an additional derivative falling on $\partial_x^\ell R_j$ due to (1.9). By using Leibniz's rule, Cauchy-Schwarz and (2.42), we see

$$\int_{\mathbb{R}} \partial_x^\ell R_j\, \partial_x^\ell N_{jj}(\psi_c, R_j)\, dx = \int_{\mathbb{R}} \partial_x^\ell R_j\, N_{jj}(\psi_c, \partial_x^\ell R_j)\, dx + \mathcal{O}(\|R_j\|_{H^\ell}^2)$$

Thanks to (2.69), we get

$$\int_{\mathbb{R}} \partial_x^\ell R_j \partial_x^\ell N_{jj}(\psi_c, R_j)\, dx = \mathcal{O}(\|R_j\|_{H^\ell}^2).$$

We now obtained

$$E_\ell = \frac{1}{2}\big(\|\partial_x^\ell R_{-1}\|_{L^2}^2 + \|\partial_x^\ell R_1\|_{L^2}^2\big) + \varepsilon\,\mathcal{O}\big(\big(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\big)^2\big).$$

and the statement follows with lemma 2.2.14.

□

Lemma 2.2.17. *For $\ell \geq 1$, we have*

$$\partial_t E_\ell = \varepsilon^2 V_\ell + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1), \tag{2.74}$$

where

$$\begin{aligned} V_\ell = & \sum_{j_1, j_2 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, i\rho \partial_x^\ell \vartheta^{-1}(R_Q \vartheta R_{j_2}) \, dx \\ & + \sum_{j_1, j_2, j_3 \in \{\pm 1\}} \Big(j_1 \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3}) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\ & \qquad + j_2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, i\rho \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big) \, dx \Big), \end{aligned} \tag{2.75}$$

and

$$R_Q = \psi_Q + \frac{1}{2} \varepsilon^{\beta - 2}(\vartheta R_{-1} + \vartheta R_1). \tag{2.76}$$

Remark 2.2.18. *Due to* (2.73) *and* (2.38) *the term $\varepsilon^2 V_\ell$ indeed has the desired ε^2-order.*

Proof. We have

$$\begin{aligned} \partial_t E_\ell = & \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_t \partial_x^\ell R_{j_1} \, dx \\ & + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_t \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\ & \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} \partial_t N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \Big). \end{aligned}$$

Using the error equations (2.19) and exploiting

$$R_\psi = \psi_c + \varepsilon R_Q,$$

we get

$$
\begin{aligned}
\partial_t E_\ell = & \sum_{j_1 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, i\omega \partial_x^\ell R_{j_1} \, dx \\
& + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \Big(j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, i\rho \partial_x^\ell \vartheta^{-1} (\psi_c \vartheta R_{j_2}) \, dx \\
& \qquad + j_1 \int_{\mathbb{R}} i\omega \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad + j_2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, i\omega R_{j_2}) \, dx \\
& \qquad - \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(i\omega \psi_c, R_{j_2}) \, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\partial_t \psi_c + i\omega \psi_c, R_{j_2}) \, dx \Big) \\
& + \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, i\rho \partial_x^\ell \vartheta^{-1} (R_Q \vartheta R_{j_2}) \, dx \\
& + \varepsilon^2 \sum_{j_1, j_2, j_3 \in \{\pm 1\}} \Big(j_1 \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3}) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad + j_2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, i\rho \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big) \, dx \Big) \\
& + \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon \Psi) \, dx \\
& + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon \Psi) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_{j_2}}(\varepsilon \Psi)\big) \, dx \Big)
\end{aligned}
$$

where

$$R_Q = \psi_Q + \frac{1}{2} \varepsilon^{\beta - 2} (\vartheta R_{-1} + \vartheta R_1).$$

Exploiting the skew symmetry of $i\omega$ in the third integral and then using (2.40) and the definition (2.75), we get

$$
\begin{aligned}
\partial_t E_\ell = \varepsilon \sum_{j_1,j_2\in\{\pm1\}} & \Big(j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho \partial_x^\ell \vartheta^{-1}\big(\psi_c(\vartheta-\vartheta_{\varepsilon,\infty})R_{j_2}\big)\, dx \\
&\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\partial_t\psi_c + i\omega\psi_c, R_{j_2})\, dx \Big) \\
& + \varepsilon^2 V_\ell \\
& + \sum_{j_1\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \varepsilon^{-\beta}\partial_x^\ell\vartheta^{-1}\mathrm{Res}_{u_{j_1}}(\varepsilon\Psi)\, dx \\
& + \varepsilon \sum_{j_1,j_2\in\{\pm1\}} \Big(\int_{\mathbb{R}} \varepsilon^{-\beta}\partial_x^\ell\vartheta^{-1}\mathrm{Res}_{u_{j_1}}(\varepsilon\Psi)\, \partial_x^\ell\vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2})\, dx \\
&\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_2}}(\varepsilon\Psi)\big)\, dx \Big)\,.
\end{aligned}
$$

We now show that all terms except the term $\varepsilon^2 V_\ell$ can be estimated against $\varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell+1)$. Thereby we will especially take advantage of corollary 2.2.16 and (2.11).
For the first integral, we can use (2.73), Cauchy-Schwarz and the fact that

$$
\big(\hat\vartheta(k) - \hat\vartheta_{\varepsilon,\infty}(k)\big) = \begin{cases} \varepsilon + (1-\varepsilon)\frac{|k|}{\delta} & \text{when } 0 \neq \pm\omega(0^+) \neq 2\omega(k_0) \text{ and } |k|\leq\varepsilon\,, \\ 0 & \text{else}\,, \end{cases}
$$

in order to get

$$
\begin{aligned}
\varepsilon \int_{\mathbb{R}} & \partial_x^\ell R_{j_1}\, i\rho\partial_x^\ell\vartheta^{-1}\big(\psi_c(\vartheta-\vartheta_{\varepsilon,\infty})R_{j_2}\big)\, dx \\
&\leq \varepsilon\,\mathcal{O}\big(\|R_{j_1}\|_{H^\ell}\, \|i\rho\big(\psi_c(\vartheta-\vartheta_{\varepsilon,\infty})R_{j_2}\big)\|_{H^\ell}\big) \\
&\leq \varepsilon\,\mathcal{O}\big(\|R_{j_1}\|_{H^\ell}\, \|\psi_c\|_{C^{\ell+\deg^*(\rho)}}\, \|(\vartheta-\vartheta_{\varepsilon,\infty})R_{j_2}\|_{H^{\ell+\deg^*(\rho)}}\big) \\
&\leq \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell).
\end{aligned}
$$

The second integral in the above evolution equality is $\varepsilon^3\,\mathcal{O}(\mathcal{E}_\ell)$ due to the estimate (2.12). In order to see this we first use (2.73), then we proceed as in the proof of (2.2.16) in order to estimate without losing regularity.

The last three integrals are $\varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell+1)$ due to (2.9). To see this, we use first (2.73), then integration by parts to shift some derivatives away from $R_{\pm1}$, and finally Cauchy-Schwarz together with (2.42) and (2.43). Here, we also exploit the estimate $\sqrt{x}\leq|x|+1$ after using corollary 2.2.16.

□

2.2.3 Closing the error estimates via energy transformations

If, for some energy $\mathcal{E}$ and $c, C > 0$, we could obtain the estimates

$$\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2 \leq c\,\mathcal{E} \leq C\left(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2\right) \tag{2.77}$$

and

$$\partial_t \mathcal{E} \leq \varepsilon^2\,\mathcal{O}\big(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2 + 1\big)\,, \tag{2.78}$$

an application of Gronwall's inequality would yield that there is a $\check{C} > 0$ such that

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|R_{-1}(t)\|_{H^s} + \|R_1(t)\|_{H^s} \leq \check{C}.$$

In the last subsection, we successfully obtained (2.77) for the energy $\mathcal{E} = \mathcal{E}_\ell$. Moreover, we could show that the evolution of the energy $\mathcal{E}_\ell$ is of quadratic ε-order.
However, we still do in general not have (2.78) for the energy $\mathcal{E} = \mathcal{E}_\ell$. Instead, we only can get

$$\partial_t \mathcal{E}_\ell = \varepsilon^2\,V_\ell + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,,$$

where $\varepsilon^2 V_\ell$ contains integrals like

$$\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho\partial_x^\ell \vartheta^{-1}(R_Q\vartheta R_{j_2})\,dx = \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho\partial_x^\ell (R_Q\vartheta R_{j_2})\,dx + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,.$$

For $\deg^*(\rho) > 0$, these integrals cannot be estimated against $\mathcal{O}(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2 + 1)$ since there are to many derivatives falling on R_{-1} or R_1.
We will solve this problem by showing that there is an expression $\mathcal{D}$ with

$$\varepsilon^2\,\mathcal{D} = \varepsilon\,\mathcal{O}\big(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2\big)$$

such that

$$\varepsilon^2\,V_\ell - \varepsilon^2\,\partial_t\mathcal{D} = \varepsilon^2\,\mathcal{O}\big(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2 + 1\big)\,,$$

i.e.

$$\partial_t\mathcal{E}_\ell - \varepsilon^2\partial_t\mathcal{D} = \varepsilon^2\,\mathcal{O}\big(\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2 + 1\big)\,.$$

Then, by defining the final energy by

$$\tilde{\mathcal{E}}_\ell := \mathcal{E}_\ell - \varepsilon^2\mathcal{D}\,,$$

we obtain (2.77) and (2.78) for $\mathcal{E} = \tilde{\mathcal{E}}_\ell$ and can finally prove theorem 1.1.1.

Looking closely at the term

$$\sum_{j_1,j_2\in\{\pm1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho\partial_x^\ell(R_Q\vartheta R_{j_2})\,dx$$

of V_ℓ, one can see that it basically can be reduced to a sum consisting of three type of integrals, which have the following form.

a) The form

$$\int_{\mathbb{R}} \gamma \partial_x^\ell R_j \, \partial_x^\ell R_{-j} \, f \, dx \tag{2.79}$$

where γ is a symmetric or skew symmetric pseudo-differential operator with

$$\deg^*(\gamma) \leq \deg^*(\rho) \, . \tag{2.80}$$

b) The form

$$\int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_j \, f \, dx, \tag{2.81}$$

where $i\sigma$ is a skew symmetric pseudo-differential operator with

$$\deg^*(\sigma) \leq \deg^*(\rho) \, . \tag{2.82}$$

c) The form

$$\int_{\mathbb{R}} \upsilon \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x f \, dx, \tag{2.83}$$

where υ is a symmetric pseudo-differential operator with

$$\deg^*(\upsilon) \leq \deg^*(\rho) - 1 \, . \tag{2.84}$$

Here f is always a function whose relevant norms can be controlled well enough.

This partition is also possible for V_ℓ, in particular since we can replace the bilinear operators $N_{j_1 j_2}$ by a sums consisting of products of linear operators due to (2.52), (2.53), (2.54) and (2.55). The fact that (2.80), (2.82) and (2.84) can also be obtained for V_ℓ is more difficult to be seen directly and is related to some good cancellations happening. The happening of such cancellations however can already be expected as a consequence of lemma 2.2.16.

The idea now is to find an energy transformation, which exploits the linear part of our system to eliminate these three type of problematic integrals that were produced by the nonlinearity.

So, the core idea is similar to the one that was behind the normal form transformations. However, while the goal of the normal form transformations was to obtain the right ε-order for our estimates in order to achieve a $\mathcal{O}(\varepsilon^{-2})$ timescale, the goal of the energy transformation is to obtain the right Sobolev norms for our estimates, such that they can be closed and Gronwall can be even applied in the first place.

Our four key observations for finding the energy transformation

$$\mathcal{E}_\ell \to \mathcal{E}_\ell - \varepsilon^2 \mathcal{D}$$

are the following ones.

a) Exploiting

$$\partial_t R_j = j i\omega R_j + H_j \, ,$$

where H_j is defined according to (2.19), and taking advantage of the skew symmetry of $i\omega$, we have

$$\begin{aligned}
&\frac{1}{2} j\, \varepsilon^2\, \partial_t \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx \\
&\quad = \frac{1}{2}\, \varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx\; - \frac{1}{2}\, \varepsilon^2 \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, i\omega \partial_x^\ell R_{-j}\, f\, dx\; + J \\
&\quad = \int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx\; + \frac{1}{2} \varepsilon^2 \int_{\mathbb{R}} \big[i\omega, f\big] \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, dx\; + J \, ,
\end{aligned}$$

where the terms coming from H_j and $\partial_t f$ were collected in the expression J. Based on this observation, we will show for (2.79) :

$$\begin{aligned}
\int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx = \frac{1}{2} j\; \partial_t \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx \\
+ J_L + \varepsilon\, J_E + \mathcal{O}(\mathcal{E}_\ell + 1) \, ,
\end{aligned} \tag{2.85}$$

where J_L consists of integrals of the form (2.79) that contain less derivatives falling on R_{-1} or R_1 than the original one and J_E consists of integrals of the form (2.79), (2.81), (2.83) that contain not more derivatives falling on R_{-1} or R_1 than the original one.

b) Looking at (2.81), we observe

$$\int_{\mathbb{R}} i\sigma \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx = \frac{1}{2} \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx - \frac{1}{2} \int_{\mathbb{R}} \partial_x^\ell R_j\, i\sigma \big(\partial_x^\ell R_j f\big)\, dx \, ,$$

i.e.

$$\int_{\mathbb{R}} i\sigma \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx = -\frac{1}{2} \int_{\mathbb{R}} \big[i\sigma, f\big] \partial_x^\ell R_j\, \partial_x^\ell R_j\, dx \, . \tag{2.86}$$

We will show that the right hand side integral can be expressed as a sum of some $\mathcal{O}(\mathcal{E}_\ell + 1)$-terms and integrals of the form (2.79), (2.81), (2.83), which contain at least a whole derivative less falling on R_{-1} or R_1 than the original integral.

c) Exploiting

$$\partial_t R_j = j i\omega R_j + H_j \, ,$$

where H_j is defined according to (2.19), and taking advantage of the skew symmetry of $i\omega$, we get

$$
\begin{aligned}
&j\,\varepsilon^2\,\partial_t \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx \\
&\quad = \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} i\omega \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx + \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, i\omega \partial_x^\ell R_j\, f\, dx + J \\
&\quad = -\varepsilon^2 \int_{\mathbb{R}} \big[i\omega, f\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, \partial_x^\ell R_j\, dx + J\,,
\end{aligned}
$$

where all the terms coming from H_j and $\partial_t f$ were collected in the expression J.

Based on this observation, we will show

$$
\begin{aligned}
\int_{\mathbb{R}} \upsilon \partial_x^\ell R_j\, \partial_x^\ell R_j\, \partial_x f\, dx \;&= -j\,\partial_t \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, \partial_x^\ell R_j\, f\, dx \\
&\qquad + J_L + \varepsilon\, J_E + \mathcal{O}(\mathcal{E}_\ell + 1)\,,
\end{aligned}
\tag{2.87}
$$

where J_L consists of integrals of the form (2.79) that contain less derivatives falling on R_{-1} or R_1 than the original one and J_E consists of integrals of the form (2.79), (2.81), (2.83) that contain not more derivatives falling on R_{-1} or R_1 than the original one.

d) Looking back at our original system (1.2), we have $\partial_t u = -i\omega v$. This means that for the diagonalized system, we have $\partial_t(u_{-1} + u_1) = -i\omega(u_{-1} - u_1)$ and can therefore easily obtain from (2.19) that

$$
\partial_t(R_1 + R_{-1}) = i\omega(R_1 - R_{-1}) + \varepsilon^{-\beta}\vartheta^{-1}\big(\mathrm{Res}_{u_1}(\varepsilon\Psi) + \mathrm{Res}_{u_{-1}}(\varepsilon\Psi)\big)\,. \tag{2.88}
$$

Thus, we have for skew symmetric operators $i\sigma$ with $\deg^*(\sigma) \le \deg(\omega)$ that

$$
\begin{aligned}
&\int_{\mathbb{R}} i\sigma \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, f\, dx \\
&\quad = \frac{1}{2}\,\partial_t \int_{\mathbb{R}} \frac{\sigma}{\omega} \partial_x^\ell (R_1 + R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, f\, dx \\
&\qquad - \frac{1}{2} \int_{\mathbb{R}} \big[\frac{\sigma}{\omega}, f\big] i\omega \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, dx \\
&\qquad - \frac{1}{2} \int_{\mathbb{R}} \frac{\sigma}{\omega} \partial_x^\ell (R_1 + R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, \partial_t f\, dx + \mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}
$$

The first term on the above right hand side is a time derivative of an integral that can be estimated against $\mathcal{O}(\mathcal{E}_\ell)$ and the last two integrals contain less derivatives falling on R_{-1} or R_1 than the one on the left hand side.

Due to (1.9), we can use these results to recursively construct the expression $\mathcal{D}$. When $\deg^*(\rho) \leq 1$ or $\deg^*(\rho) < \deg(\omega)$, we obtain $\mathcal{D}$ after a finite number of steps. When $\deg^*(\rho) > 1$ and at the same time $\deg^*(\rho) = \deg(\omega)$, we obtain $\mathcal{D}$ as the sum of an absolutely convergent series.
In order to show that we can indeed find $\mathcal{D}$, we prove that every time we proceed as in a), b) or c):

[R1] the number of additional derivatives falling on $\partial_x^\ell R_{-1}$ or $\partial_x^\ell R_1$ does not increase,

[R2] we do not generate new problematic terms for which a), b) or c) cannot be applied,

[R3] the number of the emerging integrals only depends on $\deg^*(\rho)$ and $\deg(\omega)$,

[R4] the emerging integrals εJ_E get much smaller in size.

Finally, we then preform the energy transformation

$$\mathcal{E}_\ell \to \tilde{\mathcal{E}}_\ell = \mathcal{E}_\ell - \varepsilon^2 \mathcal{D}$$

such that (2.77) and (2.78) are true for $\mathcal{E} = \tilde{\mathcal{E}}_\ell$ and we can prove theorem 1.1.1.

Remark 2.2.19. *When the functions ω or ω' have zeros, operators like $\frac{\gamma}{\omega}$ and $\frac{\gamma}{\omega'}$ could be not well-defined, however we will show that we can assume without a loss of generality that the function γ is equal to zero on some set that includes all zeros of ω and ω'.*

Naturally, integrals of the form (2.83) *are harmless for* $\deg^*(\rho) \leq 1$. *In this case one can quickly construct an energy transformation by only relying on* d) *and* b). *This also makes the energy transformations in [D17, CW17] much simpler in comparison to here, where* $\deg^*(\rho)$ *is allowed to be arbitrarily large.*

In order to apply our framework to more general quasilinear dispersive systems, it should be sufficient that either $\partial_t(R_1 + R_{-1})$ or $\partial_t(R_1 - R_{-1})$ have only a nonlinearity that loses at most $\deg(\omega')$ *derivatives,* (2.88) *is not needed.*

We will now prove our claims, what will turn out to be a rather technical procedure.

In order to handle the commutators, we saw above, we use the following lemma.

Lemma 2.2.20. *Let $n \in \mathbb{N}$, and γ be a function of $C^{n+1}(\mathbb{R})$ with* $\deg^*(\gamma) \in \mathbb{R}$ *for which*

$$\deg^*(\gamma^{(l)}) \leq \deg^*(\gamma^{(l-1)}) - 1 \qquad \text{for all } 1 \leq l \leq n+1. \tag{2.89}$$

Moreover let the operators γ and $i^l \gamma^{(l)}$ be given by their symbols in Fourier space. Then we have for $f, g \in C_c^\infty(\mathbb{R})$:

$$\left[\gamma, g\right] f = \sum_{l=1}^{n} \frac{(-1)^l}{l!} \, \partial_x^l g \, i^l \gamma^{(l)} f + \mathcal{R}(f, g). \tag{2.90}$$

For the rest-term $\mathcal{R}(f,g)$, *given through*

$$\widehat{\mathcal{R}(f,g)} = \int_{\mathbb{R}} \Big(\frac{(\cdot - m)^{n+1}}{n!} \int_0^1 \gamma^{(n+1)}\big(m + (\cdot - m)x\big)\,(1-x)\,dx \Big)\,\widehat{g}(\cdot - m)\widehat{f}(m)\,dm\,,$$

we have the estimates

$$\|\mathcal{R}(f,g)\|_{L^2} = \mathcal{O}(1)\,\|\partial_x^{n+1} g\|_{H^{p+q}} \|f\|_{H^p}, \tag{2.91}$$

$$\|\mathcal{R}(f,g)\|_{L^2} = \mathcal{O}(1)\,\|\widehat{\partial_x^{n+1} g}\|_{L^1(p)} \|f\|_{H^p}, \tag{2.92}$$

for $q > \frac{1}{2}$ *and*

$$p := \max\{\deg^*(\gamma) - n - 1,\, 0\}.$$

Remark 2.2.21. *One can think of this lemma as some sort of generalization of Leibniz's rule. To give an example, let* $\gamma = \partial_x^2$.
Then we have $\gamma(k) = -k^2$, $\gamma'(k) = -2k$, $\gamma''(k) = -2$ *and* $\gamma^{(3)}(k) = 0$. *Thus,*

$$\begin{aligned}\big[\partial_x^2, g\big] f &= \sum_{l=1}^{2} \frac{(-1)^l}{l!}\,\partial_x^l g\, i^l \gamma^{(l)} f + \mathcal{R}(f,g) \;=\; \frac{-1}{1!}\,\partial_x g\, i\gamma' f + \frac{1}{2!}\,\partial_x^2 g\, i^2 \gamma'' f + 0\\ &= 2\partial_x g \partial_x f + \partial_x^2 g f\,.\end{aligned}$$

Proof. We have

$$\widehat{\big[\gamma, g\big] f} = \widehat{\gamma(gf)} - \widehat{g\gamma f} = \int_{\mathbb{R}} \big(\gamma(\cdot) - \gamma(m)\big)\widehat{g}(\cdot - m)\widehat{f}(m)\,dm.$$

Using Taylor, we get

$$\begin{aligned}\gamma(k) - \gamma(m) &= \sum_{l=1}^{n} \frac{(k-m)^l}{l!}\gamma^{(l)}(m) + r(k, k-m, m)\\ &= \sum_{l=1}^{n} \frac{i^l (k-m)^l}{l!}(-i)^l \gamma^{(l)}(m) + r(k, k-m, m),\end{aligned}$$

where

$$\begin{aligned}r(k, k-m, m) &= \frac{(k-m)^{n+1}}{n!} \int_0^1 \gamma^{(n+1)}\big(m + (k-m)x\big)(1-x)\,dx\\ &\le \frac{(k-m)^{n+1}}{n!} \max_{x\in[0,1]} \gamma^{(n+1)}\big(m + (k-m)x\big)\\ &= \mathcal{O}(|k-m|^{n+1})\big(1 + (1+|k-m|)^{\deg^*(\gamma)-n-1} + (1+|m|)^{\deg^*(\gamma)-n-1}\big).\end{aligned}$$

For the last step, note that $\deg^*(\tilde{\gamma}) \in \mathbb{R}$ yields

$$\sup_{k\in\mathbb{R}} \frac{\tilde{\gamma}(k)}{1 + (1+|k|)^{\deg^*(\tilde{\gamma})}} \le C,$$

for some $C > 0$, which implies

$$\max_{x\in[0,1]} \frac{\tilde{\gamma}(m+(k-m)x)}{1+(1+|m+(k-m)x|)^{\deg^*(\tilde{\gamma})}} \leq C.$$

For $\deg^*(\tilde{\gamma}) \leq 0$, $\tilde{\gamma}$ is obviously bounded.
For $\deg^*(\tilde{\gamma}) > 0$, we can use the triangle inequality to get

$$\frac{\max_{x\in[0,1]} \tilde{\gamma}(m+(k-m)x)}{1+(1+|m|+|k-m|)^{\deg^*(\tilde{\gamma})}} \leq C,$$

what implies

$$\max_{x\in[0,1]} \tilde{\gamma}(m+(k-m)x) \leq C\left(1+(1+2|m|)^{\deg^*(\tilde{\gamma})}+(1+2|k-m|)^{\deg^*(\tilde{\gamma})}\right),$$

since we have $|m| \leq |k-m|$ or $|m| \geq |k-m|$.

We now get

$$\begin{aligned}
\|\mathcal{R}(f,g)\|_{L^2} &= \left\| \int_{\mathbb{R}} r(k,k-m,m)\,\widehat{g}(k-m)\,\widehat{f}(m)\,dm \right\|_{L^2} \\
&\leq \left\| \int_{\mathbb{R}} |r(k,k-m,m)\,\widehat{g}(k-m)\,\widehat{f}(m)|\,dm \right\|_{L^2} \\
&\leq \mathcal{O}(1) \left\| \int_{\mathbb{R}} |(1+|k-m|^2)^{p/2}\,\widehat{\partial_x^{n+1} g}(k-m)\,(1+|m|^2)^{p/2}\,\widehat{f}(m)|\,dm \right\|_{L^2},
\end{aligned}$$

where

$$p := \max\{\deg^*(\gamma)-n-1,\,0\}.$$

When we use Young's inequality, we get

$$\|\mathcal{R}(f,g)\|_{L^2} \leq \mathcal{O}(1)\,\|\widehat{\partial_x^{n+1} g}\|_{L^1(p)} \|f\|_{H^p}.$$

When we exploit Plancherel together with Sobolev's embedding theorem, we obtain

$$\begin{aligned}
\|\mathcal{R}(f,g)\|_{L^2} &\leq \mathcal{O}(1) \left\| \mathcal{F}^{-1}\big[|(1+|\cdot|^2)^{p/2}\,\widehat{\partial_x^{n+1} g}(\cdot)|\big]\ \mathcal{F}^{-1}\big[|(1+|\cdot|^2)^{p/2}\,\widehat{f}(\cdot)|\big] \right\|_{L^2} \\
&\leq \mathcal{O}(1) \left\| \mathcal{F}^{-1}\big[|(1+|\cdot|^2)^{p/2}\,\widehat{\partial_x^{n+1} g}(\cdot)|\big] \right\|_{\infty} \left\| \mathcal{F}^{-1}\big[|(1+|\cdot|^2)^{p/2}\,\widehat{f}(\cdot)|\big] \right\|_{L^2} \\
&\leq \mathcal{O}(1)\,\|\partial_x^{n+1} g\|_{H^{p+q}} \|f\|_{H^p}.
\end{aligned}$$

□

The next lemma will help us to address the points [R2] and [R4]. In the context of the introduction to this subsection, the mappings of this lemma map the f from the original problematic integral onto the f of an emerging integral from εJ_E.

Lemma 2.2.22. *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}_0$ with $m \le n$.*
Let D^n_m be the set of functions defined by

$$D^n_m := \{f \in H^n(\mathbb{R}) : \|\partial_x^{-1} f\|_\infty + \|\partial_t f\|_{H^m} + \|\partial_t \partial_x^{-1} f\|_{C^m} < \infty\}.$$

For $\varphi \in H^{n+1}(\mathbb{R})$ with $\|\partial_t\varphi\|_{C^m}, \|\partial_t\varphi\|_{H^{m+1}} \in \mathbb{R}$, the mappings

$$M^1_\varphi : f \mapsto \varphi f \tag{2.93}$$

and

$$M^2_\varphi : f \mapsto \partial_x(\varphi\, \partial_x^{-1} f) \tag{2.94}$$

map D^n_m into D^n_m.

Moreover, if for some $c_\varphi \in \mathbb{R}$, we have

$$\|\varphi\|_{H^{n+1}}, \|\partial_t\varphi\|_{H^{m+1}} \le \varepsilon^{-1/2} c_\varphi \qquad \text{and} \qquad \|\varphi\|_{C^n}, \|\partial_t\varphi\|_{C^m} \le c_\varphi, \tag{2.95}$$

then there is a constant $C = C(c_\varphi) \ge 0$ such that the following a priori estimates are true:

$$\begin{aligned}
&\|M^1_\varphi f\|_{H^n} + \|\partial_t M^1_\varphi f\|_{H^m} \\
&\qquad\qquad \le C\big(\|f\|_{H^n} + \|\partial_t f\|_{H^m}\big) \\
&\|\partial_x^{-1} M^1_\varphi f\|_\infty + \|\partial_t \partial_x^{-1} M^1_\varphi f\|_{C^m} \\
&\qquad\qquad \le \varepsilon^{-1/2}\, C\big(\|f\|_{L^2} + \|\partial_t f\|_{L^2}\big) + C\big(\|f\|_{C^{m-1}} + \|\partial_t f\|_{C^{m-1}}\big),
\end{aligned}$$

and

$$\begin{aligned}
&\|M^2_\varphi f\|_{H^n} + \|\partial_t M^2_\varphi f\|_{H^m} \\
&\qquad\qquad \le C\big(\|f\|_{H^n} + \|\partial_t f\|_{H^m}\big) + \varepsilon^{-1/2}\, C\big(\|\partial_x^{-1} f\|_{C^m} + \|\partial_t \partial_x^{-1} f\|_{C^m}\big), \\
&\|\partial_x^{-1} M^2_\varphi f\|_\infty + \|\partial_t \partial_x^{-1} M^2_\varphi f\|_{C^m} \\
&\qquad\qquad \le C\big(\|\partial_x^{-1} f\|_{C^m} + \|\partial_t \partial_x^{-1} f\|_{C^m}\big).
\end{aligned}$$

Remark 2.2.23. *In this context, we choose to define the operator ∂_x^{-1} by*

$$\partial_x^{-1} f := \int_{-\infty}^{(\cdot)} f\, dx + c(t)\,, \quad c(t) := 0\,.$$

Proof. For map (2.93), we have

$$\|M^1_\varphi f\|_{H^n} = \|\varphi f\|_{H^n} \le \mathcal{O}\big(\|\varphi\|_{C^n}\big)\|f\|_{H^n}$$

and

$$\begin{aligned}
\|\partial_x^{-1} M^1_\varphi f\|_\infty = \|\partial_x^{-1}(\varphi f)\|_\infty &= \Big\| \int_{-\infty}^{(\cdot)} \varphi f\, dx \Big\|_\infty \\
&\le \Big\| \int_{-\infty}^{(\cdot)} |\varphi f|\, dx \Big\|_\infty \le \int_{\mathbb{R}} |\varphi f|\, dx \\
&\le \|\varphi\|_{L^2}\|f\|_{L^2}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\|\partial_t M^1_\varphi f\|_{H^m} &= \|\partial_t(\varphi f)\|_{H^m} \le \|\partial_t \varphi f\|_{H^m} + \|\varphi \partial_t f\|_{H^m} \\ &\le \|\partial_t \varphi\|_{C^m}\|f\|_{H^m} + \|\varphi\|_{C^m}\|\partial_t f\|_{H^m},\end{aligned}$$

and

$$\begin{aligned}\|\partial_x^{-1}\partial_t M^1_\varphi f\|_\infty = \|\partial_x^{-1}\partial_t(\varphi f)\|_\infty = \Big\| \int_{-\infty}^{(\cdot)} \partial_t(\varphi f)\, dx\Big\|_\infty \le \int_{\mathbb{R}} |\partial_t(\varphi f)|\, dx \\ \le \|\partial_t\varphi\|_{L^2}\|f\|_{L^2} + \|\varphi\|_{L^2}\|\partial_t f\|_{L^2},\end{aligned}$$

while for $l = 1, \ldots, m$:

$$\begin{aligned}\|\partial_x^{l-1}\partial_t M^1_\varphi f\|_\infty &= \|\partial_x^{l-1}\partial_t(\varphi f)\|_\infty = \|\partial_x^{l-1}(\partial_t\varphi f) + \partial_x^{l-1}(\varphi\partial_t f)\|_\infty \\ &\le \mathcal{O}(\|\partial_t\varphi\|_{C^{m-1}})\|f\|_{C^{m-1}} + \mathcal{O}(\|\varphi\|_{C^{m-1}})\|\partial_t f\|_{C^{m-1}}.\end{aligned}$$

For map (2.94), we have

$$\begin{aligned}\|M^2_\varphi f\|_{H^n} &= \|\partial_x(\varphi\, \partial_x^{-1} f)\|_{H^n} \le \mathcal{O}(1) \sum_{l=1}^{n+1}\sum_{k=0}^{l} \binom{l}{k} \|\partial_x^{l-k}\varphi\, \partial_x^{k-1} f\|_{L^2} \\ &\le \mathcal{O}(1) \ \sum_{l=1}^{n+1} \Big(\sum_{k=1}^{l} \|\partial_x^{l-k}\varphi\, \partial_x^{k-1} f\|_{L^2} + \|\partial_x^{l}\varphi\, \partial_x^{-1} f\|_{L^2}\Big) \\ &\le \mathcal{O}\big(\|\varphi\|_{C^n}\big)\|f\|_{H^n} + \mathcal{O}\big(\|\varphi\|_{H^{n+1}}\big)\|\partial_x^{-1} f\|_\infty\,,\end{aligned}$$

and

$$\|\partial_x^{-1} M^2_\varphi f\|_\infty = \|\varphi\, \partial_x^{-1} f\|_\infty \le \|\varphi\|_\infty \|\partial_x^{-1} f\|_\infty.$$

Moreover, we have

$$\begin{aligned}\|\partial_t M^2_\varphi f\|_{H^m} &= \|\partial_t\partial_x(\varphi\partial_x^{-1} f)\|_{H^m} = \|\partial_t\partial_x\varphi\partial_x^{-1} f + \partial_x\varphi\partial_t\partial_x^{-1} f + \partial_t\varphi f + \varphi\partial_t f\|_{H^m} \\ &\le \|\partial_t\partial_x\varphi\|_{H^m}\|\partial_x^{-1} f\|_{C^m} + \|\partial_x\varphi\|_{H^m}\|\partial_t\partial_x^{-1} f\|_{C^m} \\ &\quad + \|\partial_t\varphi\|_{C^m}\|f\|_{H^m} + \|\varphi\|_{C^m}\|\partial_t f\|_{H^m},\end{aligned}$$

and

$$\begin{aligned}\|\partial_t\partial_x^{-1} M^2_\varphi f\|_{C^m} &= \|\partial_t(\varphi\partial_x^{-1} f)\|_{C^m} = \|\partial_t\varphi\partial_x^{-1} f + \varphi\partial_t\partial_x^{-1} f\|_{C^m} \\ &\le \|\partial_t\varphi\|_{C^m}\|\partial_x^{-1} f\|_{C^m} + \|\varphi\|_{C^m}\|\partial_t\partial_x^{-1} f\|_{C^m}.\end{aligned}$$

□

Lemma 2.2.24. *Let $N \in \mathbb{N}$ and $\ell \geq 2N+1$.*
By introducing the notation

$$\tilde{R}_\psi := \psi + \varepsilon^{\beta-1}\vartheta(R_1 + R_{-1}) \tag{2.96}$$

we obtain

$$\begin{aligned}\partial_x^\ell\big(R_\psi \vartheta(R_1+R_{-1})\big) = &\sum_{n=0}^{N}\binom{\ell}{n}\partial_x^n \tilde{R}_\psi\, \partial_x^{\ell-n}\vartheta(R_1+R_{-1}) \\ &+ \sum_{n=N+1}^{\ell-N-1}\binom{\ell}{n}\partial_x^n R_\psi\, \partial_x^{\ell-n}\vartheta(R_1+R_{-1}) \\ &+ \sum_{n=\ell-N}^{\ell}\binom{\ell}{n}\partial_x^n \psi\, \partial_x^{\ell-n}\vartheta(R_1+R_{-1})\,.\end{aligned} \tag{2.97}$$

Proof. Leibniz's rule and the definition of $R_\psi = \psi + \frac{1}{2}\varepsilon^{\beta-1}\vartheta(R_1+R_{-1})$ yield

$$\begin{aligned}&\partial_x^\ell\big(R_\psi\vartheta(R_1+R_{-1})\big) \\ &\quad= \sum_{n=0}^{N}\binom{\ell}{n}\partial_x^n\big(\psi + \frac{1}{2}\varepsilon^{\beta-1}(\vartheta R_{-1} + \vartheta R_1)\big)\,\partial_x^{\ell-n}\vartheta(R_1+R_{-1}) \\ &\quad+ \sum_{n=N+1}^{\ell-N-1}\binom{\ell}{n}\partial_x^n R_\psi\,\partial_x^{\ell-n}\vartheta(R_1+R_{-1})) \\ &\quad+ \sum_{n=\ell-N}^{\ell}\binom{\ell}{n}\partial_x^n\big(\psi + \frac{1}{2}\varepsilon^{\beta-1}(\vartheta R_{-1} + \vartheta R_1)\big)\,\partial_x^{\ell-n}\vartheta(R_1+R_{-1})\,.\end{aligned}$$

We now obtain (2.97) since

$$\sum_{n=\ell-N}^{\ell}\binom{\ell}{n}\partial_x^n f\,\partial_x^{\ell-n} f = \sum_{\tilde{n}=0}^{N}\binom{\ell}{\tilde{n}}\partial_x^{\ell-\tilde{n}} f\,\partial_x^{\tilde{n}} f\,.$$

□

Remark 2.2.25. *In the following we will assume that ε_0 is chosen such small that, for $0 < \varepsilon < \varepsilon_0$, we have*

$$\varepsilon\,\mathcal{E}_\ell \leq 1. \tag{2.98}$$

Under this assumption we can for instance make the estimates

$$\varepsilon\,\mathcal{E}_\ell^{3/2} = \mathcal{O}(\mathcal{E}_\ell)\,, \qquad \|\tilde{R}_\psi\|_{H^\ell} = \mathcal{O}(\varepsilon^{-1/2})\,, \qquad \text{or} \qquad \|\tilde{R}_\psi\|_{C^{\ell-1}} = \mathcal{O}(1).$$

We can make this assumption due to the fact that there is a some $T(\varepsilon) > 0$ such that the H^ℓ-norms of $R_{-1}(t)$ and $R_1(t)$ can be uniformly bounded for $0 \leq t \leq T(\varepsilon)$. Later, when our energy estimates close and we use Gronwall's inequality, we will obtain $T(\varepsilon) \geq T_0\,\varepsilon^{-2}$.

We now finally implement the four key observations from the introduction in such a way that [R1], [R2], [R3] and [R4] are guaranteed.

Notation. We denote the floor function by

$$\lfloor x \rfloor = \max\{z \in \mathbb{Z} : z \leq x\}$$

and the ceiling function by

$$\lceil x \rceil = \min\{z \in \mathbb{Z} : z \geq x\}.$$

Lemma 2.2.26. *Let $\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$.*
Let γ be an pseudo-differential operator that does not depend on ε and is given by its symbol in Fourier space. Furthermore, let there be some $D \geq 0$ such that we have either

$$\gamma(k) = i\sigma(k) \quad \text{for all } |k| \geq D \qquad \text{or} \qquad \gamma(k) = \upsilon(k) \quad \text{for all } |k| \geq D\,.$$

We assume the function $\sigma \in C^{\lceil \deg^*(\sigma) \rceil}(\mathbb{R},\mathbb{R})$ to be odd with $\deg^*(\sigma) \leq \deg(\omega)$, the function $\upsilon \in C^{\lceil \deg^*(\upsilon) \rceil}(\mathbb{R},\mathbb{R})$ to be even with $\deg^*(\upsilon) \leq \deg(\omega')$, and both functions to share the property (2.89).
When $\gamma(k) = i\sigma(k)$ for large $|k|$, let $f = h$ be a function with

$$\begin{aligned} \|h\|_{H^{\lceil \deg^*(\sigma) \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_t h\|_{H^{\lceil \deg^*(\sigma) \rceil - 1}} &= \mathcal{O}(\varepsilon^{-1/2}), \\ \|h\|_{C^{\lceil \deg^*(\sigma) \rceil - 1}} + \|\partial_t h\|_{C^{\lceil \deg^*(\sigma) \rceil - 1}} &= \mathcal{O}(1)\,. \end{aligned} \tag{2.99}$$

When $\gamma(k) = \upsilon(k)$ for large $|k|$, let $f = g$ be a function with

$$\begin{aligned} \|g\|_{H^{\lceil \deg^*(\upsilon) \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_t g\|_{H^{\lceil \deg^*(\upsilon) \rceil}} &= \mathcal{O}(\varepsilon^{-1/2}), \\ \|\partial_x^{-1} g\|_{C^{\lceil \deg^*(\upsilon) \rceil}} + \|\partial_t \partial_x^{-1} g\|_{C^{\lceil \deg^*(\upsilon) \rceil}} &= \mathcal{O}(1). \end{aligned} \tag{2.100}$$

Suppose

$$\|f\|_{H^{\lceil \deg^*(\gamma) \rceil + \lceil \deg(\omega) \rceil}} = \mathcal{O}(1) \qquad \text{or} \qquad \|\widehat{f}\|_{L^1(\lceil \deg^*(\gamma) \rceil + \lceil \deg(\omega) \rceil)} = \mathcal{O}(1), \tag{2.101}$$

then there is an expression $\mathcal{D}$ with

$$\mathcal{D} = \mathcal{O}(\mathcal{E}_\ell),$$

such that for $j_1, j_2 \in \{\pm 1\}$:

$$\begin{aligned} &\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &= \varepsilon^2 \partial_t \mathcal{D} + \varepsilon^2 \sum_{k=1}^{\lceil \deg^*(\gamma) \rceil - 1} \int_{\mathbb{R}} \varsigma_k \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, \partial_x^k f\, dx \\ &\quad + \varepsilon^2 \sum_{k=1}^{m} \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_k\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1), \end{aligned} \tag{2.102}$$

where ς_k and γ_k are skew symmetric or symmetric operators independent of ε and given by their symbol in Fourier space, $m = m\big(\deg^*(\gamma)\big) \in \mathbb{N}$, the f_k are some functions and $p_k, q_k \in \{-1, 1\}$. The functions $\varsigma_k \in C^{\lceil \deg^*(\varsigma_k)\rceil}(\mathbb{R},\mathbb{R})$ and $\gamma_k \in C^{\lceil \deg^*(\gamma_k)\rceil}(\mathbb{R},\mathbb{R})$ share the property (2.89) .
Furthermore, we have

$$\deg^*(\varsigma_k) \leq \deg^*(\gamma) - k\,, \tag{2.103}$$

$$\deg^*(\gamma_k) \leq \deg^*(\gamma) - \big(\deg(\omega) - \deg^*(\rho)\big)\,, \tag{2.104}$$

and

$$\begin{aligned}
&\|f_k\|_{H^{\lceil \deg^*(\gamma_k)\rceil + \lceil \deg(\omega)\rceil}} + \|\partial_t f_k\|_{H^{\lceil \deg^*(\gamma_k)\rceil}} \\
&\qquad \leq \varepsilon\, C_1 \left(\|f\|_{H^{\lceil \deg^*(\gamma_k)\rceil + \lceil \deg(\omega)\rceil}} + \|\partial_t f\|_{H^{\lceil \deg^*(\gamma_k)\rceil}}\right) \\
&\qquad\quad + \varepsilon^{1/2}\, C_2 \left(\|\partial_x^{-1} f\|_{C^{\lceil \deg^*(\gamma_k)\rceil}} + \|\partial_t \partial_x^{-1} f\|_{C^{\lceil \deg^*(\gamma_k)\rceil}}\right),
\end{aligned} \tag{2.105}$$

$$\begin{aligned}
&\|\partial_x^{-1} f_k\|_\infty + \|\partial_t \partial_x^{-1} f_k\|_{C^{\lceil \deg^*(\gamma_k)\rceil}} \\
&\quad \leq \varepsilon^{1/2}\, C_1 \left(\|f\|_{L^2} + \|\partial_t f\|_{L^2}\right) + \varepsilon\, C_2 \Big(\|\partial_x^{-1} f\|_{C^{\lceil \deg^*(\gamma_k)\rceil}} + \|\partial_t \partial_x^{-1} f\|_{C^{\lceil \deg^*(\gamma_k)\rceil}}\Big)\,,
\end{aligned} \tag{2.106}$$

where the constants C_1, C_2 depend on $\tilde{R}_\psi, f, \gamma$ but are independent of ε. We set $C_2 := 0$ when $f = h$.

Remark 2.2.27. *When γ is skew symmetric, we can even obtain*

$$\deg^*(\gamma_k) \leq \deg^*(\gamma) - 1 - \big(\deg(\omega) - \deg^*(\rho)\big)\,, \tag{2.107}$$

see a2) *and* c) *in the proof. This makes the case* $\deg^*(\rho) \leq 1$ *easy to handle since the complete energy transformation is done after one step.*

Proof. If $\deg^*(\gamma) \leq 0$, we have

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx = \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),$$

such that the lemma is trivially true.
So we will in the following assume $\deg^*(\gamma) > 0$.

Since $\deg(\omega) \geq \deg^*(\gamma) > 0$, there exist some constants $D_\omega, d_\omega > 0$ such that

$$|\omega(k)| \geq d_\omega > 0 \qquad \text{for } |k| \geq D_\omega.$$

When $\gamma = \upsilon$, we can on top of that find $D_\omega, d_\omega > 0$ such that

$$|\omega'(k)| \geq d_\omega > 0 \qquad \text{for } |k| \geq D_\omega\,,$$

due to $\deg(\omega') \geq \deg^*(\upsilon) > 0$.
There is some $D \geq D_\omega$ and some function $\tilde{\gamma} \in C^{\lceil \deg^*(\gamma) \rceil}(\mathbb{R}, \mathbb{R})$ with (2.89) such that

$$\tilde{\gamma}(k) = \gamma(k) \quad \text{for } |k| \geq D\,, \qquad \text{and} \qquad \tilde{\gamma}(k) = 0 \quad \text{for } |k| \leq D_\omega\,.$$

Since we have

$$\begin{aligned}
\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx &= \varepsilon^2 \int_{\mathbb{R}} \tilde{\gamma} \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx + \varepsilon^2 \int_{\mathbb{R}} (\gamma - \tilde{\gamma}) \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\
&= \varepsilon^2 \int_{\mathbb{R}} \tilde{\gamma} \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),
\end{aligned}$$

we can in the following assume that we have $\gamma \in C^{\lceil \deg^*(\gamma) \rceil}(\mathbb{R}, \mathbb{R})$ with (2.89) and $\gamma(k) = 0$ for $|k| \leq D_\omega$. Therefore, the operators given by the expressions $\frac{\gamma}{\omega}$ and $\frac{\gamma}{\omega'}$ will make sense, and, we also will be able to make use of (2.90).
In this proof we will sometimes implicitly assume $\omega, \rho \in C^{m_\omega}(\mathbb{R}, \mathbb{R})$ when we apply (2.90), we can do this for the same reasons as above.
a) Handling integrals of the form

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx\,. \tag{2.108}$$

By exploiting the skew symmetry of $i\omega$ and (2.19), we have

$$\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx \\
&= \frac{1}{2} j\, \varepsilon^2\, \partial_t \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, f\, dx \\
&\quad - \frac{1}{2} \varepsilon^2 \int_{\mathbb{R}} \big[i\omega, f\big] \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, dx \\
&\quad - \frac{1}{2} \varepsilon^3 \int_{\mathbb{R}} \frac{\gamma}{i\omega} i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1})\big)\, \partial_x^\ell R_{-j}\, f\, dx \\
&\quad + \frac{1}{2} \varepsilon^3 \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1})\big)\, f\, dx \\
&\quad - \frac{1}{2} j\, \varepsilon^2 \int_{\mathbb{R}} \frac{\gamma}{i\omega} \partial_x^\ell R_j\, \partial_x^\ell R_{-j}\, \partial_t f\, dx \\
&\quad - \frac{1}{2} j\, \varepsilon^{2-\beta} \int_{\mathbb{R}} \gamma \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_j}(\varepsilon\Psi)\, \partial_x^\ell R_{-j}\, f\, dx \\
&\quad - \frac{1}{2} j\, \varepsilon^{2-\beta} \int_{\mathbb{R}} \gamma \partial_x^\ell R_j\, \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{-j}}(\varepsilon\Psi)\, f\, dx.
\end{aligned}$$

The first term is the time derivative of an integral, which can be estimated against $\varepsilon^2 \mathcal{O}(\mathcal{E}_\ell)$ by using Cauchy-Schwarz.
The last three integrals can be estimated against $\varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell+1)$ since $\|\partial_t f\|_\infty = \mathcal{O}(1)$ and due to (2.9).
For the second integral, applying (2.90) gives us

$$
\begin{aligned}
&-\frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \big[i\omega, f\big]\frac{\gamma}{i\omega}\partial_x^\ell R_j\,\partial_x^\ell R_{-j}\,dx \\
&\qquad = -\frac{1}{2}\varepsilon^2 \sum_{n=1}^{\lceil \deg^*(\gamma)\rceil -1} \frac{1}{n!}\int_{\mathbb{R}} (-i)^n \omega^{(n)}\frac{\gamma}{\omega}\partial_x^\ell R_j\,\partial_x^\ell R_{-j}\,\partial_x^n f\,dx \\
&\qquad\quad + \mathcal{O}(\varepsilon^2)\,\|\mathcal{R}(\frac{\gamma}{\omega}\partial_x^\ell R_j, f)\|_{L^2}\|\partial_x^\ell R_{-j}\|_{L^2},
\end{aligned}
$$

where we estimate $\|\mathcal{R}(\frac{\gamma}{\omega}\partial_x^\ell R_j, f)\|_{L^2} = \mathcal{O}(\|f\|_{H^{\lceil \deg(\omega)\rceil+1}}\|R_j\|_{H^\ell}) = \mathcal{O}(1)$ as in (2.91) or $\|\mathcal{R}(\frac{\gamma}{\omega}\partial_x^\ell R_j, f)\|_{L^2} = \mathcal{O}(\|\widehat{f}\|_{L^1(\lceil \deg(\omega)\rceil)}\|R_j\|_{H^\ell}) = \mathcal{O}(1)$ as in (2.92).
We now have obtained

$$
\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \gamma\partial_x^\ell R_j\,\partial_x^\ell R_{-j}\,f\,dx \\
&\quad = \varepsilon^2\,\partial_t\tilde{\mathcal{D}} \\
&\quad\quad -\frac{1}{2}\varepsilon^2 \sum_{n=1}^{\lceil \deg^*(\gamma)\rceil -1} \frac{1}{n!}\int_{\mathbb{R}} (-i)^n \omega^{(n)}\frac{\gamma}{\omega}\partial_x^\ell R_j\,\partial_x^\ell R_{-j}\,\partial_x^n f\,dx \\
&\quad\quad -\frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\gamma}{i\omega} i\rho\partial_x^\ell\vartheta^{-1}\big(R_\psi\vartheta(R_1+R_{-1})\big)\,\partial_x^\ell R_{-j}\,f\,dx \\
&\quad\quad +\frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\gamma}{i\omega}\partial_x^\ell R_j\,i\rho\partial_x^\ell\vartheta^{-1}\big(R_\psi\vartheta(R_1+R_{-1})\big)\,f\,dx \\
&\quad\quad + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell+1),
\end{aligned}
$$

for some $\tilde{\mathcal{D}} = \mathcal{O}(\mathcal{E}_\ell)$.
The second term already has the desired form.
The integrals in the third and the forth place can be written as a sum of some $\varepsilon^3\,\mathcal{O}(\mathcal{E}_\ell+1)$-terms and m many integrals of the form

$$
\varepsilon^2 \int_{\mathbb{R}} \gamma_k\partial_x^\ell R_{p_k}\,\partial_x^\ell R_{q_k}\,f_k\,dx
$$

with m, γ_k, f_k, p_k and q_k just as in the lemma.
We see this by using (2.73), Leibniz's rule, (2.90) and (2.72). Additionally, we also have to rely on the results of lemma 2.2.22 for map (2.93) in order to see that the functions of the form $f_k = \varepsilon\,\partial_x^n\tilde{R}_\psi\partial_x^m f$ in the resulting integrals do fulfill (2.105)

and (2.106). We show this lengthy calculation once in detail for the third term, the forth term can be handled analogously.
Using (2.73), then Leibniz's rule and (2.97) with $N := \lceil \deg^*(\gamma) \rceil - 1$ in order to extract all terms with more than ℓ spatial derivatives falling on R_1 or R_{-1}, we get

$$
\begin{aligned}
-\frac{1}{2}\varepsilon^3 & \int_{\mathbb{R}} \frac{\gamma}{i\omega} i\rho \partial_x^\ell \vartheta^{-1}\big(R_\psi \vartheta(R_1 + R_{-1})\big)\, \partial_x^\ell R_{-j}\, f\, dx \\
&= -\frac{1}{2}\varepsilon^3 \sum_{n=0}^{N} \binom{\ell}{n} \int_{\mathbb{R}} \frac{\rho}{\omega}\gamma\big(\partial_x^n \tilde{R}_\psi\, \partial_x^{\ell-n}\vartheta(R_1 + R_{-1})\big)\, \partial_x^\ell R_{-j}\, f\, dx \\
&\quad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1).
\end{aligned}
$$

With the help of (2.90), (2.91), (2.92) and (2.72) we now arrive at

$$
\begin{aligned}
-\frac{1}{2}\varepsilon^3 & \int_{\mathbb{R}} \frac{\gamma}{i\omega} i\rho \partial_x^\ell \vartheta^{-1}\big(R_\psi \vartheta(R_1 + R_{-1})\big)\, \partial_x^\ell R_{-j}\, f\, dx \\
&= -\frac{1}{2}\varepsilon^3 \sum_{n=0}^{N} \binom{\ell}{n} \Big(\int_{\mathbb{R}} \frac{\rho}{\omega}\gamma \partial_x^{\ell-n}\vartheta(R_1 + R_{-1})\, \partial_x^\ell R_{-j}\, \partial_x^n \tilde{R}_\psi\, f\, dx \\
&\qquad\qquad + \int_{\mathbb{R}} \big[\frac{\rho}{\omega}\gamma, \partial_x^n \tilde{R}_\psi\big] \partial_x^{\ell-n}\vartheta(R_1 + R_{-1})\, \partial_x^\ell R_{-j}\, f\, dx \Big) \\
&\quad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1) \\
&= -\frac{1}{2}\varepsilon^3 \sum_{n=0}^{N} \binom{\ell}{n} \Big(\int_{\mathbb{R}} \frac{\rho}{\omega}\gamma \partial_x^{\ell-n}(R_1 + R_{-1})\, \partial_x^\ell R_{-j}\, \partial_x^n \tilde{R}_\psi\, f\, dx \\
&\qquad\qquad + \sum_{m=1}^{M_n} \frac{1}{m!} \int_{\mathbb{R}} (-i)^m \big(\frac{\rho}{\omega}\gamma\big)^{(m)} \partial_x^{\ell-n}(R_1 + R_{-1})\, \partial_x^\ell R_{-j}\, \partial_x^{m+n} \tilde{R}_\psi\, f\, dx \Big) \\
&\quad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1) \\
&\quad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell)\, \|\widehat{\psi}\|_{L^1(2\lceil \deg^*(\gamma)\rceil - 1)} + \varepsilon^{3+\beta-1}\, \mathcal{O}(\mathcal{E}_\ell)\, \|R_1 + R_{-1}\|_{H^{2\lceil \deg^*(\gamma)\rceil}},
\end{aligned}
$$

where $M_n := \lceil \deg^*(\gamma) \rceil - n - 1$ and the operators $(-i)^m \big(\frac{\rho}{\omega}\gamma\big)^{(m)}$ are given by their symbols just as in (2.90). We have

$$
\|\widehat{\psi}\|_{L^1(2\lceil \deg^*(\gamma)\rceil - 1)} = \mathcal{O}(1),
$$

due to (2.11), and

$$
\varepsilon^{\beta-1} \|R_1 + R_{-1}\|_{H^{2\lceil \deg^*(\gamma)\rceil}} = \varepsilon^{\beta-1}\, \mathcal{O}(\sqrt{\mathcal{E}_\ell}) = \mathcal{O}(1)
$$

due to $\ell \geq 2\lceil \deg^*(\gamma) \rceil$.
The above sum of integrals is now of the form

$$\varepsilon^2 \sum_{k=1}^{\tilde{m}} \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k} \, \partial_x^\ell R_{q_k} \, f_k \, dx$$

with γ_k, f_k, p_k and q_k just as in the lemma.
This is in particular true, since, due to

$$\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1 \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\gamma) \rceil + 1 \,, \tag{2.109}$$

we have for $0 \leq p \leq \lceil \deg^*(\gamma) \rceil - 1$:

$$\|\partial_x^p \tilde{R}_\psi\|_{H^{\lceil \deg^*(\gamma)-p \rceil + \lceil \deg(\omega) \rceil + 1}} + \|\partial_t \partial_x^p \tilde{R}_\psi\|_{H^{\lceil \deg^*(\gamma)-p \rceil + 1}} \leq \varepsilon^{-1/2} \, c$$

and

$$\|\partial_x^p \tilde{R}_\psi\|_{C^{\lceil \deg^*(\gamma)-p \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_t \partial_x^p \tilde{R}_\psi\|_{C^{\lceil \deg^*(\gamma)-p \rceil}} \leq c \;,$$

for some constant $c = c(\tilde{R}_\psi)$ that is independent of ε (due to assumption (2.107)).
Therefore the estimates of Lemma 2.2.22 for the map (2.93) give us

$$\begin{aligned} &\|\varepsilon \, \partial_x^p \tilde{R}_\psi f\|_{H^{\lceil \deg^*(\gamma)-p \rceil + \lceil \deg(\omega) \rceil}} + \|\varepsilon \, \partial_t (\partial_x^p \tilde{R}_\psi f)\|_{H^{\lceil \deg^*(\gamma)-p \rceil}} \\ &\qquad \leq \varepsilon \, C \left(\|f\|_{H^{\lceil \deg^*(\gamma)-p \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_t f\|_{H^{\lceil \deg^*(\gamma)-p \rceil}} \right) \end{aligned}$$

and

$$\begin{aligned} &\|\varepsilon \, \partial_x^{-1} (\partial_x^p \tilde{R}_\psi f)\|_\infty + \|\varepsilon \, \partial_t \partial_x^{-1} (\partial_x^p \tilde{R}_\psi f)\|_{C^{\lceil \deg^*(\gamma)-p \rceil}} \\ &\qquad \leq \varepsilon^{1/2} \, C \left(\|f\|_{L^2} + \|\partial_t f\|_{L^2} \right) + \varepsilon \, C \left(\|f\|_{C^{\lceil \deg^*(\gamma)-p \rceil - 1}} + \|\partial_t f\|_{C^{\lceil \deg^*(\gamma)-p \rceil - 1}} + 1 \right), \end{aligned}$$

for some constant $C = C(\tilde{R}_\psi)$ that is independent of ε, what verifies (2.105) and (2.106).

a2) Handling integrals of the form

$$\varepsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_{-j} \, h \, dx \,. \tag{2.110}$$

We already proved lemma 2.2.26 for these kind of integrals, but when $i\sigma$ is skew symmetric we can still show some more. In a), we showed that

$$
\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_{-j} \, h \, dx \\
&\quad = \varepsilon^2 \, \partial_t \tilde{\mathcal{D}} \\
&\qquad - \frac{1}{2}\varepsilon^2 \sum_{n=1}^{\lceil \deg^*(\sigma) \rceil - 1} \frac{1}{n!} \int_{\mathbb{R}} (-i)^n \omega^{(n)} \frac{i\sigma}{\omega} \partial_x^\ell R_j \, \partial_x^\ell R_{-j} \, \partial_x^n h \, dx \\
&\qquad - \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, \partial_x^\ell R_{-j} \, h \, dx \\
&\qquad + \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} \partial_x^\ell R_j \, i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, h \, dx \\
&\qquad + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \, ,
\end{aligned}
$$

for some $\tilde{\mathcal{D}} = \mathcal{O}(\mathcal{E}_\ell)$. Since $i\sigma$ is skew symmetric, we can handle the last two integrals better than in a). By using (2.73) and the skew symmetry of $i\rho$ and $i\rho\frac{\sigma}{\omega}$, we get

$$
\begin{aligned}
&- \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, \partial_x^\ell R_{-j} \, h \, dx \\
&\quad + \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} \partial_x^\ell R_j \, i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, h \, dx \\
&\qquad = -j\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell (R_1 - R_{-1}) \, \partial_x^\ell \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, h \, dx \\
&\qquad\quad + \varepsilon^3 \int_{\mathbb{R}} \big[\frac{\sigma}{\omega} i\rho, h \big] \partial_x^\ell R_{-j} \, \partial_x^\ell \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, dx. \\
&\qquad\quad - \varepsilon^3 \int_{\mathbb{R}} \big[i\rho, h \big] \frac{\sigma}{\omega} \partial_x^\ell R_j \, \partial_x^\ell \big(R_\psi \vartheta (R_1 + R_{-1}) \big) \, dx \\
&\qquad\quad + \varepsilon^3 \, \mathcal{O}(\mathcal{E}_\ell + 1) \, ,
\end{aligned}
$$

The last two integrals can be written as some $\varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1)$-terms plus a sum of integrals of the form

$$
\varepsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k} \, \partial_x^\ell R_{q_k} \, f_k \, dx
$$

with γ_k, f_k, p_k, q_k just as in the lemma and

$$
\deg^*(\gamma_k) \leq \deg^*(\sigma) - 1 \, .
$$

We see this by using (2.90), Leibniz's rule, (2.97), (2.72). By using the result of lemma 2.2.22 for the map (2.93) we see that the emerging functions of the form $f_k = \varepsilon \, \partial_x^n h \partial_x^p \tilde{R}_\psi$, with $p \geq 0$, fulfill (2.105) and (2.106).

Exploiting Leibniz's rule, (2.97), (2.72) for the first integral on the above right hand side, we get

$$
\begin{aligned}
& -j\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell \big(R_\psi \vartheta (R_1 + R_{-1})\big)\, h\, dx \\
&\quad = -j\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, \tilde{R}_\psi h\, dx \\
&\qquad - j\varepsilon^3 \sum_{n=1}^{N} \binom{\ell}{n} \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell (R_1 - R_{-1})\, \partial_x^{\ell-n} (R_1 + R_{-1}\big)\, \partial_x^n \tilde{R}_\psi h\, dx \\
&\qquad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1)\,,
\end{aligned}
$$

where $N := \lfloor \deg^*(\rho\sigma) - \deg(\omega) \rfloor$.
The second term can be written as a sum of integrals of the form

$$
\varepsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_k\, dx
$$

with γ_k, f_k, p_k, q_k just as in the lemma and $\deg^*(\gamma_k) \leq \deg^*(\sigma) - 1$. The functions of the form $f_k = \varepsilon \partial_x^n \tilde{R}_\psi h$, fulfill (2.105) and (2.106) due to the result of lemma 2.2.22 for the map (2.93).
For the other integral, we can exploit (2.88), i.e.

$$
\partial_t (R_1 + R_{-1}) = i\omega (R_1 - R_{-1}) + \varepsilon^{-\beta} \vartheta^{-1} \big(\mathrm{Res}_{u_1}(\varepsilon\Psi) + \mathrm{Res}_{u_{-1}}(\varepsilon\Psi)\big)\,,
$$

and the symmetry of the operator $\frac{\sigma\rho}{\omega^2}$ in order to get

$$
\begin{aligned}
& -j\varepsilon^3 \int_{\mathbb{R}} \frac{\sigma}{\omega} i\rho \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell (R_1 + R_{-1})\, \tilde{R_\psi}\, h\, dx \\
&\quad = -\frac{1}{2} j\, \varepsilon^3\, \partial_t \int_{\mathbb{R}} \frac{\rho\sigma}{\omega^2} \partial_x^\ell (R_{-1} + R_1)\, \partial_x^\ell (R_{-1} + R_1)\, \tilde{R_\psi}\, h\, dx \\
&\qquad + \frac{1}{2} j\, \varepsilon^3 \int_{\mathbb{R}} \Big[\frac{\rho\sigma}{\omega^2}, \tilde{R_\psi} h\Big] i\omega \partial_x^\ell (R_1 - R_{-1})\, \partial_x^\ell (R_{-1} + R_1)\, dx \\
&\qquad + \frac{1}{2} j\, \varepsilon^3 \int_{\mathbb{R}} \frac{\rho\sigma}{\omega^2} \partial_x^\ell (R_{-1} + R_1)\, \partial_x^\ell (R_{-1} + R_1)\, \partial_t \big(\tilde{R_\psi}\, h\big)\, dx \\
&\qquad + \varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}
$$

The first term is the time derivative of an integral $\varepsilon^3\, \tilde{\mathcal{D}}_2$, which can be estimated against $\varepsilon^3 \mathcal{O}(\mathcal{E}_\ell)$ since $\|\tilde{R_\psi} h\|_\infty = \mathcal{O}(1)$.
The last integral can be estimated against $\varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1)$ since $\deg^*(\frac{\rho\sigma}{\omega^2}) \leq 0$ and $\|\partial_t (\tilde{R_\psi} h)\|_\infty = \mathcal{O}(1)$.
The second integral can be written as $\varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)$ plus a sum of integrals of the form

$$
\varepsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_k\, dx
$$

with γ_k, f_k, p_k, q_k just as in the lemma and $\deg^*(\gamma_k) \leq \deg^*(\sigma) - 1$. We see this by using (2.90) and then the result of lemma 2.2.22 for the map (2.93) to see that the emerging functions of the form $f_k = \varepsilon \partial_x^n (h \tilde{R}_\psi)$ fulfill (2.105) and (2.106).
Thus, we obtain

$$\begin{aligned} &\varepsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_{-j} \, h \, dx \\ &\quad = \varepsilon^2 \partial_t \mathcal{D} + \varepsilon^2 \sum_{k=1}^{\deg^*(\sigma)-1} \int_{\mathbb{R}} \varsigma_k \partial_x^\ell R_{j_1} \, \partial_x^\ell R_{j_2} \, \partial_x^k h \, dx \\ &\qquad + \varepsilon^2 \sum_{k=1}^{m} \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k} \, \partial_x^\ell R_{q_k} \, f_k \, dx + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1), \end{aligned}$$

with $\mathcal{D}$, ς_k, m, γ_k, p_k, q_k, f_k just as in the lemma, and have on top of that

$$\deg^*(\gamma_k) \leq \deg^*(\sigma) - 1.$$

b) Handling integrals of the form

$$\varepsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_j \, h \, dx \,. \tag{2.111}$$

Since $i\sigma$ is skew symmetric and due to (2.90) and, (2.91) or (2.92), we have

$$\begin{aligned} \varepsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^\ell R_j \, \partial_x^\ell R_j \, h \, dx &= -\frac{1}{2} \varepsilon^2 \int_{\mathbb{R}} \big[i\sigma, h \big] \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx, \\ &= \varepsilon^2 \sum_{k=1}^{\deg^*(\sigma)-1} \int_{\mathbb{R}} \varsigma_k \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^k h \, dx + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,, \end{aligned}$$

with ς_k just as in the lemma.
More precisely, we could even write this term as a sum of some $\varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1)$-terms and integrals of the form (2.112), which have at least a whole derivative less than the original term.

c) Handling integrals of the form

$$\varepsilon^2 \int_{\mathbb{R}} \upsilon \partial_x^\ell R_j \, \partial_x^\ell R_j \, g \, dx \,. \tag{2.112}$$

By using (2.90), we get

$$
\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \upsilon \partial_x^\ell R_j \, \partial_x^\ell R_j \, g \, dx \\
&\quad = \varepsilon^2 \int_{\mathbb{R}} \big[i\omega, \partial_x^{-1} g \big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx \\
&\qquad + \varepsilon^2 \sum_{n=2}^{\lceil \deg^*(\upsilon) \rceil} \frac{(-1)^n}{(n)!} \int_{\mathbb{R}} i^{n+1} \omega^{(n)} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{n-1} g \, dx \\
&\qquad + \mathcal{O}(\varepsilon^2) \, \| \mathcal{R}(\frac{\upsilon}{\omega'} \partial_x^\ell R_j, \partial_x^{-1} g) \|_{L^2} \| \partial_x^\ell R_j \|_{L^2} \,,
\end{aligned}
$$

where $\|\mathcal{R}(\frac{\upsilon}{\omega'}\partial_x^\ell R_j, \partial_x^{-1} g)\|_{L^2} = \mathcal{O}(\|g\|_{H^{\lceil \deg(\omega) \rceil - 1 + q}} \|R_j\|_{H^\ell})$ (for some $q > 1/2$) as in (2.91) or $\|\mathcal{R}(\frac{\upsilon}{\omega'}\partial_x^\ell R_j, \partial_x^{-1} g)\|_{L^2} = \mathcal{O}(\|\widehat{g}\|_{L^1(\lceil \deg(\omega) \rceil - 1)} \|R_j\|_{H^\ell})$ as in (2.92). The estimate works without any problems since $\deg(\omega) - 1 \geq \deg^*(\upsilon) > 0$.
Now, the second term already has the desired form and the last term is $\varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1)$ such that we only have to look at the first term.
By exploiting the skew symmetry of $i\omega$ and (2.19) (and (2.9)), we have

$$
\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \big[i\omega, \partial_x^{-1} g \big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx \\
&\quad = -j \, \varepsilon^2 \, \partial_t \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_{-1} + R_1) \big) \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_{-1} + R_1) \big) \, \partial_x^{-1} g \, dx \\
&\qquad + j \, \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_t \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,.
\end{aligned}
$$

Due to (2.73), the skew symmetry of $i\rho$ and the symmetry of ω' and υ, we get

$$\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \big[i\omega, \partial_x^{-1} g\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx \\
&\qquad = -j\, \varepsilon^2\, \partial_t \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx \\
&\qquad\quad - 2\varepsilon^3 \int_{\mathbb{R}} i\rho \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, \partial_x^{-1} g \, dx \\
&\qquad\quad - \varepsilon^3 \int_{\mathbb{R}} \big[i\rho \frac{\upsilon}{\omega'}, \partial_x^{-1} g\big] \partial_x^\ell R_j \, \partial_x^\ell \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, dx \\
&\qquad\quad - \varepsilon^3 \int_{\mathbb{R}} \big[i\rho, \partial_x^{-1} g\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, dx \\
&\qquad\quad + j\, \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_t \partial_x^{-1} g \, dx \\
&\qquad\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1).
\end{aligned}$$

The first term is a time derivative of an integral $\varepsilon^2 \tilde{D}$, which can be estimated against $\varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell)$ since $\|\partial_x^{-1} g\|_\infty = \mathcal{O}(1)$.
The last integral can be estimated against $\varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)$.
By using (2.90) and Leibniz's rule, we can write the third and the fourth integral as a sum of some $\varepsilon^3\, \mathcal{O}(\mathcal{E}_\ell + 1)$-terms and integrals of the form

$$\varepsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k} \, \partial_x^\ell R_{q_k} \, f_k \, dx$$

with γ_k, f_k, p_k, q_k just as in the lemma. In order to see that the functions of the form $f_k = \varepsilon\, \partial_x^n \tilde{R}_\psi \partial_x^m g$, with $n, m \geq 0$, in the resulting integrals do fulfill (2.105), (2.106) and (2.100), we use the results of lemma 2.2.22 for the map (2.93).

We now have arrived at

$$\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \upsilon \partial_x^\ell R_j \, \partial_x^\ell R_j \, g \, dx \\
&\qquad = \varepsilon^2 \partial_t \tilde{\mathcal{D}} + \varepsilon^2 \sum_{k=1}^{\deg^*(\upsilon)-1} \int_{\mathbb{R}} \varsigma_k \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^k g \, dx \\
&\qquad\quad + \varepsilon^2 \sum_{k=1}^{\tilde{m}} \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k} \, \partial_x^\ell R_{q_k} \, f_k \, dx \\
&\qquad\quad - 2\varepsilon^3 \int_{\mathbb{R}} i\rho \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, \partial_x^{-1} g \, dx \\
&\qquad\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),
\end{aligned}$$

where $\tilde{D}$, ς_k, γ_k and f_k already satisfy the conditions of lemma 2.2.26.

Using Leibniz's rule and afterwards (2.97) and (2.72), we obtain

$$
\begin{aligned}
-2\varepsilon^3 & \int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^\ell\big(R_\psi\vartheta(R_{-1}+R_1)\big)\,\partial_x^{-1}g\,dx \\
&= -2\varepsilon^3\sum_{n=0}^{\ell}\binom{\ell}{n}\int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^{\ell-n}\vartheta(R_{-1}+R_1)\,\partial_x^n R_\psi\,\partial_x^{-1}g\,dx \\
&= -2\varepsilon^3\int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^\ell(R_{-1}+R_1)\,\tilde{R}_\psi\,\partial_x^{-1}g\,dx \\
&\quad - 2\varepsilon^3\sum_{n=1}^{N}\binom{\ell}{n}\int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^{\ell-n}(R_{-1}+R_1)\,\partial_x^n\tilde{R}_\psi\,\partial_x^{-1}g\,dx \\
&\quad + \varepsilon^3\,\mathcal{O}(\mathcal{E}_\ell+1),
\end{aligned}
$$

where $N := \lfloor \deg^*(\rho\upsilon) - \deg(\omega')\rfloor$.
The second term is a sum of integrals of the form

$$
\varepsilon^2\int_{\mathbb{R}}\gamma_k\partial_x^\ell R_{p_k}\,\partial_x^\ell R_{q_k}\,f_k\,dx
$$

with γ_k, f_k, p_k, q_k just as in the lemma. Moreover, (2.105) and (2.106) are true for the $f_k = \varepsilon\,\partial_x^n\tilde{R}_\psi\partial_x^{-1}g$ with $1\le n\le N$. This follows due to lemma 2.2.22, since we can exploit the fact that

$$
\varepsilon\,\partial_x^n\tilde{R}_\psi\partial_x^{-1}g = \varepsilon\,\partial_x\big(\partial_x^{n-1}\tilde{R}_\psi\partial_x^{-1}g\big) - \varepsilon\,\partial_x^{n-1}\tilde{R}_\psi g
$$

such that we can use the estimates of lemma 2.2.22 for the maps (2.93) and (2.94).

Thus, we now only have to examine the term

$$
\begin{aligned}
-2\varepsilon^3 & \int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^\ell(R_{-1}+R_1)\,\tilde{R}_\psi\,\partial_x^{-1}g\,dx \\
&= -2\varepsilon^3\int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^\ell R_{-j}\,\tilde{R}_\psi\,\partial_x^{-1}g\,dx \\
&\quad - 2\varepsilon^3\int_{\mathbb{R}} i\rho\frac{\upsilon}{\omega'}\partial_x^\ell R_j\,\partial_x^\ell R_j\,\tilde{R}_\psi\,\partial_x^{-1}g\,dx \\
&.
\end{aligned}
$$

The first integral is of the form (2.111) and the second integral of the form (2.110). However, this is not as trivial as it first may seem, since there is the possibility that $\deg^*(\rho\frac{\upsilon}{\omega'}) > \deg^*(\upsilon)$. Thus, in order to see that the function $h = \varepsilon\tilde{R}_\psi\,\partial_x^{-1}g$ does indeed satisfy the conditions (2.99) and (2.101), we also have to use the fact that

$$
\deg^*(\rho\frac{\upsilon}{\omega'}) = \deg^*(\upsilon) + \deg^*(\rho) - \deg(\omega) + 1 \le \deg^*(\upsilon) + 1 \le \deg(\omega) \qquad (2.113)
$$

and exploit lemma 2.2.22 for the map (2.94).
Due to (2.113), the lemma is now finally proven by applying the result of paragraph a2) in this proof to the first integral and the result of paragraph b) to the second integral. □

Corollary 2.2.28. *Let $\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$.*
Let the pseudo-differential operator γ and the function f be exactly as in lemma 2.2.26. Then, for $0 < \varepsilon < \varepsilon_0$ and ε_0 small enough there exists an expression $\mathcal{D}$ with

$$\mathcal{D} = \mathcal{O}(\mathcal{E}_\ell)\,,$$

such that

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx = \varepsilon^2\, \partial_t \mathcal{D} + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1). \tag{2.114}$$

Proof. Due to lemma 2.2.26, we have

$$\begin{aligned} &\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &\quad = \varepsilon^2 \partial_t \mathcal{D} + \varepsilon^2 \sum_{k=1}^{\lceil \deg^*(\gamma) \rceil - 1} \int_{\mathbb{R}} \varsigma_k \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, \partial_x^k f\, dx \\ &\quad\quad + \varepsilon^2 \sum_{k=1}^{m} \int_{\mathbb{R}} \gamma_k \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_k\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1) \end{aligned}$$

with $\varsigma_k, \gamma_k, m, f_k, p_q$ and q_k just as in the lemma. Moreover, we can apply 2.2.26 repeatedly, i.e. we can always apply 2.2.26 again to every integral on above right hand side.
If $\deg^*(\gamma) < \deg(\omega)$, we can repeatedly use (2.102) until we obtain

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx = \varepsilon^2 \partial_t \tilde{\mathcal{D}} + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),$$

for some $\tilde{\mathcal{D}}$ with $\tilde{\mathcal{D}} = \mathcal{O}(\mathcal{E}_\ell)$. This is because due to (2.104) and (2.103), every time we apply (2.102), the resulting integrals will contain $\big(\deg(\omega) - \deg^*(\rho)\big)$ derivatives or an whole derivative less falling on $\partial_x^\ell R_{-1}$ or $\partial_x^\ell R_1$ than the previous integrals such that the above result is achieved after a finite number of steps.

If $\deg^*(\gamma) = \deg(\omega)$, we can use (2.102) and exploit (2.103) in order to get

$$\begin{aligned} &\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &\quad = \varepsilon^2 \partial_t \tilde{\mathcal{D}} + \varepsilon^2 \sum_{k=1}^{m_\gamma} \int_{\mathbb{R}} \tilde{\gamma} \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, \tilde{f}_k\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1), \end{aligned}$$

for some $\tilde{\mathcal{D}}$ with $\tilde{\mathcal{D}} = \mathcal{O}(\mathcal{E}_\ell)$ and some $m_\gamma = m_\gamma(\deg^*(\gamma)) \in \mathbb{N}$. Herby, we have

$$\|\tilde{f}_k\|_{H^{\lceil \deg^*(\tilde{\gamma}_k)\rceil \lceil \deg(\omega)\rceil}} + \|\partial_t \tilde{f}_k\|_{H^{\lceil \deg^*(\tilde{\gamma}_k)\rceil}} \leq C_f,$$

$$\|\partial_x^{-1} \tilde{f}_k\|_\infty + \|\partial_t \partial_x^{-1} \tilde{f}_k\|_{C^{\lceil \deg^*(\tilde{\gamma}_k)\rceil}} \leq C_f\,,$$

for some constant $C_f = C_f(\tilde{R}_\psi, f, \gamma) > 1$ due to (2.105) and (2.106).
By using (2.102) and exploiting (2.103) again for every integral on the above right-hand side, we can obtain an expression $\check{\mathcal{D}} = \mathcal{O}(\mathcal{E}_\ell)$ such that we have

$$\begin{aligned}\varepsilon^2 &\int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &= \varepsilon^2 \partial_t \check{\mathcal{D}} + \varepsilon^2 \sum_{k=1}^{\check{m}} \int_{\mathbb{R}} \check{\gamma}_k \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, \check{f}_k\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),\end{aligned}$$

where $\check{m} = m_\gamma^2$,

$$\|\check{f}_k\|_{H^{\lceil \deg^*(\check{\gamma}_k)\rceil \lceil \deg(\omega)\rceil}} + \|\partial_t \check{f}_k\|_{H^{\lceil \deg^*(\check{\gamma}_k)\rceil}} \leq \varepsilon^{1/2}\, C_f^2\,,$$

$$\|\partial_x^{-1} \check{f}_k\|_\infty + \|\partial_t \partial_x^{-1} \check{f}_k\|_{C^{\lceil \deg^*(\check{\gamma}_k)\rceil}} \leq C_f^2\,.$$

By repeating the last step one more time, we now get

$$\begin{aligned}\varepsilon^2 &\int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &= \varepsilon^2 \partial_t \mathcal{D}_0 + \varepsilon^{2+1/2} \sum_{k=1}^{m_0} \int_{\mathbb{R}} \gamma_{k,0} \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_{k,0}\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),\end{aligned}$$

for some $\mathcal{D}_0$ with $\mathcal{D}_0 = \mathcal{O}(\mathcal{E}_\ell)$, $m_0 = m_\gamma^3$ and

$$\|f_{k,0}\|_{H^{\lceil \deg^*(\gamma_{k,0})\rceil + \lceil \deg(\omega)\rceil}} + \|\partial_t f_{k,0}\|_{H^{\lceil \deg^*(\gamma_{k,0})\rceil}} \leq \varepsilon^{1/2}\, C_f^3\,,$$

$$\|\partial_x^{-1} f_{k,0}\|_\infty + \|\partial_t \partial_x^{-1} f_{k,0}\|_{C^{\lceil \deg^*(\gamma_{k,0})\rceil}} \leq C_f^3\,.$$

After N additional steps, we get

$$\begin{aligned}\varepsilon^2 &\int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx \\ &= \varepsilon^2 \sum_{p=0}^{N} \varepsilon^{p/2}\, \partial_t \mathcal{D}_p + \varepsilon^2\, \varepsilon^{\frac{N+1}{2}} \sum_{k=1}^{m_N} \int_{\mathbb{R}} \gamma_{k,N} \partial_x^\ell R_{p_k}\, \partial_x^\ell R_{q_k}\, f_{k,N}\, dx + \varepsilon^2 \sum_{p=0}^{N} \varepsilon^{p/2}\, \mathcal{C}_p,\end{aligned}$$

for some expressions $\mathcal{D}_p$ with $\mathcal{D}_p = \mathcal{O}(\mathcal{E}_\ell)$, some $\mathcal{C}_p = \mathcal{O}(\mathcal{E}_\ell + 1)$, $m_N = m_\gamma^{3+N}$ and

$$\|f_{k,N}\|_{H^{\lceil \deg^*(\gamma_{k,N})\rceil + \lceil \deg(\omega)\rceil}} + \|\partial_t f_{k,N}\|_{H^{\lceil \deg^*(\gamma_{k,N})\rceil}} \leq \varepsilon^{1/2}\, C_f^{3+N}\,,$$

$$\|\partial_x^{-1} f_{k,N}\|_\infty + \|\partial_t \partial_x^{-1} f_{k,N}\|_{C^{\lceil \deg^*(\gamma_{k,N})\rceil}} \leq C_f^{3+N}\,.$$

Moreover, we have $\deg^*(\gamma_{k,N}) \leq \deg^*(\gamma)$ due to (2.104).

We will now show that

$$\mathcal{D}^\infty := \sum_{p=0}^{\infty} \varepsilon^{p/2}\, \mathcal{D}_p$$

does exist, that $\mathcal{D}^\infty = \mathcal{O}(\mathcal{E}_\ell)$, and that

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx = \varepsilon^2\, \partial_t \mathcal{D}^\infty + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1).$$

By taking a close look at the proof of (2.102), we find that

$$\begin{aligned} \varepsilon^{\frac{p}{2}}\, \mathcal{D}_p &\leq \varepsilon^{\frac{p}{2}}\, m_\gamma^{p+3} c_\gamma^{p+3} C_f^{p+3}\, \mathcal{E}_\ell, \\ \varepsilon^{\frac{p}{2}}\, \mathcal{C}_p &\leq \varepsilon^{\frac{p}{2}}\, m_\gamma^{p+3} c_\gamma^{p+3} C_f^{p+3}\, \big(\mathcal{E}_\ell + 1\big), \end{aligned}$$

for some $c_\gamma = c_\gamma(\deg^*(\gamma)) > 1$ as long as $f, i\rho, i\omega$ and ℓ are fixed. We emphasize that this is in particular possible due to the fact that $\deg^*(\gamma_{k,N}) \leq \deg^*(\gamma)$.
Now, by choosing ε small enough, for instance such that

$$\varepsilon^{1/4}\, m_\gamma c_\gamma C_f \leq 1\,,$$

we get the following.
There is a $c \in \mathbb{R}$ such that

$$\mathcal{D}^\infty = \sum_{p=0}^{\infty} \varepsilon^{p/2}\, \mathcal{D}_p \leq \sum_{p=0}^{\infty} \varepsilon^{p/2}\, |\mathcal{D}_p| \leq \sum_{p=0}^{\infty} \varepsilon^{p/4}\, c\, \mathcal{E}_\ell = c\, \mathcal{E}_\ell \sum_{p=0}^{\infty} \varepsilon^{p/4} = \mathcal{O}(\mathcal{E}_\ell)\,,$$

analogously we get

$$\sum_{p=0}^{\infty} \varepsilon^{p/2}\, C_p \leq \sum_{p=0}^{\infty} \varepsilon^{p/2}\, |C_p| = \mathcal{O}(\mathcal{E}_\ell + 1)\,.$$

Moreover,

$$\begin{aligned} &\varepsilon^{\frac{N+1}{2}} \sum_{k=1}^{m_N} \int_{\mathbb{R}} \gamma_{k,N} \partial_x^\ell R_{j_k}\, \partial_x^\ell R_{l_k}\, f_{k,N}\, dx \\ &\qquad \leq \varepsilon^{\frac{N+1}{4}}\, C_f^2 \Big(\|R_1\|_{H^\ell} \|R_1\|_{C^{\ell + \lceil \deg^*(\gamma)\rceil}} + \|R_1\|_{H^\ell} \|R_{-1}\|_{C^{\ell + \lceil \deg^*(\gamma)\rceil}} \\ &\qquad\qquad + \|R_{-1}\|_{H^\ell} \|R_1\|_{C^{\ell + \lceil \deg^*(\gamma)\rceil}} + \|R_{-1}\|_{H^\ell} \|R_{-1}\|_{C^{\ell + \lceil \deg^*(\gamma)\rceil}} \Big) \\ &\qquad = 0, \quad \text{for } N \to \infty. \end{aligned}$$

We now obtain

$$\begin{aligned}\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx &= \varepsilon^2\, \partial_t \mathcal{D}^\infty + \varepsilon^2 \sum_{p=0}^{\infty} \varepsilon^{p/2}\, C_p \\ &\quad + \varepsilon^2 \lim_{N\to\infty} \varepsilon^{\frac{N+1}{2}} \sum_{k=1}^{m_N} \int_{\mathbb{R}} \gamma_{k,N} \partial_x^\ell R_{q_k}\, \partial_x^\ell R_{p_k}\, f_{k,N}\, dx \\ &= \varepsilon^2\, \partial_t \mathcal{D}^\infty + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1).\end{aligned}$$

□

Remark 2.2.29. *The short involvement of the $C^{\ell+\lceil \deg^* \rho\rceil}$-norm is not problematic since the final estimates here do no longer involve this norm. More precisely, using some mollifiers ι_m and looking at $R_j^m := R_j * \iota_m$, one would take the above limit of $\varepsilon^{\frac{N+1}{4}}$ for $N \to \infty$ before going over to the limit $R_j^m \to R_j$ for the final energy estimates. Therefore, $\tilde{\mathcal{E}}_\ell(R_{-1}^m, R_1^m)$ would converge against $\tilde{\mathcal{E}}_\ell(R_{-1}, R_1)$ and $\partial_t \tilde{\mathcal{E}}_\ell(R_{-1}^m, R_1^m)$ would converge uniformly against $\partial_t \tilde{\mathcal{E}}_\ell(R_{-1}, R_1)$.*

Corollary 2.2.30. *Let $\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$.*
For $\varepsilon < \varepsilon_0$ and ε_0 sufficiently small, there exists an energy $\tilde{\mathcal{E}}_\ell$ and some constants $c, C > 0$ such that

$$\big(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\big)^2 \leq c\,\tilde{\mathcal{E}}_\ell \leq C\,\big(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\big)^2 \tag{2.115}$$

and

$$\partial_t \tilde{\mathcal{E}}_\ell \leq \varepsilon^2\,\mathcal{O}\big(\tilde{\mathcal{E}}_\ell + 1\big)\,.$$

Proof. According to the definition of $\mathcal{E}_\ell$ in (2.65) and due to lemma 2.2.9, we have

$$\partial_t \mathcal{E}_\ell = \partial_t E_0 + \partial_t E_\ell = \partial_t E_\ell + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,,$$

where, due to lemma 2.2.17,

$$\begin{aligned}
\partial_t E_\ell \;&=\; \varepsilon^2\,V_\ell + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1),\\
&=\; \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} \Big(j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho \partial_x^\ell \vartheta^{-1} (R_Q \vartheta R_{j_3})\,dx\\
&\qquad\qquad + j_1 \int_{\mathbb{R}} i\rho\partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\,\partial_x^\ell \vartheta^{-1} N_{j_1 j_1}(\psi_c, R_{j_1})\,dx\\
&\qquad\qquad + j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\,\partial_x^\ell \vartheta^{-1} N_{j_1 j_1}(\psi_c, i\rho\vartheta^{-1}(R_\psi \vartheta R_{j_3}))\,dx\\
&\qquad\qquad + j_1 \int_{\mathbb{R}} i\rho\partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\,\partial_x^\ell \vartheta^{-1} N_{j_1 -j_1}(\psi_c, R_{-j_1})\,dx\\
&\qquad\qquad - j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\,\partial_x^\ell \vartheta^{-1} N_{j_1 -j_1}(\psi_c, i\rho\vartheta^{-1}(R_\psi \vartheta R_{j_3}))\,dx \Big)\\
&\qquad + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\\
&=:\; \sum_{i=0}^{4} I_i + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}$$

First, we analyze the term I_0.
Using (2.73), we get

$$\begin{aligned}
I_0 :&= \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, i\rho\partial_x^\ell \vartheta^{-1}(R_Q \vartheta R_{j_3})\,dx \qquad (2.116)\\
&= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell (R_1 - R_{-1})\, i\rho \partial_x^\ell \big(R_Q \vartheta (R_{-1} + R_1)\big)\,dx + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}$$

Due to Leibniz's rule and the definition $R_Q = \psi_Q + \frac{1}{2}\varepsilon^{\beta-2}\vartheta(R_{-1}+R_1)$, we obtain by proceeding analogously as in (2.97), setting

$$\tilde{R}_Q := \psi_Q + \varepsilon^{\beta-2}\vartheta(R_{-1}+R_1) \tag{2.117}$$

and using (2.72), that

$$I_0 := \varepsilon^2 \sum_{n=0}^{N} \binom{\ell}{n} \int_{\mathbb{R}} \partial_x^{\ell}(R_1 - R_{-1})\, i\rho\big(\partial_x^n \tilde{R}_Q \partial_x^{\ell-n}(R_{-1}+R_1)\big)\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)\ ,$$

where $N := \lceil \deg^*(\rho\rceil - 1$.
After replacing the expression

$$i\rho\big(\partial_x^n \tilde{R}_Q \partial_x^{\ell-n}(R_{-1}+R_1)\big)$$

by

$$\partial_x^n \tilde{R}_Q\, i\rho \partial_x^{\ell-n}(R_{-1}+R_1) + \big[i\rho, \partial_x^n \tilde{R}_Q\big]\partial_x^{\ell-n}(R_{-1}+R_1)$$

and using (2.90), we can use corollary 2.2.28 in order to obtain

$$I_0 = \varepsilon^2\, \partial_t \mathcal{D}_0 + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)$$

for some $\mathcal{D}_0$ with $\varepsilon^2\mathcal{D}_0 = \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell)$.

Now, we analyze the term $I_1 + I_2$.
Using (2.73) we get

$$\begin{aligned} I_1 + I_2 := \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} & \Big(j_1 \int_{\mathbb{R}} i\rho \partial_x^{\ell}\vartheta^{-1}(R_\psi \vartheta R_{j_3})\, \partial_x^{\ell}\vartheta^{-1} N_{j_1 j_1}(\psi_c, R_{j_1})\, dx \\ & + j_1 \int_{\mathbb{R}} \partial_x^{\ell} R_{j_1}\, \partial_x^{\ell}\vartheta^{-1} N_{j_1 j_1}\big(\psi_c, i\rho\vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\ = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} & j_1 \Big(\int_{\mathbb{R}} i\rho \partial_x^{\ell}\vartheta^{-1}(R_\psi \vartheta R_{j_3})\, \partial_x^{\ell} N_{j_1 j_1}(\psi_c, R_{j_1})\, dx \\ & + \int_{\mathbb{R}} \partial_x^{\ell} R_{j_1}\, \partial_x^{\ell} N_{j_1 j_1}\big(\psi_c, i\rho\vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\ & + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,. \end{aligned} \tag{2.118}$$

In order to extract all terms with more than ℓ spatial derivatives falling on R_1 or

R_{-1}, we apply Leibniz's rule and get

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(& \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N_{j_1 j_1}(\partial_x^m \psi_c, \partial_x^{\ell-m} R_{j_1})\, dx \\
& + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 j_1}\big(\psi_c, i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \\
& + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil} \binom{\ell}{m} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 j_1}\big(\partial_x^m \psi_c, i\rho \partial_x^{\ell-m} \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\
& + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2})\,.
\end{aligned}
$$

Notice that this is since the $N_{j_1 j_1}(\partial_x^m \psi_c, \cdot)$ map $H^1(\mathbb{R})$ on $L^2(\mathbb{R})$ due to (1.9) (see lemma 2.2.5).

By using (2.68), we get

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(& \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& + \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N^*_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N_{j_1 j_1}(\partial_x^m \psi_c, \partial_x^{\ell-m} R_{j_1})\, dx \\
& + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^{\ell-m} \vartheta^{-1}(R_\psi \vartheta R_{j_3})\, N^*_{j_1 j_1}(\partial_x^m \psi_c, \partial_x^\ell R_{j_1})\, dx \Big) \\
& + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2})\,,
\end{aligned}
$$

where

$$
\widehat{N}^*_{j_1 j_2}(\psi_c, f)(k) := \int_{\mathbb{R}} n_{j_1 j_2}(-m, k-m, -k)\, \widehat{\psi}_c(k-m)\, \widehat{f}(m)\, dm\,.
$$

For all integrals except the first two, we can now exploit (2.52), or respectively (2.53) or (2.54), together with (2.73), Leibniz's rule and (2.72) such that corollary 2.2.28 can be applied. The exact details on how this is done should soon become clear in this proof. We get

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(& \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3}) \, N_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1}) \, dx \\
& + \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3}) \, N^*_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1}) \, dx \Big) \\
& + \varepsilon^2 \, \partial_t \mathcal{D}_{1,2} + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,,
\end{aligned}
$$

for some $\mathcal{D}_{1,2}$ with $\varepsilon^2 \, \mathcal{D}_{1,2} = \varepsilon \, \mathcal{O}(\mathcal{E}_\ell)$.

For the remaining terms, we cannot apply corollary 2.2.28 since in these integrals there are more than $\deg^*(\rho)$ derivatives falling on $\partial_x^\ell R_{-1}$ or $\partial_x^\ell R_1$. We use the skew symmetry of $i\rho$ and exploit (2.73) in order to get

$$
\begin{aligned}
I_1 + I_2 = -\varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell (R_\psi \vartheta R_{j_3}) \, i\rho \big(N_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1}) + N^*_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1}) \big) \, dx \\
+ \varepsilon^2 \, \partial_t \mathcal{D}_{1,2} + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,.
\end{aligned}
$$

If we now look at

$$
\begin{aligned}
& i\rho(k) \, \big(n_{jj}(k, k-m, m) + n_{jj}(-m, k-m, -k) \big) \\
& \qquad = i\rho(k) \, (\rho(k) - \rho(m)) \, \frac{\chi_c(k-m)}{\omega(k) - \omega(m) + j \, \omega(k-m)} \qquad (\text{for } |k| \to \infty) \,,
\end{aligned}
$$

and use Taylor, the same cancellation as in the proof of corollary 2.2.16 occurs such that by exploiting (2.52), or respectively (2.53) or (2.54), we obtain

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \sum_{n=1}^{N} \int_{\mathbb{R}} \partial_x^\ell (R_\psi \vartheta R_{j_3}) \, \beta_n \psi_c \, \alpha_n \partial_x^\ell R_{j_1} \, dx \\
+ \varepsilon^2 \, \partial_t \mathcal{D}_{1,2} + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,,
\end{aligned}
$$

for some $N \in \mathbb{N}$ and some pseudo-differential operators β_n and α_n. Here α_n is either skew-symmetric with $\deg^*(\alpha_n) \leq \deg^*(\rho)$ or symmetric with $\deg^*(\alpha_n) \leq \deg^*(\rho) - 1$. Since we now have a derivative less falling on $\partial_x^\ell R_{-1}$ or $\partial_x^\ell R_1$, we can apply corollary 2.2.28 (after we used Leibniz's rule and (2.72)) and obtain

$$
I_1 + I_2 = \varepsilon^2 \, \partial_t \tilde{\mathcal{D}}_{1,2} + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1)
$$

for some $\tilde{\mathcal{D}}_{1,2}$ with $\varepsilon^2 \, \tilde{\mathcal{D}}_{1,2} = \varepsilon \, \mathcal{O}(\mathcal{E}_\ell)$.
When we applied corollary 2.2.28 to the integrals of the form

$$
\int_{\mathbb{R}} \partial_x^\ell R_{j_3} \, \alpha_n \partial_x^{\ell - m} R_{j_1} \, \partial_x^{m_1} \tilde{R}_\psi \, \partial_x^{m_2} \beta_n \psi_c \, dx \,,
$$

with $m, m_1, m_2 \geq 0$, we had to proceed considerately. One small thing was that to obtain (2.101) we spited the functions $f = \partial_x^{m_1} \tilde{R}_\psi \partial_x^{m_2} \beta_n \psi_c$ into

$$f = \partial_x^{m_1} (\psi_c + \varepsilon \tilde{R}_Q) \partial_x^{m_2} \beta_n \psi_c = \partial_x^{m_1} \psi_c \partial_x^{m_2} \beta_n \psi_c + \partial_x^{m_1} \varepsilon \tilde{R}_Q \partial_x^{m_2} \beta_n \psi_c =: f_1 + f_2$$

such that $\|f_1\|_{L^1(s)} = \mathcal{O}(1)$ and $\|f_2\|_{H^s} = \mathcal{O}(1)$.
Another thing and also the reason for which we only get $\varepsilon^2 \tilde{\mathcal{D}}_{1,2} = \varepsilon\, \mathcal{O}(\mathcal{E}_\ell)$ is that, for functions of the form $g = \gamma \psi_c \beta \psi_c$, we have to rely on the estimate

$$\|\partial_x^{-1} g\|_\infty = \|\partial_x^{-1}(\gamma \psi_c \beta \psi_c)\|_\infty \leq \int_{\mathbb{R}} |\gamma \psi_c \beta \psi_c| \, dx \leq \|\gamma \psi_c\|_{L^2} \|\beta \psi_c\|_{L^2} \,,$$

such that $\|\partial_x^{-1} g\|_\infty \leq \mathcal{O}(\varepsilon^{-1})$. A similarly bad estimate for $\|\partial_t \partial_x^{-1} g\|_\infty$ could be avoided by estimating in the following way. We exploit that $\psi_c = \psi_{-1} + \psi_1$ and

$$\gamma \psi_j = \gamma(j k_0) \psi_j + \mathcal{O}(\varepsilon) \,, \tag{2.119}$$

for operators γ that are given in Fourier space by their sufficiently smooth symbol γ, see section 2.1. Thus we have

$$\gamma \psi_c \beta \psi_c = \sum_{j_1, j_2 \in \{\pm 1\}} \gamma(j k_0) \beta(j k_0) \psi_{j_1} \psi_{j_2} + \mathcal{O}(\varepsilon)$$

with

$$\psi_{j_1}(x,t) \psi_{j_2}(x,t) = A_{j_1}(\varepsilon(x - c_g t), \varepsilon^2 t) A_{j_2}(\varepsilon(x - c_g t), \varepsilon^2 t) e^{i(j_1 + j_2)(k_0 x - \omega_0 t)},$$

where $A_{-1} := \overline{A}_1$ and A_1 is as in (2.8).
Since $\psi_j \psi_j$ is strictly concentrated around $k = \pm 2k_0$ in Fourier space, we can compute $\|\partial_x^{-1}(\psi_j \psi_j)\|_{L^1(s)} = \mathcal{O}(1)$ and $\|\partial_t \partial_x^{-1}(\psi_j \psi_j)\|_{L^1(s)} = \mathcal{O}(1)$.
Due to the fact that $\psi_j \psi_{-j}$ is strictly concentrated around $k = 0$ in Fourier space, we have to stick with the estimate $\|\partial_x^{-1}(\psi_j \psi_{-j})\|_\infty = \mathcal{O}(\varepsilon^{-1})$. However, we can obtain $\|\partial_t \partial_x^{-1}(\psi_j \psi_{-j})\|_\infty = \mathcal{O}(1)$ because

$$\partial_t(\psi_j \psi_{-j}) = \partial_t |A_1(\varepsilon(x - c_g t), \varepsilon^2 t)|^2 = \mathcal{O}(\varepsilon).$$

Thus, we have the estimates

$$\|\partial_x^{-1}(\gamma \psi_c \beta \psi_c)\|_\infty \leq \mathcal{O}(\varepsilon^{-1}), \qquad \|\partial_t \partial_x^{-1}(\gamma \psi_c \beta \psi_c)\|_\infty = \mathcal{O}(1) \,.$$

Due to the first of the above estimates, an extra step was needed to enable the application of corollary 2.2.28, we give the full details on this technicality directly after this proof in lemma 2.2.31.

Now, we analyze the term $I_3 + I_4$.
Using (2.73), we have

$$
\begin{aligned}
I_3 + I_4 :&= \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} \Big(j_1 \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3})\, \partial_x^\ell \vartheta^{-1} N_{j_1 - j_1}(\psi_c, R_{-j_1})\, dx \qquad (2.120) \\
&\qquad - j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 - j_1}\big(\psi_c, i\rho \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\
&= \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(\int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3})\, \partial_x^\ell N_{j_1 - j_1}(\psi_c, R_{-j_1})\, dx \\
&\qquad - \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell N_{j_1 - j_1}\big(\psi_c, i\rho \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\
&\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta - 1} \mathcal{E}_\ell^{3/2}).
\end{aligned}
$$

Applying Leibniz's rule, (2.38) and exploiting that the $N_{j_1 - j_1}(\partial_x^m \psi_c, \cdot)$ always map $L^2(\mathbb{R})$ on $L^2(\mathbb{R})$ due to (1.9) (see lemma 2.2.5), we get

$$
\begin{aligned}
&I_3 + I_4 \\
&= \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(\int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3})\, N_{j_1 - j_1}(\psi_c, \partial_x^\ell R_{-j_1})\, dx \\
&\qquad + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil - 1} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^\ell \vartheta^{-1} (R_\psi \vartheta R_{j_3})\, N_{j_1 - j_1}(\partial_x^m \psi_c, \partial_x^{\ell - m} R_{-j_1})\, dx \\
&\qquad - \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 - j_1}\big(\psi_c, i\rho \partial_x^\ell \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \\
&\qquad - \sum_{m=1}^{\lceil \deg^*(\rho) \rceil - 1} \binom{\ell}{m} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 - j_1}\big(\partial_x^m \psi_c, i\rho \partial_x^{\ell - m} \vartheta^{-1}(R_\psi \vartheta R_{j_3})\big)\, dx \Big) \\
&\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta - 1} \mathcal{E}_\ell^{3/2}).
\end{aligned}
$$

Using (2.66) and (2.73), we get

$$
\begin{aligned}
I_3 + I_4 = \varepsilon^2 \sum_{j_1, j_3 \in \{\pm 1\}} j_1 \Big(& \int_{\mathbb{R}} i\rho \partial_x^\ell (R_\psi \vartheta R_{j_3})\, N_{j_1 - j_1}(\psi_c, \partial_x^\ell R_{-j_1})\, dx \\
& - \int_{\mathbb{R}} i\rho \partial_x^\ell (R_\psi \vartheta R_{j_3})\, N^*_{j_1 - j_1}(\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& + \sum_{m=1}^{\lceil \deg^*(\rho) \rceil - 1} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^\ell (R_\psi \vartheta R_{j_3})\, N_{j_1 - j_1}(\partial_x^m \psi_c, \partial_x^{\ell - m} R_{-j_1})\, dx \\
& - \sum_{m=1}^{\lceil \deg^*(\rho) \rceil - 1} \binom{\ell}{m} \int_{\mathbb{R}} i\rho \partial_x^{\ell - m} (R_\psi \vartheta R_{j_3})\ N^*_{j_1 - j_1}(\partial_x^m \psi_c, \partial_x^\ell R_{j_1})\, dx \Big) \\
+ \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta - 1} \mathcal{E}_\ell^{3/2}).
\end{aligned}
$$

After exploiting (2.55), Leibniz's rule and (2.72), we can now apply corollary 2.2.28 in order to obtain

$$I_3 + I_4 = \varepsilon^2\, \partial_t \mathcal{D}_{3,4} + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1),$$

for some $\mathcal{D}_{3,4}$ with $\varepsilon^2\, \mathcal{D}_{3,4} = \varepsilon\, \mathcal{O}(\mathcal{E}_\ell)$.

Hence, by choosing ε_0 small enough and summing up our results for I_0-I_4, we can define a modified energy

$$\tilde{\mathcal{E}}_\ell = \mathcal{E}_\ell - \varepsilon^2 (\mathcal{D}_0 + \tilde{\mathcal{D}}_{1,2} + \mathcal{D}_{3,4}) \,,$$

with

$$\varepsilon^2\, (\mathcal{D}_0 + \tilde{\mathcal{D}}_{1,2} + \mathcal{D}_{3,4}) = \varepsilon\, \mathcal{O}(\mathcal{E}_\ell)$$

such that

$$\partial_t \tilde{\mathcal{E}}_\ell \lesssim \varepsilon^2 \big(1 + \mathcal{E}_\ell\big) \,.$$

Since $\tilde{\mathcal{E}}_\ell = \mathcal{E}_\ell + \varepsilon\, \mathcal{O}(\mathcal{E}_\ell)$, the statement is now proven with corollary 2.2.16. □

Here are the details on the application of corollary 2.2.28 for $I_1 + I_2$ and $I_4 + I_4$.

Lemma 2.2.31. *Let $\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$.*
Let the pseudo-differential operator γ and the function f be as in lemma 2.2.26 with the only exception being that $\|\partial_x^{-1} g\|_\infty = \mathcal{O}(\varepsilon^{-1})$ for $f = g$. Then, for $0 < \varepsilon < \varepsilon_0$ and ε_0 small enough there exists an expression $\mathcal{D}$ with

$$\varepsilon^2\, \mathcal{D} = \varepsilon\, \mathcal{O}(\mathcal{E}_\ell) \,,$$

such that

$$\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\ell R_{j_1}\, \partial_x^\ell R_{j_2}\, f\, dx = \varepsilon^2\, \partial_t \mathcal{D} + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1). \tag{2.121}$$

Proof. We proceed as in b) in the proof of lemma 2.2.26.
By using (2.90), we get

$$\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \upsilon \partial_x^\ell R_j \, \partial_x^\ell R_j \, g \, dx \\
&\quad = \varepsilon^2 \int_{\mathbb{R}} \big[i\omega, \partial_x^{-1} g\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx \\
&\qquad + \varepsilon^2 \sum_{n=2}^{\lceil \deg^*(\upsilon) \rceil} \frac{(-1)^n}{(n)!} \int_{\mathbb{R}} i^n \omega^{(n)} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{n-1} g \, dx \\
&\qquad + \mathcal{O}(\varepsilon^2) \, \|\mathcal{R}(\frac{\upsilon}{\omega'} \partial_x^\ell R_j, \partial_x^{-1} g)\|_{L^2} \|\partial_x^\ell R_j\|_{L^2} \,.
\end{aligned}$$

The last term can be estimated against $\varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1)$ with (2.91) and (2.92), especially since only derivatives of $\partial_x^{-1} g$ are involved. The integrals of the sum

$$\varepsilon^2 \sum_{n=2}^{\lceil \deg^*(\upsilon) \rceil} \frac{(-1)^n}{(n)!} \int_{\mathbb{R}} i^n \omega^{(n)} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{n-1} g \, dx$$

are also no longer problematic since there is at least one derivative falling on g.
Proceeding further as in b) in the proof of lemma 2.2.26, we have

$$\begin{aligned}
&\varepsilon^2 \int_{\mathbb{R}} \big[i\omega, \partial_x^{-1} g\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, dx \\
&\quad = -j \, \varepsilon^2 \, \partial_t \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, i\rho \partial_x^\ell \vartheta^{-1} \big(R_\psi \vartheta (R_{-1} + R_1)\big) \, \partial_x^{-1} g \, dx \\
&\qquad + j \, \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_t \partial_x^{-1} g \, dx \\
&\qquad + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,.
\end{aligned}$$

We obtain

$$\varepsilon^2 D := \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_x^{-1} g \, dx = \varepsilon \mathcal{O}(\mathcal{E}_\ell) \,,$$

since $\|\partial_x^{-1} g\|_\infty = \mathcal{O}(\varepsilon^{-1})$.
For the last integral, we have

$$j \, \varepsilon^2 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j \, \partial_x^\ell R_j \, \partial_t \partial_x^{-1} g \, dx \leq \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \,,$$

since $\|\partial_t \partial_x^{-1} g\|_\infty = \mathcal{O}(1)$.
For the integrals

$$\begin{aligned}
&\varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} i\rho \partial_x^\ell \vartheta^{-1}\big(R_\psi \vartheta(R_{-1}+R_1)\big)\, \partial_x^\ell R_j\, \partial_x^{-1} g\, dx \\
&\quad + \varepsilon^3 \int_{\mathbb{R}} \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, i\rho \partial_x^\ell \vartheta^{-1}\big(R_\psi \vartheta(R_{-1}+R_1)\big)\, \partial_x^{-1} g\, dx \\
&\qquad\qquad = -2\varepsilon^3 \int_{\mathbb{R}} i\rho \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, \partial_x^\ell\big(R_\psi \vartheta(R_{-1}+R_1)\big)\, \partial_x^{-1} g\, dx \\
&\qquad\qquad\quad - \varepsilon^3 \int_{\mathbb{R}} \big[i\rho \frac{\upsilon}{\omega'}, \partial_x^{-1} g\big] \partial_x^\ell R_j\, \partial_x^\ell\big(R_\psi \vartheta(R_{-1}+R_1)\big)\, dx \\
&\qquad\qquad\quad - \varepsilon^3 \int_{\mathbb{R}} \big[i\rho, \partial_x^{-1} g\big] \frac{\upsilon}{\omega'} \partial_x^\ell R_j\, \partial_x^\ell\big(R_\psi \vartheta(R_{-1}+R_1)\big)\, dx \\
&\qquad\qquad\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}$$

corollary 2.2.28 can now be applied since $\varepsilon \|\partial_x^{-1} g\|_\infty = \mathcal{O}(1)$. In detail, we have the following estimates.
We have (2.99) and (2.100), where we need it:

- For $m \geq 0$ and n as required, we have

  $$\|\varepsilon \partial_x^m \tilde{R}_\Psi \partial_x^{-1} g\|_{H^n} \leq \varepsilon\, \|\partial_x^m \tilde{R}_\Psi\|_{H^n} \|\partial_x^{-1} g\|_{C^n} = \mathcal{O}(\varepsilon^{-1/2})\,,$$

  and similar

  $$\|\varepsilon \partial_t (\partial_x^m \tilde{R}_\Psi \partial_x^{-1} g)\|_{H^n} = \mathcal{O}(\varepsilon^{-1/2})\,.$$

- We also have for $m \geq 0$ and n as required,

  $$\|\varepsilon \partial_x^m \tilde{R}_\Psi \partial_x^{-1} g\|_{C^n} \leq \varepsilon\, \|\partial_x^m \tilde{R}_\Psi\|_{C^n} \|\partial_x^{-1} g\|_{C^n} = \mathcal{O}(1)\,,$$

  and similar

  $$\|\varepsilon \partial_t (\partial_x^m \tilde{R}_\Psi \partial_x^{-1} g)\|_{C^n} = \mathcal{O}(1)\,.$$

- For for $m > 0$ and n as required, we have

  $$\begin{aligned}
  \|\varepsilon \partial_x^{-1} (\partial_x^m \tilde{R}_\Psi \partial_x^{-1} g)\|_\infty &\leq \varepsilon \|\partial_x^{m-1} \tilde{R}_\Psi \partial_x^{-1} g - \partial_x^{-1}(\partial_x^{m-1} \tilde{R}_\Psi g)\|_\infty \\
  &\leq \varepsilon \|\partial_x^{m-1} \tilde{R}_\Psi\|_\infty \|\partial_x^{-1} g\|_\infty + \varepsilon \|\partial_x^{m-1} \tilde{R}_\Psi\|_{L^2} \|g\|_{L^2} \\
  &\leq \mathcal{O}(1)\,,
  \end{aligned}$$

  similar

  $$\|\varepsilon \partial_t \partial_x^{-1} (\partial_x^m \tilde{R}_\Psi \partial_x^{-1} g)\|_\infty \leq \mathcal{O}(1)\,.$$

  The term $\tilde{R}_\Psi \partial_x^{-1} g$ only occurs in the skew symmetric case.

Regarding (2.101), we split

$$\tilde{R}_\Psi = \psi_c + \varepsilon \tilde{R}_Q$$

such that we have $\varepsilon^2 \|\partial_x^m \tilde{R}_Q \partial_x^{-1} g\|_{H^n} = \mathcal{O}(1)$ for $m \geq 0$ and n as required. For $f = \varepsilon \psi_c \partial_x^{-1} g$, we can get

$$\left\| \mathcal{F}^{-1}\left[|(1+|\cdot|^2)^{p/2}\, \widehat{\partial_x^{n+1} f}(\cdot)| \right] \right\|_\infty = \mathcal{O}(1)\,, \tag{2.122}$$

what is sufficient since (2.101) is only needed to apply lemma 2.2.20. We had estimated this ∞-norm by using Sobolev's embedding theorem to get (2.91) in the proof of lemma 2.2.20. For the shake of completeness, here are the full details on how this last estimate is obtained.
We split $\psi_c = \psi_{-1} + \psi_1$, set $A_{-1} := \overline{A}_1$ where A_1 is as in (2.8) and then we compute

$$\begin{aligned}
&\varepsilon \sup_{x\in\mathbb{R}} \left| \mathcal{F}^{-1}\left[|(1+|\cdot|^2)^{p/2} (\cdot)^{n+1} \widehat{\psi_j \partial_x^{-1} g}(\cdot)| \right] \right| \\
&\quad = \varepsilon \frac{1}{4\pi^2} \sup_{(\varepsilon^{-1}x)\in\mathbb{R}} \Big| \int_{\mathbb{R}} e^{ik\varepsilon x} (1+|k|^2)^{p/2} |k|^{n+1} \times \\
&\qquad\qquad \times \Big| \int_{\mathbb{R}} e^{-i(k-jk_0)y} A_j(\varepsilon(y-c_g t), \varepsilon^2 t) e^{ji\omega_0 t} (\partial_x^{-1} g)(y)\, dy \Big|\, dk \Big| \\
&\quad = \varepsilon \frac{1}{4\pi^2} \sup_{x\in\mathbb{R}} \Big| \int_{\mathbb{R}} e^{ikx} (1+|\varepsilon k + jk_0|^2)^{p/2} |\varepsilon k + jk_0|^{n+1} \times \\
&\qquad\qquad \times \Big| \int_{\mathbb{R}} e^{-iky} A_j(y - \varepsilon c_g t, \varepsilon^2 t) e^{ji\omega_0 t} (\partial_x^{-1} g)(\varepsilon^{-1} y)\, dy \Big|\, dk \Big|\,.
\end{aligned}$$

Here, we exploited the property of the supremum on $\mathbb{R}$ and made the substitutions

$$\varepsilon y \mapsto y \qquad \text{and} \qquad \varepsilon^{-1}(k - jk_0) \mapsto k\,.$$

Using Sobolev's embedding theorem, we now get

$$\begin{aligned}
&\varepsilon \sup_{x\in\mathbb{R}} \left| \mathcal{F}^{-1}\left[|(1+|\cdot|^2)^{p/2} (\cdot)^{n+1} \widehat{\psi_j \partial_x^{-1} g}(\cdot)| \right] \right| \\
&\quad = \varepsilon \frac{1}{2\pi} \Big\| (1+|(\cdot)|^2)(1+|\varepsilon(\cdot) + jk_0|^2)^{p/2} |\varepsilon(\cdot) + jk_0|^{n+1} \times \\
&\qquad\qquad \times \int_{\mathbb{R}} e^{-i(\cdot)y} A_j(y - \varepsilon c_g t, \varepsilon^2 t) e^{ji\omega_0 t} (\partial_x^{-1} g)(\varepsilon^{-1} y)\, dy \Big\|_{L^2} \\
&\quad \leq \varepsilon\, \mathcal{O}\Big(\|(\partial_x^{-1} g)(\varepsilon^{-1} y)\|_{C^1} + \sum_{m=1}^{n+p+2} \varepsilon^{m-1} \|(\partial_x^{-1} g)(\varepsilon^{-1} y)\|_{C^m} \Big) \|A_1\|_{H^{n+p+2}} \\
&\quad = \mathcal{O}(1).
\end{aligned}$$

□

Corollary 2.2.30 now allows us to prove theorem 1.1.1.

Proof of theorem 1.1.1 When $\ell \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$, we can use corollary 2.2.30 together with Gronwall's inequality in order to obtain the $\mathcal{O}(1)$-boundedness of $\tilde{\mathcal{E}}_\ell$ for all $t \in [0, T_0/\varepsilon^2]$ as long as $\varepsilon_0 > 0$ is chosen sufficiently small:
Since we assumed the local existence of solutions to (1.2), there is a some $T(\varepsilon) > 0$ such that the H^ℓ-norms of $R_{-1}(t)$ and $R_1(t)$ can be uniformly bounded as long as $0 \leq t \leq T(\varepsilon)$. Due to corollary 2.2.16, we know that for sufficiently small ε_0, (2.98) is true for $0 \leq t \leq T(\varepsilon)$, i.e.

$$\varepsilon \mathcal{E}_\ell(t) \leq 1 \,,$$

and thus corollary 2.2.30 does indeed hold for $0 \leq t \leq T(\varepsilon)$.
In particular, we have

$$\partial_t \tilde{\mathcal{E}}_\ell(t) \leq \varepsilon^2 \, C \left(\tilde{\mathcal{E}}_\ell(t) + 1 \right)$$

for some $C \geq 0$ and $0 \leq t \leq T(\varepsilon)$.
Gronwall's inequality now yields

$$\tilde{\mathcal{E}}_\ell(t) \leq \big(\tilde{\mathcal{E}}_\ell(0) + \varepsilon^2 \, Ct \big) e^{\varepsilon^2 \, Ct}$$

for $0 \leq t \leq T(\varepsilon)$.
Choosing ε_0 such small that

$$\big(\tilde{\mathcal{E}}_\ell(0) + CT_0 \big) e^{CT_0} \leq \varepsilon_0^{-1} \,,$$

we can now obtain $T(\varepsilon) \geq \varepsilon^{-2} \, T_0$, i.e. in particular

$$\tilde{\mathcal{E}}_\ell(t) \leq \big(\tilde{\mathcal{E}}_\ell(0) + \varepsilon^2 \, Ct \big) e^{\varepsilon^2 \, Ct}$$

for $0 \leq t \leq \varepsilon^{-2} \, T_0$.

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

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

due to corollary 2.2.30.
For $\ell \geq s_A$, we can now, due to estimate (2.10), conclude

$$\begin{aligned}
&\sup_{[0, T_0/\varepsilon^2]} \| u - \varepsilon \psi_{NLS} \|_{H^{s_A}} \\
&\qquad = \sup_{[0, T_0/\varepsilon^2]} \| u_{-1} + u_1 - \varepsilon \psi_{NLS} \|_{H^{s_A}} \\
&\qquad \leq \sup_{[0, T_0/\varepsilon^2]} \left\| \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon \begin{pmatrix} \psi_{NLS} \\ 0 \end{pmatrix} \right\|_{H^{s_A}} \\
&\qquad \leq \sup_{[0, T_0/\varepsilon^2]} \varepsilon^\beta \left\| \begin{pmatrix} \vartheta R_{-1} \\ \vartheta R_1 \end{pmatrix} \right\|_{H^{s_A}} + \sup_{[0, T_0/\varepsilon^2]} \left\| \varepsilon \Psi - \varepsilon \begin{pmatrix} \psi_{NLS} \\ 0 \end{pmatrix} \right\|_{H^{s_A}} \\
&\qquad \lesssim \varepsilon^{3/2}.
\end{aligned}$$

□

Remark 2.2.32. *In the above proof, we clearly see that, due to our method of proof, our estimate for the size of the error*

$$\sup_{[0,T_0/\varepsilon^2]} \|u - \varepsilon\psi_{NLS}\|_{H^{s_A}}$$

cannot be better than the estimate, which we have for the difference between the NLS-approximation $\varepsilon\psi_{NLS}$ and the improved NLS-approximation $\varepsilon\Psi$.
We however showed that the error between a original solution of (1.2) *and the improved NLS-approximation $\varepsilon\Psi$,*

$$\sup_{[0,T_0/\varepsilon^2]} \varepsilon^{\beta} \Big\| \begin{pmatrix} \vartheta R_{-1} \\ \vartheta R_1 \end{pmatrix} \Big\|_{H^{s_A}},$$

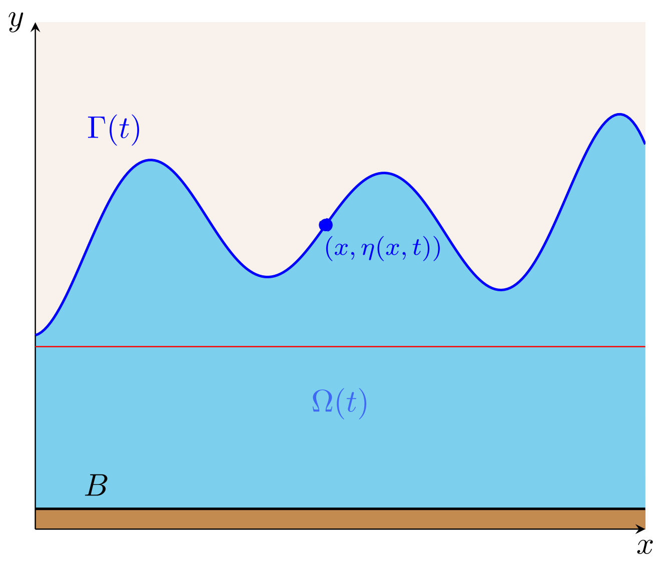
is of the size $\mathcal{O}(\varepsilon^{\beta})$, where $\beta = 5/2$. If the residual estimate (2.9) *can be improved, for instance as in remark 2.1.4, we can make β even larger. In other words, our estimate for the error between a original solution of* (1.2) *and the improved NLS-approximation $\varepsilon\Psi$ is much smaller than our estimate for the error between a original solution of* (1.2) *and the NLS-approximation $\varepsilon\psi_{NLS}$. In some cases it can even be made arbitrarily small. That being said, we cannot increase the time interval $[0, T_0/\varepsilon^2]$ on which the estimate does hold this way.*

Chapter 3

A reduced system for the water wave problem

3.1 Motivation

For our introduction of the water wave problem here, we will follow [D19].
The 2D water wave problem consists in finding the flow of an incompressible inviscid fluid in an infinitely long canal of finite or infinite depth with a free top surface under the influence of gravity and possibly of surface tension. The 2D water wave problem with finite depth (formulated in Eulerian coordinates) has the following form.
In an infinitely long canal of finite depth, an incompressible, inviscid fluid fills a domain $\Omega(t) = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, -h < y < \eta(x, t)\}$ in between the impermeable bottom $B = \{(x, y) : x \in \mathbb{R}, y = -h\}$ and the free top surface $\Gamma(t) = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, y = \eta(x, t)\}$. All, under the influence of gravity and surface tension.



The velocity field $V = (v_1, v_2)$ of the fluid is governed by the incompressible Euler equations

$$\partial_t V + \big(V \cdot \nabla\big) V = \nabla p + g \begin{pmatrix} 0 \\ -1 \end{pmatrix} \qquad \text{in } \Omega(t), \tag{3.1}$$

$$\nabla \cdot V = 0 \qquad \text{in } \Omega(t), \tag{3.2}$$

where p is the pressure and g is the constant of gravity.
Under the assumptions that fluid particles on the top surface stay on the top surface, that the pressure at the top surface is determined by the Laplace-Young jump condition and that the bottom is impermeable, one obtains the boundary conditions

$$\eta_t = V \cdot \begin{pmatrix} -\eta_{x_1} \\ 1 \end{pmatrix} \qquad \text{at}\, \Gamma(t), \tag{3.3}$$

$$p = -bgh^2\kappa \qquad \text{at}\, \Gamma(t), \tag{3.4}$$

$$v_2 = 0 \qquad \text{at}\, B, \tag{3.5}$$

where $b \geq 0$ is the Bond number, which is proportional to the strength of surface tension, and κ is the curvature of $\Gamma(t)$.
One further assumes the flow to be irrotational, such that one can now show that there exists a harmonic velocity potential ϕ with vanishing normal derivative at B and an operator $\mathcal{K} = \mathcal{K}(\eta)$ such that

$$V = \nabla\phi \qquad \text{and} \qquad \phi_y \;=\; \mathcal{K}\phi_x. \tag{3.6}$$

Thus, the system (3.1)-(3.5) can be reduced to

$$\eta_t = V \cdot \begin{pmatrix} -\eta_{x_1} \\ 1 \end{pmatrix} \qquad \text{at}\, \Gamma(t), \tag{3.7}$$

$$(\phi)_t = -\frac{1}{2}\big((\phi_x)^2 + (\mathcal{K}\phi_x)^2\big)_x - g\eta + bgh^2\Big(\frac{\eta_x}{\sqrt{1+\eta_x^2}}\Big)_{xx} \qquad \text{at } \Gamma(t) \tag{3.8}$$

or to

$$\eta_t = \mathcal{K}v_1 - v_1\eta_x \qquad \text{at}\, \Gamma(t), \tag{3.9}$$

$$(v_1)_t = -g\eta_x - \frac{1}{2}\big((v_1)^2 + (\mathcal{K}v_1)^2\big)_x + bgh^2\Big(\frac{\eta_x}{\sqrt{1+\eta_x^2}}\Big)_{xx} \qquad \text{at } \Gamma(t). \tag{3.10}$$

One can assume time and space to be rescaled in such a way that $h = 1$ and $g = 1$.

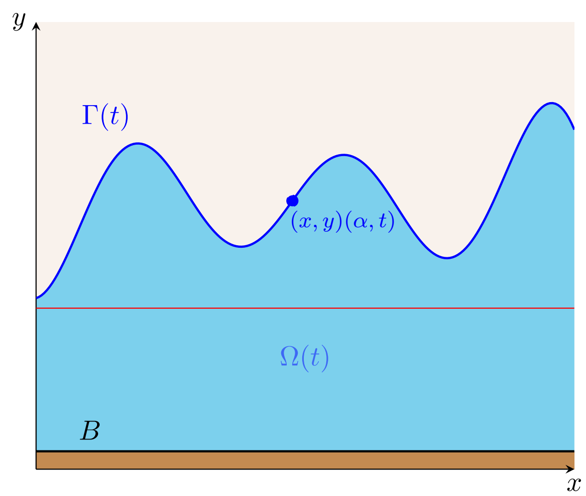
Besides this formulation of the water wave problem in Eulerian coordinates, there exist also other formulations of the water wave problem and each of these formulations has its own advantages and disadvantages. Including the arc length formulation, there are local and global well-posedness results for almost all of these formulations, we refer to [D18] for a quick overview.

Since the water wave problem and its solutions are not expected to be solved or qualitatively understood in near future, approximations for the water wave problem are of great importance. The two most promising ones are the Korteweg-de Vries approximation and the NLS approximation.

In this section, we will now show that our techniques acquired in the last section will be useful to prove the validity of the NLS approximation for the 2D water wave problem in case of finite depth and with or without surface tension, in the arc length formulation. We choose the arc length formulation, since this formulation has the advantage that the linear part of the equation is the one with the most derivatives in the presence of nonzero surface tension. I.e. we have $\deg(\omega) > \deg^*(\rho)$ in case of nonzero surface tension.
Until recently, the validity of the NLS approximation for the 2D water wave problem with finite depth and surface tension was an open problem, regardless of the chosen formulation. The validity of the NLS approximation for the 2D water wave problem with finite depth and without surface tension has already been proved in [DSW16] by using Lagrangian coordinates, however the result was not optimal in the sense that the validity of the approximation could be proven on the right timescale but not for the full modulation interval.

3.1.1 The 2D water wave problem in arc length formulation

In order to obtain the arc length formulation of the water wave problem, the free top surface $\Gamma(t)$ gets parametrized by arc length.



Let $P(t) : \mathbb{R} \to \Gamma(t) : \alpha \mapsto P(\alpha, t) = \big(x(\alpha, t), y(\alpha, t)\big)$ be such a parametrization, i.e. let $P(t)$ be such that

$$\sqrt{(\partial_\alpha x)^2(\alpha,t) + (\partial_\alpha y)^2(\alpha,t)} = 1\,. \tag{3.11}$$

Let

$$\theta = \arctan\left(\frac{\partial_\alpha x}{\partial_\alpha y}\right) \tag{3.12}$$

denote the tangent angles , U the normal velocity and T the tangential velocity on the free top surface, i.e.

$$\partial_t\big(x(\alpha,t), y(\alpha,t)\big) = U(\alpha,t)\hat{n}(\alpha,t) + T(\alpha,t)\hat{t}(\alpha,t)\,, \tag{3.13}$$

where $\hat{n} = \big(-\sin(\theta), \cos(\theta)\big)$ are the upward unit normal vectors to the free surface and $\hat{t} = \big(\cos(\theta), \sin(\theta)\big)$ the upward unit tangential vectors to the free surface.
Due to (3.11), one can show that T is determined by ϑ and U (up to a constant that can be set to zero without a loss of generality). One can show

$$T(\alpha,t) = \int_{-\infty}^{\alpha} \partial_\beta\theta(\beta,t)\,U(\beta,t)\,d\beta\,. \tag{3.14}$$

Since irrotational flows are considered, the normal velocity U can be expressed in terms of the free top surface and the physical tangential velocity v. The physical tangential velocity v can be expressed by using the velocity field $V = (v_1, v_2)$ and the evolution of v is determined by the equations (3.1)-(3.5) and the form of the free top surface.
Under the assumption that $y(\cdot,t)$, $\theta(\cdot,t)$ and $v(\cdot,t)$ are sufficiently regular, for example $y(\cdot,t), v(\cdot,t) \in L^2$ and $\theta(\cdot,t) \in H^2$, the evolution of x is completely determined by the one of θ due to (3.11). Thus $U(\cdot,t)$ can be written as a function of $y(\cdot,t)$, $\theta(\cdot,t)$ and $v(\cdot,t)$ and one can obtain the system

$$\begin{aligned}
\partial_t y &= U\cos(\theta) + T\partial_\alpha y, && (3.15)\\
\partial_t v &= -\partial_\alpha y + b\partial_\alpha^2\theta - \delta\partial_\alpha\delta + U(\partial_\alpha U - T\partial_\alpha\theta), && (3.16)\\
\partial_t\theta &= \partial_\alpha U + T\partial_\alpha\theta && (3.17)\\
\partial_t\partial_\alpha\delta &= -c\partial_\alpha\theta + b\partial_\alpha^3\theta - \partial_\alpha(\delta\partial_\alpha\delta) + (\partial_\alpha U + v\partial_\alpha\theta)^2 && (3.18)\\
\partial_\alpha y &= \sin(\theta) && (3.19)\\
\delta &= v - T\,, && (3.20)
\end{aligned}$$

where

$$c = \partial_t U + v\partial_t\theta + \delta\partial_\alpha U + \delta v\partial_\alpha\theta + \cos(\theta). \tag{3.21}$$

The evolution equations (3.17) and (3.18) get included since they have better regularity properties than the spatial derivatives of y and v.

In [D19], the linear and quadratic terms, which are also the most troublesome terms, are extracted. Then, θ and $\partial_\alpha \delta$ get derived one time in space and the system gets diagonalized by

$$\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & \sigma^{-1} \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \tilde{u}_{-1} \\ \tilde{u}_1 \end{pmatrix}, \tag{3.22}$$

$$\begin{pmatrix} \partial_\alpha \theta \\ \partial_\alpha^2 \delta \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & \sigma^{-1} \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \tilde{u}_{-2} \\ \tilde{u}_2 \end{pmatrix}, \tag{3.23}$$

where σ^{-1} is the inverse of the operator σ, which is defined in Fourier space by its symbol

$$\sigma(k) = \sigma(k; b) = \sqrt{\frac{k + bk^3}{\tanh(k)}}\,. \tag{3.24}$$

By doing so, the following diagonalized system is obtained

$$\begin{aligned} (\tilde{u}_{-1})_t &= -i\omega\tilde{u}_{-1} \\ &\quad +\partial_\alpha\Big(-\frac{1}{4}(\tilde{u}_{-1}+\tilde{u}_1)^2 + \frac{1}{4}(K_0(\tilde{u}_{-1}+\tilde{u}_1))^2 \\ &\qquad +\frac{1}{2}\sigma K_0[K_0, \sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)](\tilde{u}_{-1}+\tilde{u}_1) \\ &\qquad -\frac{1}{2}\sigma(1+K_0^2)(\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)(\tilde{u}_{-1}+\tilde{u}_1))\Big) \\ &\quad +m_{-1}, \end{aligned} \tag{3.25}$$

$$\begin{aligned} (\tilde{u}_1)_t &= i\omega\tilde{u}_1 \\ &\quad +\partial_\alpha\Big(-\frac{1}{4}(\tilde{u}_{-1}+\tilde{u}_1)^2) + \frac{1}{4}(K_0(\tilde{u}_{-1}+\tilde{u}_1))^2 \\ &\qquad -\frac{1}{2}\sigma K_0[K_0, \sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)](\tilde{u}_{-1}+\tilde{u}_1) \\ &\qquad +\frac{1}{2}\sigma(1+K_0^2)(\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)(\tilde{u}_{-1}+\tilde{u}_1))\Big) \\ &\quad +m_1, \end{aligned} \tag{3.26}$$

$$
\begin{aligned}
(\tilde{u}_{-2})_t \;=\; & -i\omega\tilde{u}_{-2} \\
& +\partial_\alpha\Big(-\partial_\alpha^{-2}(\tilde{u}_{-2}+\tilde{u}_2)\tilde{u}_{-2} \\
& \qquad -\frac{1}{2}[\sigma,\partial_\alpha^{-2}(\tilde{u}_{-2}+\tilde{u}_2)]\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad +\frac{1}{2}K_0\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad -\frac{1}{2}b\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)K_0\sigma^{-1}\partial_\alpha(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad -\frac{1}{2}(\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))^2+\frac{1}{2}(K_0\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))^2 \\
& \qquad +\frac{1}{2}\partial_\alpha\big(\sigma K_0[K_0,\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)]\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2)\big) \\
& \qquad -\frac{1}{2}\partial_\alpha\big(\sigma(1+K_0^2)(\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))\big) \\
& \qquad +\frac{1}{2}\sigma K_0[K_0,\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)]\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2) \\
& \qquad -\frac{1}{2}\sigma(1+K_0^2)(\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2)) \\
& \qquad +\frac{1}{2}c_1\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)+m_{-2}\Big),
\end{aligned}
\tag{3.27}
$$

$$
\begin{aligned}
(\tilde{u}_2)_t = \; & i\omega\tilde{u}_2 \qquad (3.28)\\
& +\partial_\alpha\Big(-\partial_\alpha^{-2}(\tilde{u}_{-2}+\tilde{u}_2)\tilde{u}_2 \\
& \qquad +\frac{1}{2}[\sigma,\partial_\alpha^{-2}(\tilde{u}_{-2}+\tilde{u}_2)]\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad +\frac{1}{2}K_0\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad -\frac{1}{2}b\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)K_0\sigma^{-1}\partial_\alpha(\tilde{u}_{-2}-\tilde{u}_2) \\
& \qquad -\frac{1}{2}(\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))^2+\frac{1}{2}(K_0\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))^2 \\
& \qquad -\frac{1}{2}\partial_\alpha\big(\sigma K_0[K_0,\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)]\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2)\big) \\
& \qquad +\frac{1}{2}\partial_\alpha\big(\sigma(1+K_0^2)(\sigma^{-1}(\tilde{u}_{-1}-\tilde{u}_1)\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2))\big) \\
& \qquad -\frac{1}{2}\sigma K_0[K_0,\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)]\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2) \\
& \qquad +\frac{1}{2}\sigma(1+K_0^2)(\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)\partial_\alpha^{-1}(\tilde{u}_{-2}+\tilde{u}_2)) \\
& \qquad +\frac{1}{2}c_1\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)+m_2\Big),
\end{aligned}
$$

with

$$\partial_\alpha^{-1}\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)=\sigma^{-1}\partial_\alpha(\tilde{u}_{-1}-\tilde{u}_1)+m_3, \qquad (3.29)$$

$$\partial_\alpha^{-2}(\tilde{u}_{-2}+\tilde{u}_2)=(\tilde{u}_{-1}+\tilde{u}_1)-\partial_\alpha^{-1}\big(K_0(\tilde{u}_{-1}+\tilde{u}_1)\sigma^{-1}(\tilde{u}_{-2}-\tilde{u}_2)\big)+m_4, \qquad (3.30)$$

where ω is given in Fourier space by its symbol

$$\omega(k)=\omega(k;b)=\mathrm{sign}(k)\sqrt{(k+bk^3)\tanh(k)}\,, \qquad (3.31)$$

K_0 is defined by its symbol $K_0(k)=-i\tanh(k)$ and ∂_α^{-1} by the multiplier $-ik^{-1}$. The relevant norms of the nonlinear terms $m_{-1},m_1,\partial_\alpha m_{-2},\partial_\alpha m_2,\partial_\alpha m_3,m_4,c_1$ can be controlled. For more details on this and on the whole derivation, we refer to [D19] and the references therein.

3.2 The Reduced system and its properties

Preventing a loss of regularity stemming from the evolution of $\tilde{u}_{-2}$ and $\tilde{u}_2$ is crucial for justifying a NLS approximation for the above arc length formulation of the full

2D water wave problem. Instead of the above full water wave problem, which is a rather complicated system, we will now consider a reduced system, which shares crucial properties with the full water wave problem. That is in particular the structure of the linear and quadratic terms of the evolution from $\tilde{u}_{-2}$ and $\tilde{u}_2$. What we will neglect in this reduced system, are all terms, which are neither linear nor quadratic. On top of that, we will only look at a system with two evolution parameters, since the evolution parameters $\tilde{u}_{-2}, \tilde{u}_2$ and $\tilde{u}_{-1}, \tilde{u}_1$ are deeply connected, see (3.29)- (3.30) (also remember that $\partial_\alpha \theta$ and $\partial_\alpha^2 \delta$ were chosen as a substitute for higher spatial derivatives of y and v).

The reduced system we are looking at is

$$\begin{aligned}
\partial_t u_{-1} = -i\omega u_{-1} + \partial_\alpha \Big(& - D_\alpha^{-2}(u_{-1} + u_1)u_{-1} \\
& - \frac{1}{2}[\sigma, D_\alpha^{-2}(u_{-1} + u_1)]\sigma^{-1}(u_{-1} - u_1) \\
& + \frac{1}{2}K_0 D_\alpha^{-1}\sigma^{-1}(u_{-1} - u_1)\sigma^{-1}(u_{-1} - u_1) \\
& - \frac{1}{2}b\sigma^{-1}(u_{-1} - u_1)K_0\sigma^{-1}\partial_\alpha(u_{-1} - u_1) \\
& - \frac{1}{2}(D_\alpha^{-1}(u_{-1} + u_1))^2 + \frac{1}{2}(K_0 D_\alpha^{-1}(u_{-1} + u_1))^2 \Big), \\
\partial_t u_1 = i\omega u_1 + \partial_\alpha \Big(& - D_\alpha^{-2}(u_{-1} + u_1)u_1 \\
& + \frac{1}{2}[\sigma, D_\alpha^{-2}(u_{-1} + u_1)]\sigma^{-1}(u_{-1} - u_1) \\
& + \frac{1}{2}K_0 D_\alpha^{-1}\sigma^{-1}(u_{-1} - u_1)\sigma^{-1}(u_{-1} - u_1) \\
& - \frac{1}{2}b\sigma^{-1}(u_{-1} - u_1)K_0\sigma^{-1}\partial_\alpha(u_{-1} - u_1) \\
& - \frac{1}{2}(D_\alpha^{-1}(u_{-1} + u_1))^2 + \frac{1}{2}(K_0 D_\alpha^{-1}(u_{-1} + u_1))^2 \Big),
\end{aligned} \tag{3.32}$$

where the linear operator $i\omega$ is given in Fourier space by

$$\omega(k) = \omega(k; b) = \text{sign}(k)\sqrt{(k + bk^3)\tanh(k)}\,. \tag{3.33}$$

In order to avoid resonances being caused by the nonlinear terms, we do not use $D_\alpha^{-1} = \partial_\alpha^{-1}$. Instead, the operator $D_\alpha^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is given in Fourier space by some fixed function $\hat{D}_\alpha^{-1}$, which is smooth, odd and fulfills $\hat{D}_\alpha^{-1}(k) = \mathcal{O}(-ik^{-1})$ for $|k| \to \infty$. Like before, the operators σ, K_0 and σ^{-1} are given in Fourier space by

$$\sigma(k) = \sigma(k; b) = \sqrt{\frac{k + bk^3}{\tanh(k)}}\,,$$

$K_0(k) = -i\tanh(k)$ and $\sigma^{-1}(k) = (\sigma(k))^{-1}$.

We obtain this reduced system by modifying the evolution equations of $\tilde{u}_{-2}$ and $\tilde{u}_2$ from the arc length formulation of the full 2D water wave problem in the following way:

- We replace the operators ∂_α^{-1} by the operators D_α^{-1} in order to avoid resonances produced by the nonlinear terms.
  This has to be done since we only want to consider a system with two evolution parameters and therefore automatically miss out the special interaction between $\tilde{u}_{-1}, \tilde{u}_1$ and $\tilde{u}_{-2}, \tilde{u}_2$, which could ensure that expressions like $\partial_\alpha^{-2}(\tilde{u}_{-2} + \tilde{u}_2)$ make sense.

- We drop all terms that are neither linear nor quadratic and on top of that also some quadratic terms, which cannot cause a loss of regularity.

In the following, we assume the local existence of real-valued solutions to our system (1.2) in H^{s_A} with s_A as in theorem 1.2.1.
We chose $k_0 > 0$ such that (1.4), i.e.

$$\omega''(k_0) \neq 0, \tag{3.34}$$

(1.5), i.e.

$$\omega'(k_0) \neq \pm\omega'(0) \tag{3.35}$$

and (1.6), i.e.

$$m\omega(k_0) \neq \pm\omega(mk_0) \qquad \text{for } m = \pm 2, \dots, \pm 5\,, \tag{3.36}$$

(1.14), i.e.

$$\omega(k_0) \neq 0 \qquad \text{when } \deg(\omega) < 1\,, \tag{3.37}$$

are true. Moreover, we chose $k_0 > 0$ such that for $j_1, j_2 \in \{\pm 1\}$ the only possible solutions of the equations (1.12), i.e.

$$\omega(k) - j_1 j_2 \omega(k \mp k_0) + j_1\omega(\pm k_0) = 0\,, \tag{3.38}$$

are $k = \pm k_0$ and $k = 0$.
Solutions of (3.38) will correspond to resonances in our normal form transforms.

Remark 3.2.1. *In literature it was shown that when* $b = 0$ *or* $b > 1/3$ *there can always only occur resonances in* $k = \pm k_0$ *and* $k = 0$ *for all* $k_0 > 0$. *When* $b \in]0, 1/3[$ *there can occur more than three resonances for some* $k_0 > 0$.

Instead of deriving the NLS equation rigorously by proving residual estimates, what could get pretty exhaustive, we will just assume that a NLS equation of the form

$$\partial_T A = i\,\frac{\omega''(k_0)}{2}\,\partial_X^2 A + i\nu_2 A|A|^2\,, \tag{3.39}$$

with $\nu_2 = \nu_2(k_0; b) \in \mathbb{R}$ can be derived via an ansatz of the form

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \begin{pmatrix} \varepsilon\psi_{NLS} \\ 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2), \tag{3.40}$$

where

$$\varepsilon\psi_{NLS}(x,t) = \varepsilon A\big(\varepsilon(x - c_g t), \varepsilon^2 t\big)e^{i(k_0 x - \omega_0 t)} + c.c.\,, \tag{3.41}$$

$\omega_0 = \omega(k_0)$ and $c_g = \omega'(k_0)$.
Moreover, we assume to have an improved approximation

$$\varepsilon\Psi = \varepsilon\Psi_c + \varepsilon^2\Psi_q\,, \tag{3.42}$$

where

$$\begin{aligned} \varepsilon\Psi_c &= \varepsilon\psi_c \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \varepsilon(\psi_1 + \psi_{-1}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \varepsilon\,\big(\,A_1(\varepsilon(x - c_g t), \varepsilon^2 t)\, e^{i(k_0 x - \omega_0 t)} + c.c.\big) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \varepsilon^2\Psi_q &= \varepsilon^2 \begin{pmatrix} \psi_{q_{-1}} \\ \psi_{q_1} \end{pmatrix}. \end{aligned}$$

The functions $\psi_{q_{-1}}$ and $\psi_{q_{-1}}$ shall have a finite support in Fourier space, which is restricted to small neighborhoods of integer multiples from the basic wave numbers k_0. Therefore, $A_1\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)$ denotes the restriction of $A\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)$ in Fourier space to the interval $\{k \in \mathbb{R} : |k| \le \delta < k_0/20\}$ by some cut-off function:

$$A_1\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big) := \mathcal{F}^{-1}\Big[\chi_{[-\delta,\delta]}(\cdot)\mathcal{F}\big[A\big(\varepsilon(\cdot - c_g t), \varepsilon^2 t\big)\big](\cdot)\Big]\,,$$

where $\chi_{[-\delta,\delta]}$ is the characteristic function on the interval $[-\delta, \delta]$. One can again think of $\varepsilon\psi_c$ as $\varepsilon\psi_{NLS}$, just with a support in Fourier space which is restricted to small neighborhoods of the wave numbers $\pm k_0$.
The improved approximation $\varepsilon\Psi$ shall have the following properties.

Lemma 3.2.2. *Let* $s_A \geq 7$ *and* $A \in C([0,T_0], H^{s_A}(\mathbb{R},\mathbb{C}))$ *be a solution of the NLS equation* (3.39) *with*

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

Then for all $s \geq 0$ *there exist* $C_{Res}, C_\Psi, \varepsilon_0 > 0$ *depending on* C_A *such that for all* $\varepsilon \in (0,\varepsilon_0)$ *the approximation* $\varepsilon\Psi = \varepsilon\Psi_c + \varepsilon^2\Psi_q$ *satisfies*

$$\sup_{t\in[0,T_0/\varepsilon^2]} \|\mathrm{Res}_u(\varepsilon\Psi)\|_{H^s} \leq C_{\mathrm{Res}}\, \varepsilon^{11/2}, \tag{3.43}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} \Big\|\varepsilon\Psi - \varepsilon\psi_{NLS}\begin{pmatrix}1\\0\end{pmatrix}\Big\|_{H^{s_A}} \leq C_\Psi\, \varepsilon^{3/2}, \tag{3.44}$$

$$\sup_{t\in[0,T_0/\varepsilon^2]} (\|\widehat{\Psi}_c\|_{L^1(s+1)(\mathbb{R},\mathbb{C})} + \|\widehat{\Psi}_q\|_{L^1(s+1)(\mathbb{R},\mathbb{C})}) \leq C_\Psi\,, \tag{3.45}$$

$$\|\partial_t\widehat{\psi}_{\pm1} + i\widehat{\omega\psi}_{\pm1}\|_{L^1(s)} \leq C_\Psi\, \varepsilon^2\,. \tag{3.46}$$

Similar estimates can be shown for the full water wave problem, see [D19].

3.2.1 Key properties of the reduced system

The following properties of the system will now be sufficient for justifying the NLS equation, i.e. to prove error estimates on the right time scale.
The reduced system is an abstract system of the form

$$\partial_t u_{-1} = -i\omega u_{-1} + \mathcal{A}_{-1}(u_{-1},u_{-1}) + \mathcal{B}_{-1}(u_1,u_1) + \mathcal{C}_{-1}(u_{-1},u_1) \tag{3.47}$$

$$\partial_t u_1 = i\omega u_1 + \mathcal{A}_1(u_{-1},u_{-1}) + \mathcal{B}_1(u_1,u_1) + \mathcal{C}_1(u_{-1},u_1)\,, \tag{3.48}$$

where the linear operator $i\omega$ is given exactly like before by

$$\omega(k) = \omega(k;b) = \mathrm{sign}(k)\sqrt{(k+bk^3)\tanh(k)}\,. \tag{3.49}$$

The quadratic terms are given in Fourier space by

$$\widehat{\mathcal{A}}_j(u_{-1},u_{-1})(k) := \int_{\mathbb{R}} a_j(k,k-m,m)\,\widehat{u}_{-1}(k-m)\widehat{u}_{-1}(m)\,dm\,, \tag{3.50}$$

$$\widehat{\mathcal{B}}_j(u_1,u_1)(k) := \int_{\mathbb{R}} b_j(k,k-m,m)\,\widehat{u}_1(k-m)\widehat{u}_1(m)\,dm\,, \tag{3.51}$$

$$\widehat{\mathcal{C}}_j(u_{-1},u_1)(k) := \int_{\mathbb{R}} c_j(k,k-m,m)\,\widehat{u}_{-1}(k-m)\widehat{u}_1(m)\,dm\,, \tag{3.52}$$

where $j \in \{\pm1\}$ and the functions a_j, b_j and c_j are assumed to be sufficiently smooth.
For $\mathcal{Z} \in \{\mathcal{A}_{-1},\mathcal{A}_1,\mathcal{B}_{-1},\mathcal{B}_1,\mathcal{C}_{-1},\mathcal{C}_1\}$ and accordingly chosen $z \in \{a_{-1},a_1,b_{-1},b_1,c_{-1},c_1\}$, we have

$$z(k,k-m,m) = \mathcal{O}(k) \qquad \text{for } |k|\to 0\,, \tag{3.53}$$

as long as $|k-m|$ gets uniformly bounded.
The $z(k,k-m,m)$ suffice the conditions of lemma 3.3.10.
Moreover, the operators $\mathcal{Z}$ always map a pair of real-valued functions on a real-valued function and satisfy a priori estimates of the form

$$\|\mathcal{Z}(u,v)\|_{L^2} \lesssim \begin{cases} \|u\|_{H^2}\|v\|_{H^1}\,, \\ \|u\|_{H^1}\|v\|_{H^2}\,, \\ \|\widehat{u}\|_{L^1(4)}\|v\|_{H^1}, \\ \|u\|_{H^1}\|\widehat{v}\|_{L^1(4)} \end{cases} . \tag{3.54}$$

On top of that, we have the a priori estimates

$$\begin{aligned} \|\mathcal{A}_{-1,s}(f,g)+\mathcal{A}^*_{-1,s}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\ \|\mathcal{B}_{1,s}(f,g)+\mathcal{B}^*_{1,s}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\ \|\mathcal{C}_{-1}(g,f)+\mathcal{C}_{-1,*}(g,f)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \\ \|\mathcal{C}_{1}(f,g)+\mathcal{C}^*_{1}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{L^2}\,, \end{aligned} \tag{3.55}$$

and

$$\begin{aligned} \|\mathcal{A}^*_{1,s}(f,g)+\mathcal{C}_{-1}(f,g)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{H^{1/2}}\,, \\ \|\mathcal{B}_{-1,s}(f,g)+\mathcal{C}_{1,*}(g,f)\|_{L^2} &\lesssim \min\left\{\|f\|_{H^4},\ \|\widehat{f}\|_{L^1(4)}\right\}\|g\|_{H^{1/2}}\,. \end{aligned} \tag{3.56}$$

Here, we are using the notations

$$\mathcal{Z}_s(f,\cdot) := \mathcal{Z}(f,\cdot)+\mathcal{Z}(\cdot,f), \quad z^s := z(k,k-m,m)+z(k,m,k-m)\,, \tag{3.57}$$

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{Z}^*(g,f)\,h\,dx &:= \int_{\mathbb{R}} f\,\mathcal{Z}(g,h)\,dx\,, \\ \int_{\mathbb{R}} \mathcal{Z}_*(g,f)\,h\,dx &:= \int_{\mathbb{R}} g\,\mathcal{Z}(h,f)\,dx\,. \end{aligned} \tag{3.58}$$

Remark 3.2.3. *Due to lemma 2.2.11, $\mathcal{Z}^*$ and $\mathcal{Z}_*$ exists and we can write*

$$\begin{aligned} \mathcal{Z}^*(g,f) &= \int_{\mathbb{R}} z(-m,k-m,-k)\,\widehat{g}(k-m)\widehat{f}(m)\,dm\,, \\ \mathcal{Z}_*(g,f) &= \int_{\mathbb{R}} z(-m,-k,k-m)\,\widehat{g}(m)\widehat{f}(k-m)\,dm\,. \end{aligned} \tag{3.59}$$

Remark 3.2.4. *There is no special reason for which we picked the integer 4, i.e. the H^4-norm and $L^1(4)$-norm for f. A larger real number can also be allowed, however then the condition for corollary 3.3.14 may has to be modified accordingly.*

From our derivation of the NLS equation in section 2.1, we already expected a condition like (3.53) to be crucial, since in the case $\omega(0) = 0$ we could only rigorously derive or justify the NLS equation when $\rho(0) = 0$. The relevance of this condition is indeed also the reason for which θ and $\partial_\alpha \delta$ were derived one time in space before the water wave problem was diagonalized.
The condition (3.53) as well as the assumed smoothness of a_j, b_j and c_j are not entirely true for $\tilde{u}_{-2}$ and $\tilde{u}_2$ in the full water wave problem (3.25)-(3.30) due to the presence of the operator ∂_α^{-1}. This however should be evened out by the special interaction between $\tilde{u}_{-1}$, $\tilde{u}_1$, $\tilde{u}_{-2}$ and $\tilde{u}_2$. The same holds for (3.54), (3.55) and (3.56).

The structure enforced by (3.55) and (3.56) can also be found in the full water wave problem for $\tilde{u}_{-2}$ and $\tilde{u}_2$ and stems from the dynamics of (3.17)-(3.18). To make this structure a bit more tangible, one can rewrite the system (3.47)-(3.48) as

$$\partial_t u_j = ji\omega u_j + Q_{j-1} u_{-1} + Q_{j1} u_1 ,$$

$$\begin{aligned} Q_{j-1} = Q_{j-1}(u_{-1}, u_1) &:= \frac{1}{2}\big(\mathcal{A}_j(u_{-1}, \cdot) + \mathcal{A}_j(\cdot, u_{-1})\big) + \frac{1}{2}\,\mathcal{C}_j(\cdot, u_1) , \\ Q_{j1} = Q_{j1}(u_{-1}, u_1) &:= \frac{1}{2}\big(\mathcal{B}_j(u_1, \cdot) + \mathcal{B}_j(\cdot, u_1)\big) + \frac{1}{2}\,\mathcal{C}_j(u_{-1}, \cdot) \end{aligned}$$

where $j \in \{\pm 1\}$. and the $Q_{j_1 j_2}$ are linear operators for a fixed argument.
Then, we can write the system (3.47)-(3.48) as

$$\partial_t \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \mathcal{Q} \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} := \begin{pmatrix} -i\omega + Q_{-1-1} & Q_{-11} \\ Q_{1-1} & i\omega + Q_{11} \end{pmatrix} \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} . \tag{3.60}$$

and the conditions (3.55) and (3.56) can now be understood as a condition regarding the imaging behavior of the matrix

$$\mathcal{Q} + \mathcal{Q}^* = \begin{pmatrix} Q_{-1-1} + Q^*_{-1-1} & Q_{-11} + Q^*_{1-1} \\ Q_{1-1} + Q^*_{-11} & Q_{11} + Q^*_{11} \end{pmatrix} ,$$

where $Q^*_{j_1 j_2}$ is given through the L^2-product, i.e.

$$\int_{\mathbb{R}} Q_{j_1 j_2} f \; g \, dx = \int_{\mathbb{R}} f \, Q^*_{j_1 j_2} g \, dx .$$

To give some more details on this, we have

$$\begin{aligned} 2\big(Q_{-1-1} + Q^*_{-1-1}\big) &= \mathcal{A}_{-1,s}(u_{-1}, \cdot) + \mathcal{A}^*_{-1,s}(u_{-1}, \cdot) + \mathcal{C}_{-1}(\cdot, u_1) + \mathcal{C}_{-1,*}(\cdot, u_1) , \\ 2\big(Q_{-11} + Q^*_{1-1}\big) &= \mathcal{A}^*_{1,s}(u_{-1}, \cdot) + \mathcal{C}_{-1}(u_{-1}, \cdot) + \mathcal{B}_{-1,s}(u_1, \cdot) + \mathcal{C}_{1,*}(\cdot, u_1) , \\ 2\big(Q_{1-1} + Q^*_{-11}\big) &= \mathcal{A}_{1,s}(u_{-1}, \cdot) + \mathcal{C}^*_{-1}(u_{-1}, \cdot) + \mathcal{B}^*_{-1,s}(u_1, \cdot) + \mathcal{C}_1(\cdot, u_1) , \\ 2\big(Q_{11} + Q^*_{11}\big) &= \mathcal{B}_{1,s}(u_1, \cdot) + \mathcal{B}^*_{1,s}(u_1, \cdot) + \mathcal{C}_1(u_{-1}, \cdot) + \mathcal{C}^*_1(u_{-1}, \cdot) , \end{aligned}$$

where $\mathcal{A}_{j,s}(u_{-1},\cdot) := \mathcal{A}_j(u_{-1},\cdot)+\mathcal{A}_j(\cdot,u_{-1})$, $\mathcal{B}_{j,s}(u_{-1},\cdot)$ analogously. Thus, (3.55) refers to the diagonal entries of the matrix $\mathcal{Q}+\mathcal{Q}^*$ and (3.56) to the minor diagonal entries of the matrix $\mathcal{Q}+\mathcal{Q}^*$. One could therefore also formulate (3.55) and (3.56) as a property, which the matrix $\mathcal{Q}+\mathcal{Q}^*$ has to fulfill.

As a side note, conditions similar to (3.55) and (3.56) were automatically fulfilled for the system (1.2) from the last section, i.e. the system (2.3). This was in particular, since the function ρ was odd and (1.10) held.

Lemma 3.2.5. *The system* (3.32) *indeed satisfies the conditions* (3.53), (3.54), (3.55) *and* (3.56).

Proof. We write the system (3.32) as

$$\begin{aligned}\partial_t u_{-1} = -i\omega u_{-1} - T_0(u_{-1}+u_1,u_{-1}) - T_1(u_{-1}+u_1,u_{-1}-u_1)&\\ + T_2(u_{-1}-u_1,u_{-1}-u_1) - T_3(u_{-1}-u_1,u_{-1}-u_1)&\\ - T_4(u_{-1}+u_1,u_{-1}+u_1)\,,&\end{aligned}$$

$$\begin{aligned}\partial_t u_1 = i\omega u_1 - T_0(u_{-1}+u_1,u_1) + T_1(u_{-1}+u_1,u_{-1}-u_1)&\\ + T_2(u_{-1}-u_1,u_{-1}-u_1) - T_3(u_{-1}-u_1,u_{-1}-u_1)&\\ - T_4(u_{-1}+u_1,u_{-1}+u_1)\,,&\end{aligned}$$

where

$$\begin{aligned}T_0(f,g) &= \partial_\alpha(D_\alpha^{-2} fg)\,,\\ T_1(f,g) &= \frac{1}{2}\,\partial_\alpha[\sigma, D_\alpha^{-2} f]\sigma^{-1} g\,,\\ T_2(f,g) &= \frac{1}{2}\,\partial_\alpha(K_0 D_\alpha^{-1}\sigma^{-1} f\,\sigma^{-1} g)\,,\\ T_3(f,g) &= \frac{b}{2}\,\partial_\alpha(\sigma^{-1} f\,K_0\sigma^{-1}\partial_\alpha g)\,,\\ T_4(f,g) &= \frac{1}{2}\,\partial_\alpha\big(D_\alpha^{-1} f\,D_\alpha^{-1} g - K_0 D_\alpha^{-1} f\,K_0 D_\alpha^{-1} g\big)\,.\end{aligned}$$

Bringing the system into the form (3.47)-(3.48), we now get

$$\begin{aligned}\mathcal{A}_{-1}(u_{-1},u_{-1}) &= -T_0(u_{-1},u_{-1}) - T_1(u_{-1},u_{-1}) + T_2(u_{-1},u_{-1})\\ &\quad - T_3(u_{-1},u_{-1}) - T_4(u_{-1},u_{-1})\,,\\ \mathcal{B}_{-1}(u_1,u_1) &= T_1(u_1,u_1) + T_2(u_1,u_1)\\ &\quad - T_3(u_1,u_1) - T_4(u_1,u_1)\,,\\ \mathcal{C}_{-1}(u_{-1},u_1) &= -T_0(u_1,u_{-1}) + T_1(u_{-1},u_1) - T_1(u_1,u_{-1}) - T_2(u_{-1},u_1)\\ &\quad - T_2(u_1,u_{-1}) + T_3(u_{-1},u_1) + T_3(u_1,u_{-1}) - 2\,T_4(u_{-1},u_{-1})\,,\end{aligned}$$

$$
\begin{aligned}
\mathcal{A}_1(u_{-1},u_{-1}) &= T_1(u_{-1},u_{-1}) + T_2(u_{-1},u_{-1}) \\
&\qquad - T_3(u_{-1},u_{-1}) - T_4(u_{-1},u_{-1})\,, \\
\mathcal{B}_1(u_1,u_1) &= -T_0(u_1,u_1) - T_1(u_1,u_1) + T_2(u_1,u_1) \\
&\qquad - T_3(u_1,u_1) - T_4(u_1,u_1)\,, \\
\mathcal{C}_1(u_{-1},u_1) &= -T_0(u_{-1},u_1) - T_1(u_{-1},u_1) + T_1(u_1,u_{-1}) - T_2(u_{-1},u_1) \\
&\qquad - T_2(u_1,u_{-1}) + T_3(u_{-1},u_1) + T_3(u_1,u_{-1}) - 2\,T_4(u_{-1},u_{-1})\,.
\end{aligned}
$$

We immediately notice the following interesting structure

$$
\begin{aligned}
\mathcal{A}_{-1}(f,g) &= \mathcal{B}_1(f,g) \\
\mathcal{B}_{-1}(f,g) &= \mathcal{A}_1(f,g) \\
\mathcal{C}_{-1}(f,g) &= \mathcal{C}_1(g,f)\,.
\end{aligned}
\tag{3.61}
$$

This special structure is also present in the quadratic terms of (3.29)-(3.30), i.e. in the arc length formulation of the full water wave problem. We decided to not include this property into our key properties since we do not have to exploit it in order to prove theorem (1.2.1). Nevertheless it could be an important feature of of the water wave problem.

In order to prove the lemma, we will now take advantage of the operators $T_0 - T_5$.

The property (3.53) is obviously true since all nonlinear terms of the system are basically a derivative.

Looking at the linear operators

$$
\widehat{T_j}(f,g) := \int_{\mathbb{R}} t_j(k,k-m,m)\,\widehat{f}(k-m)\,\widehat{g}(m)\,dm\,,
$$

we see that (3.54) is true for every T_j and thus also for $\mathcal{A}_{-1}, \mathcal{B}_{-1}, \mathcal{C}_{-1}, \mathcal{A}_1, \mathcal{B}_1$ and $\mathcal{C}_1$. This can for example be shown by finding $s, r \in \mathbb{R}$ such that

$$
\sup_{k,m\in\mathbb{R}} \frac{|t_j(k,k-m,m)|}{(1+|m|^2)^{s/2}(1+|k-m|^2)^{r/2}} \leq C
$$

and then exploiting Plancherel together with Cauchy-Schwarz and Young's inequality.
One can in fact show that even better estimates than (3.54) hold. So we can obviously estimate T_4 without a loss of regularity. We can estimate the L^2-norm of $T_0(f,g)$ only against $\|g\|_{H^1}$, i.e. the non-linearity will always lose one derivative. For $b = 0$ the L^2-norm of $T_2(f,g)$ can only be estimated against $\|g\|_{H^{1/2}}$. While for $b > 0$ the L^2-norm of $T_3(f,g)$ can only be estimated against $\|g\|_{H^{1/2}}$. Otherwise

T_2 or T_3 are harmless. The operator T_1 is harmless for $b = 0$, otherwise it causes a loss of half a derivative.

In order to prove (3.55), we prove

$$\begin{aligned} \|T_j(f,g) + T_j^*(f,g)\|_{L^2} &\le \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\} \|g\|_{L^2}\,, \\ \|T_j(g,f) + T_{j*}(g,f)\|_{L^2} &\le \min\left\{\|f\|_{H^4},\, \|\widehat{f}\|_{L^1(4)}\right\} \|g\|_{L^2}\,, \end{aligned} \tag{3.62}$$

such that (3.55) directly can be followed.

Due to lemma 2.2.11, we have

$$\widehat{T_j}(f,g) + \widehat{T_j^*}(f,g) = \int_{\mathbb{R}} \big(t_j(k,k-m,m) + t(-m,k-m,-k)\big)\, \widehat{f}(k-m)\,\widehat{g}(m)\,dm,$$

$$\widehat{T_j}(g,f) + \widehat{T_{j*}}(g,f) = \int_{\mathbb{R}} \big(t_j(k,m,k-m) + t(-m,-k,k-m)\big)\, \widehat{f}(k-m)\,\widehat{g}(m)\,dm.$$

We have $t_0(k,k-m,m) = ik\hat{D}_\alpha^{-2}(k-m)$ such that

$$\begin{aligned} t_0(k,k-m,m) + t_0(-m,k-m,-k) &= ik\hat{D}_\alpha^{-2}(k-m) - im\hat{D}_\alpha^{-2}(k-m) \\ &= i(k-m)\hat{D}_\alpha^{-2}(k-m) \end{aligned}$$

and

$$\begin{aligned} t_0(k,m,k-m) &= ik\hat{D}_\alpha^{-2}(m) \\ &= i(k-m)\hat{D}_\alpha^{-2}(m) + im\hat{D}_\alpha^{-2}(m)\,, \\ t_0(-m,-k,k-m) &= i(k-m)\hat{D}_\alpha^{-2}(k) - ik\hat{D}_\alpha^{-2}(k)\,. \end{aligned}$$

Therefore we see that (3.62) is true for $T_j = T_0$, e.g. by exploiting Plancherel together with Cauchy-Schwarz and Young's inequality like before.
Exactly like this, we can prove (3.62) for the other operators T_j. There are always either similar cancellations happening as in the first equation, or there are no derivatives falling on g to begin with as in the other two equations above.

Concerning (3.56), we observe that neither $\mathcal{A}_{1,s}^*(f,g)$, $\mathcal{C}_{-1}(f,g)$, $B_{-1,s}(f,g)$ nor $\mathcal{C}_{1,*}(g,f)$ do include the dangerous term

$$T_0(f,g) = \partial_\alpha(D_\alpha^{-2}fg) \qquad \text{or} \quad T_{0*}(f,g) = -D_\alpha^{-2}f\partial_\alpha g\,.$$

As we have seen above the terms $T_0(g,f)$ and $T_{0*}(g,f)$ are harmless and the operators T_j can for $j \in \{1,2,3,4\}$ not cause us to lose more than half a derivative. Thus (3.56) is easily proven.

□

3.3 Error estimates for the reduced model

In this section, we ultimately will prove theorem 1.2.1. Therefore, our aim is to prove that the H^{s_A}-norm of the error

$$\mathcal{R}_{err} = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon \begin{pmatrix} \psi_{NLS} \\ 0 \end{pmatrix}$$

remains bounded by some $\mathcal{O}(\varepsilon^{3/2})$-term on the $\mathcal{O}(\varepsilon^{-2})$-time interval $[0, T_0/\varepsilon^2]$.
In order to achieve this, we will, just as in section 2.2, first estimate the error

$$\mathcal{R} = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi \tag{3.63}$$

that the improved approximation $\varepsilon\Psi$ makes.

Let

$$\Psi = \begin{pmatrix} \Psi_{-1} \\ \Psi_1 \end{pmatrix}$$

be the improved approximation. We write the error (3.63) as

$$\varepsilon^{\beta} \begin{pmatrix} \vartheta R_{-1} \\ \vartheta R_1 \end{pmatrix} = \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi \tag{3.64}$$

where $\beta = 5/2$ and ϑ is an invertible operator on $L^2(\mathbb{R})$ that later will be given by some weight function $\hat{\vartheta}$ in Fourier space.
Throughout this section, we will now work with the rescaled error

$$\begin{pmatrix} R_{-1} \\ R_1 \end{pmatrix} = \varepsilon^{-\beta}\vartheta^{-1}\Big(\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} - \varepsilon\Psi \Big), \tag{3.65}$$

where $\vartheta^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the inverse of the operator ϑ.
Plugging in (3.64) into our original system, we obtain the following dynamics for the rescaled error

$$\partial_t R_{-1} = -i\omega R_{-1} + \varepsilon\vartheta^{-1}\mathcal{G}_{-1}(R^{\Psi}_{-1}, R^{\Psi}_1, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{-1}}(\varepsilon\Psi) \tag{3.66}$$

$$\partial_t R_1 = i\omega R_1 + \varepsilon\vartheta^{-1}\mathcal{G}_1(R^{\Psi}_{-1}, R^{\Psi}_1, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_1}(\varepsilon\Psi)\,, \tag{3.67}$$

where

$$\begin{aligned}\mathcal{G}_j(R^{\Psi}_{-1}, R^{\Psi}_1, \vartheta R_{-1}, \vartheta R_1) :=& \mathcal{A}_j(R^{\Psi}_{-1}, \vartheta R_{-1}) + \mathcal{A}_j(\vartheta R_{-1}, R^{\Psi}_{-1}) + \mathcal{B}_j(R^{\Psi}_1, \vartheta R_1) \\ &+ \mathcal{B}_j(\vartheta R_1, R^{\Psi}_1) + \mathcal{C}_j(R^{\Psi}_{-1}, \vartheta R_1) + \mathcal{C}_j(\vartheta R_{-1}, R^{\Psi}_1)\,,\end{aligned} \tag{3.68}$$

and

$$R^{\Psi}_j := \Psi_j + \frac{1}{2}\varepsilon^{\beta-1}\vartheta R_j\,. \tag{3.69}$$

Remark 3.3.1. *The mappings $\mathcal{G}_j$ are bilinear, when they are viewed as mappings $\left(\mathbb{R}^{\mathbb{R}}\right)^2 \times \left(\mathbb{R}^{\mathbb{R}}\right)^2 \to \mathbb{R}^{\mathbb{R}}$.*

3.3.1 Achieving $\mathcal{O}(\varepsilon^{-2})$-time scale while preserving regularity

We have chosen $\beta = 5/2$ large enough and assume $\mathrm{Res}_{u_j}(\varepsilon\Psi)$ to be small enough such that we formally have

$$\begin{aligned}\partial_t R_j &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\mathcal{G}_j(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi)\\ &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\mathcal{G}_j(\Psi_{-1}, \Psi_1, \vartheta R_{-1}, \vartheta R_1)\\ &\quad + \frac{1}{2}\varepsilon^{\beta}\vartheta^{-1}\mathcal{G}_j(\vartheta R_{-1}, \vartheta R_1, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi)\\ &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\mathcal{G}_j(\Psi_{-1}, \Psi_1, \vartheta R_{-1}, \vartheta R_1) + \mathcal{O}(\varepsilon^2)\,.\end{aligned}$$

We assume that ϑ and the combination $\vartheta^{-1}\mathcal{G}_j$ cannot cause a loss of ε-powers. Thus, by exploiting

$$\Psi = \begin{pmatrix} \Psi_{-1} \\ \Psi_1 \end{pmatrix} = \begin{pmatrix} \psi_c + \varepsilon\psi_{q_{-1}} \\ \varepsilon\psi_{q_1} \end{pmatrix}, \tag{3.70}$$

we obtain

$$\begin{aligned}\partial_t R_j &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\mathcal{G}_j(\psi_c, 0, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^2\vartheta^{-1}\mathcal{G}_j(\psi_{q_{-1}}, \psi_{q_1}, \vartheta R_{-1}, \vartheta R_1) + \mathcal{O}(\varepsilon^2)\\ &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\mathcal{G}_j(\psi_c, 0, \vartheta R_{-1}, \vartheta R_1) + \mathcal{O}(\varepsilon^2)\\ &= j\, i\omega R_j + \varepsilon\vartheta^{-1}\Big(\mathcal{A}_j(\psi_c, \vartheta R_{-1}) + \mathcal{A}_j(\vartheta R_{-1}, \psi_c) + \mathcal{C}_j(\psi_c, \vartheta R_1)\Big) + \mathcal{O}(\varepsilon^2)\,.\end{aligned}$$

Our system has only resonances in $k = \pm k_0$ and $k = 0$, i.e. equation (1.12) is only solved by $k = \pm k_0$ and $k = 0$. Thus, we define ϑ exactly as in (2.30). Thanks to our experience gained from section 2.2.1, we now expect to formally obtain a $\mathcal{O}(\varepsilon^{-2})$-time scale for the error via the normal form transformations

$$\widehat{N}_{j_1 j_2}(\psi_c, R_{j_2})(k) = \int_{\mathbb{R}} \widehat{n}_{j_1 j_2}(k, k-m, m)\widehat{\psi}_c(k-m)\widehat{R}_{j_2}(m)\, dm\,, \tag{3.71}$$

and

$$\widehat{\mathcal{T}}_{j_1 j_2 j_3 j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3})(k) = \int_{\mathbb{R}} t_{j_1, j_2, j_3, j_4}(k)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_{j_4}(m-n)\widehat{R}_{j_3}(n)\, dn\, dm\,, \tag{3.72}$$

where

$$\widehat{n}_{j_1 j_2}(k, k-m, m) = \frac{\rho_{j_1, j_2}(k, k-m, m)\, \widehat{\vartheta}_{\varepsilon,\infty}(m)\, \chi_c(k-m)}{\omega(k) - j_1 j_2\omega(m) + j_1\omega(k-m)}\,,$$

$$t_{j_1, j_2, j_3, j_4}(k) = \frac{-j_2\, \hat{P}_{0,\delta}(k)\, n_{j_1 j_2}(k, j_4 k_0, k - j_4 k_0)\, \rho_{j_2, j_3}(k - j_4 k_0, j_4 k_0, k - 2 j_4 k_0)}{\big(-j_1\omega(k) - 2\omega(j_4 k_0) + j_3\omega(k - 2 j_4 k_0)\big)}\,,$$

$$
\begin{aligned}
\rho_{-1,-1}(k,k-m,m) &:= i\left(a_{-1}(k,k-m,m)+a_{-1}(k,m,k-m)\right),\\
\rho_{-1,1}(k,k-m,m) &:= i\,c_{-1}(k,k-m,m),\\
\rho_{1,-1}(k,k-m,m) &:= -i\left(a_{1}(k,k-m,m)+a_{1}(k,m,k-m)\right),\\
\rho_{1,1}(k,k-m,m) &:= -i\,c_{1}(k,k-m,m)).
\end{aligned}
$$

Just as in the last section, χ_c is the characteristic function on supp $\widehat{\psi}_c$, the function $\widehat{\vartheta}_{\varepsilon,\infty}$ is given as in (2.31) and $\widehat{P}_{a,b}$ denotes the characteristic function of the set $\{k : a \leq |k| \leq b\}$.

Considering (2.2.6), we expect the normal form transformations N_{j_1,j_2} to lose regularity. For this reason, we use the modified energy from section 2.2.2 to preserve regularity. We define

$$
\mathcal{E}_\ell = E_0 + E_\ell\,, \tag{3.73}
$$

$$
E_\ell = \sum_{j_1\in\{\pm1\}}\Big(\frac{1}{2}\left\|\partial_x^\ell R_{j_1}\right\|_{L^2}^2 + \varepsilon\sum_{j_2\in\{\pm1\}}\int_{\mathbb{R}}\partial_x^\ell R_{j_1}\partial_x^\ell\vartheta^{-1}N_{j_1j_2}(\psi_c,R_{j_2})\,dx\Big)\,,
$$

$$
E_0(R) = \|\check{R}_{-1}\|_{L^2}^2 + \|\check{R}_1\|_{L^2}^2\,,
$$

$$
\check{R}_j := R_j + \varepsilon\sum_{j_2\in\{\pm1\}}\vartheta^{-1}N_{jj_2}(\psi_c,R_{j_2}) + \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}}\vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3})\,.
$$

Herby $\vartheta^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the inverse of the operator ϑ, ϑ is just as in (2.30) and ψ_{j_4} is defined similar as in (2.8).

We now expect that this energy is equivalent to the energy $\|R_{-1}\|_{H^\ell}^2 + \|R_1\|_{H^\ell}^2$ for $\ell \geq 1$, just as in the one in section 2.2.1. Moreover, we expect that the evolution of this energy only contains terms that are of at least quadratic ε-order.
In the following, we will now show that these expectations are indeed met.

We start by looking at the normal form transformations.

Lemma 3.3.2. *The normal-form transforms $N_{j_1j_2}$ were constructed such that*

$$
\sum_{j_2\in\{-1,1\}}\Big(-j_1 i\omega N_{j_1j_2}(\psi_c,R_{j_2}) - N_{j_1j_2}(i\omega\psi_c,R_{j_2}) + j_2N_{j_1j_2}(\psi_c,i\omega R_{j_2})\Big) \tag{3.74}
$$

$$
= -\mathcal{G}_{j_1}(\psi_c,0,\vartheta_{\varepsilon,\infty}R_{-1},\vartheta_{\varepsilon,\infty}R_1)
$$

where

$$
\begin{aligned}
\mathcal{G}_{j_1}(\Psi_{-1},\Psi_1,\vartheta R_{-1},\vartheta R_1) - \mathcal{G}_{j_1}(\psi_c,0,\vartheta_{\varepsilon,\infty}R_{-1},\vartheta_{\varepsilon,\infty}R_1)& \\
\leq \varepsilon\,\mathcal{O}(\|R_{-1}\|_{H^2} + \|R_1\|_{H^2}).&
\end{aligned} \tag{3.75}
$$

Moreover, for every fix $h \in L^2(\mathbb{R},\mathbb{R})$ the operators $N_{j_1j_2}(h,\cdot)$ are continuous linear operators which map $H^1(\mathbb{R},\mathbb{R})$ into $L^2(\mathbb{R},\mathbb{R})$. In particular, there is a $C = C(\|\widehat{h}(\cdot)\chi_c(\cdot)\|_{L^1})$ such that for all $g \in H^1(\mathbb{R})$ we have

$$\|N_{jj}(h,g)\|_{L^2} \leq C\|g\|_{H^1}\,, \tag{3.76}$$

$$\|N_{j-j}(h,g)\|_{L^2} \leq C\|g\|_{H^{1/2}}. \tag{3.77}$$

Proof.

In order to find possible resonances for $N_{j_1j_2}$, we have to look at the zeros of the denominator of $n_{j_1j_2}$, i.e. of

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m),$$

for $\chi_c(k-m) \neq 0$, i.e. for $|k-m \mp k_0| \leq \delta$.
Due to (3.38), we can chose δ such small that for $|k-m \mp k_0| \leq \delta$ the equation

$$\omega(k) - j_1j_2\omega(m) + j_1\omega(k-m) = 0, \tag{3.78}$$

can have no other solutions than $k=0$ or $m=0$.
Since we also have

$$\rho_{j_1j_2}(k,k-m,m)\,\chi_c(k) = \mathcal{O}(k) \qquad \text{for } |k| \to 0\,,$$

due to (3.53), we can now proceed exactly as the proof of lemma 2.2.5 in order to show that the normal-form transform $N_{j_1j_2}$ has no nontrivial resonances.

The property (3.74) can be easily checked in Fourier space.

The estimate (3.75) follows by exploiting 3.3.1, i.e.

$$\begin{aligned}&\mathcal{G}_{j_1}(\Psi_{-1},\Psi_1,\vartheta R_{-1},\vartheta R_1) - \mathcal{G}_{j_1}(\psi_c,0,\vartheta_{\varepsilon,\infty}R_{-1},\vartheta_{\varepsilon,\infty}R_1)\\ &\quad= \varepsilon\,\mathcal{G}_{j_1}(\psi_{q_{-1}},\psi_{q_1},\vartheta R_{-1},\vartheta R_1) + \mathcal{G}_{j_1}(\psi_c,0,(\vartheta-\vartheta_{\varepsilon,\infty})R_{-1},(\vartheta-\vartheta_{\varepsilon,\infty})R_1)\,,\end{aligned}$$

together with $\hat{\vartheta}(k) - \hat{\vartheta}_{\varepsilon,\infty}(k) = \mathcal{O}(\varepsilon)$ and estimating $\mathcal{G}_{j_1}$ with (3.54).

The estimates (3.76) and (3.77) can be shown by using (1.11), i.e. the expansions (2.52) and (2.54), together with (3.54).

The bilinear operators A_j and C_j map pairs of real-valued functions on real-valued functions and

$$\begin{aligned}\frac{-j_1 i\,\widehat{\vartheta}_{\varepsilon,\infty}(m)\,\chi_c(k-m)}{\omega(k)-j_1j_2\omega(m)+j_1\omega(k-m)} &= -\overline{\Big(\frac{-j_1 i\,\widehat{\vartheta}_{\varepsilon,\infty}(m)\,\chi_c(k-m)}{\omega(k)-j_1j_2\omega(m)+j_1\omega(k-m)}\Big)}\\ &= -\Big(\frac{-j_1 i\,\widehat{P}_{\varepsilon,\infty}(-m)\widehat{\vartheta}(-m)\,\chi_c(-k+m)}{\omega(-k)-j_1j_2\omega(-m)+j_1\omega(-k+m)}\Big).\end{aligned}$$

Thus,

$$n_{j_1j_2}(k,k-m,m) = \overline{n_{j_1j_2}(k,k-m,m)} = n_{j_1j_2}(-k,-(k-m),-m)$$

and the mappings $f \mapsto N_{j_1j_2}(h,f)$ map real-valued functions on real-valued functions. $\square$

Lemma 3.3.3. *The normal-form transforms $\mathcal{T}_{j_1j_2j_3j_4}$ were constructed such that for all $j_1, j_2, j_3, j_4 \in \{\pm 1\}$, we have*

$$\|\vartheta^{-1} Y_{j_1,j_2}\|_{L^2} \le \mathcal{O}\big(\|R_{-1}\|_{H^2} + \|R_1\|_{H^2}\big). \tag{3.79}$$

where

$$\begin{aligned} Y_{j_1,j_2} = &N_{j_1j_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(\Psi_{-1}, \Psi_1, \vartheta R_{-1}, \vartheta R_1)\big) \\ &+ \sum_{j_3,j_4=\pm 1} \Big(- j_1\, i\omega \mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1j_2j_3j_4}(-i\omega\psi_{j_4}, \psi_{j_4}, R_{j_3}) \\ &\qquad + \mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, -i\omega\psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1j_2j_3j_4}(\psi_{j_4}, \psi_{j_4}, j_3\, i\omega R_{j_3})\Big)\,. \end{aligned} \tag{3.80}$$

Furthermore, for every fix functions g, h with $\widehat{g}, \widehat{h} \in L^1(\mathbb{R},\mathbb{C})$, the mapping $f \mapsto \mathcal{T}_{jj_3}(g, h, f)$ defines a continuous linear map from $L^2(\mathbb{R},\mathbb{C})$ into $L^2(\mathbb{R},\mathbb{C})$ and there exists a constant $C = C\big(\|\widehat{g}\|_{L^1}\|\widehat{h}\|_{L^1}\big)$ such that for all $f \in L^2(\mathbb{R},\mathbb{C})$, we have

$$\|\mathcal{T}_{j_1j_2j_3j_4}(g, h, f)\|_{L^2} \le C\|f\|_{L^2}\,. \tag{3.81}$$

Proof. The proof is analogous to the one of 2.2.8.
That $\mathcal{T}_{j_1j_2j_3j_4}$ is well-defined and estimate (3.81) does hold, can be shown exactly as in the proof of 2.2.8.
Considering estimate (3.79), one shows

$$\begin{aligned} &\vartheta^{-1} N_{j_1j_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(\Psi_{-1}, \Psi_1, \vartheta R_{-1}, \vartheta R_1)\big) \\ &\quad = \sum_{j_4=\pm 1} \Big(P_{0,\delta}\vartheta^{-1} N_{j_1j_2}\big(\psi_{j_4}, \vartheta^{-1}\mathcal{G}_{j_2}(\psi_{j_4}, 0, \vartheta R_{-1}, \vartheta R_1)\big) \\ &\qquad\qquad + P_{0,\delta}\vartheta^{-1} N_{j_1j_2}\big(\psi_{j_4}, \vartheta^{-1}\mathcal{G}_{j_2}(\psi_{-j_4}, 0, \vartheta R_{-1}, \vartheta R_1)\big)\Big) \\ &\qquad + \varepsilon P_{0,\delta}\vartheta^{-1} N_{j_1j_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(\psi_{q_{-1}}, \psi_{q_1}, \vartheta R_{-1}, \vartheta R_1)\big) \\ &\qquad + P_{\delta,\infty} N_{j_1j_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(\Psi_{-1}, \Psi_1, \vartheta R_{-1}, \vartheta R_1)\big) \end{aligned}$$

by exploiting the fact that $\vartheta^{-1} = P_{0,\delta}\vartheta^{-1} + P_{\delta,\infty}$, $\Psi = \big(\psi_c + \varepsilon\psi_{q_{-1}}, \varepsilon\psi_{q_1}\big)^T$ and $\psi_c = \psi_1 + \psi_{-1}$.
Using (2.32), (3.76) and (3.77), (2.37) and (3.54), we see that the L^2-norm of the last two summands can be estimated against $\mathcal{O}(\|R_{-1}\|_{H^2} + \|R_1\|_{H^2})$.
For the remaining summands, we have in Fourier space

$$\begin{aligned} &\mathcal{F}\Big[P_{0,\delta}\vartheta^{-1} N_{j_1j_2}\big(\psi_{j_4}, \vartheta^{-1}\mathcal{G}_{j_2}(\psi_\ell, 0, \vartheta R_{-1}, \vartheta R_1)\big)\Big](k) \\ &\quad = \sum_{j_3=\pm 1} \hat{P}_{0,\delta}(k) \int_{\mathbb{R}}\int_{\mathbb{R}} K_{j_3,\varepsilon}(k, k-m, m, n)\widehat{\psi}_{j_4}(k-m)\widehat{\psi}_\ell(m-n)\widehat{R_{j_3}}(n)\, dndm \end{aligned}$$

where

$$K_{j_3,\varepsilon}(k,k-m,m,n) = \frac{\rho_{j_1,j_2}(k,k-m,m)\left(j_2 i\,\rho_{j_2,j_3}(m,m-n,n)\right)\hat{\vartheta}(n)}{\hat{\vartheta}(k)\left(\omega(k)-j_1j_2\omega(m)+j_1\omega(k-m)\right)}\,.$$

We could replace the term $\hat{\vartheta}_{\varepsilon,\infty}(m)\hat{\vartheta}^{-1}(m)$ by 1, since $|k|\leq\delta$ and $|k-m-j_4k_0|\leq\delta$ implies $|m|>k_0/2>\varepsilon$.
Since the required Lipschitz continuity is given for $K_{j_3,\varepsilon}$, we can now proceed as in the proof of 2.2.8.

□

Lemma 3.3.4. *For E_0 as in* (3.73) *and $m\geq 2$, we have*

$$\partial_t E_0 \leq \varepsilon^2\,\mathcal{O}\Big(\varepsilon^{1/2}\left(\|R_{-1}\|^2_{H^m}+\|R_1\|^2_{H^m}\right)^{3/2}+\|R_{-1}\|^2_{H^m}+\|R_1\|^2_{H^m}+1\Big)\,, \quad (3.82)$$

Proof. The proof is analogous to the one of lemma 2.2.9.
Exploiting the skew symmetry of $i\omega$ and then (2.32), (3.76), (3.77) and (3.81), we obtain

$$\begin{aligned}
\partial_t E_0 &= \sum_{j=\pm1}\int_{\mathbb{R}} \overline{\check{R}_j}\,\partial_t\check{R}_j + \check{R}_j\,\partial_t\overline{\check{R}_j}\,dx \\
&= \sum_{j=\pm1}\int_{\mathbb{R}} \overline{\check{R}_j}\,\partial_t\check{R}_j + \check{R}_j\,\partial_t\overline{\check{R}_j} - \overline{\check{R}_j}\,ji\omega\check{R}_j - \check{R}_j\,ji\omega\overline{\check{R}_j}\,dx \\
&\leq 2\sum_{j=\pm1}\|\check{R}_j\|_{L^2}\|\partial_t\check{R}_j - ji\omega\check{R}_j\|_{L^2} \\
&\leq \mathcal{O}\Big(\left(\|R_{-1}\|^2_{H^1}+\|R_1\|^2_{H^1}\right)^{1/2}\Big)\sum_{j=\pm1}\|\partial_t\check{R}_j - ji\omega\check{R}_j\|_{L^2},
\end{aligned}$$

where

$$\partial_t\check{R}_j = \partial_t R_j + \varepsilon\sum_{j_2\in\{\pm1\}}\vartheta^{-1}\partial_t N_{jj_2}(\psi_c,R_{j_2}) + \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}}\vartheta^{-1}\partial_t\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3}).$$

Due to

$$\begin{aligned}
j\,i\omega R_j = j\,i\omega\check{R}_j &- \varepsilon\sum_{j_2\in\{\pm1\}} j\,i\omega\vartheta^{-1}N_{jj_2}(\psi_c,R_{j_2}) \\
&- \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}} j\,i\omega\vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3})\,,
\end{aligned}$$

we get

$$
\begin{aligned}
\partial_t \check{R}_j =& j\, i\omega \check{R}_j \\
&+ \varepsilon\vartheta^{-1}\Big(\mathcal{G}_j(\Psi_{-1},\Psi_1,\vartheta R_{-1},\vartheta R_1)\big) \\
&\qquad + \sum_{j_2\in\{\pm1\}} \big(-j\, i\omega N_{jj_2}(\psi_c, R_{j_2}) + N_{jj_2}(\partial_t\psi_c, R_{j_2}) + N_{jj_2}(\psi_c, j_2\, i\omega R_{j_2})\big)\Big) \\
&+ \varepsilon^2\vartheta^{-1}\Big(\sum_{j_2\in\{\pm1\}} N_{jj_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(\Psi_{-1},\Psi_1,\vartheta R_{-1},\vartheta R_1)\big) \\
&\qquad + \sum_{j_2,j_3,j_4\in\{\pm1\}} \big(-j\, i\omega \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},R_{j_3}) + \mathcal{T}_{jj_2j_3j_4}(\partial_t\psi_{j_4},\psi_{j_4},R_{j_3}) \\
&\qquad\qquad + \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\partial_t\psi_{j_4},R_{j_3}) + \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4},\psi_{j_4},j_3\, i\omega R_{j_3})\big)\Big) \\
&+ \frac{\varepsilon^\beta}{2}\vartheta^{-1}\mathcal{G}_j(\vartheta R_{-1},\vartheta R_1,\vartheta R_{-1},\vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_j}(\varepsilon\Psi) \\
&+ \varepsilon\sum_{j_2\in\{\pm1\}} \vartheta^{-1}N_{jj_2}\big(\psi_c, \frac{\varepsilon^\beta}{2}\vartheta^{-1}\mathcal{G}_{j_2}(\vartheta R_{-1},\vartheta R_1,\vartheta R_{-1},\vartheta R_1) + \varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_2}}(\varepsilon\Psi)\big) \\
&+ \varepsilon^3\sum_{j_2,j_3,j_4\in\{\pm1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4},\psi_{j_4},\vartheta^{-1}\mathcal{G}_{j_3}(R^\Psi_{-1},R^\Psi_1,\vartheta R_{-1},\vartheta R_1)\big) \\
&+ \varepsilon^2\sum_{j_2,j_3,j_4\in\{\pm1\}} \vartheta^{-1}\mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4},\psi_{j_4},\varepsilon^{-\beta}\vartheta^{-1}\mathrm{Res}_{u_{j_3}}(\varepsilon\Psi)\big).
\end{aligned}
$$

By construction of our normal-form transforms, i.e. due to (3.74) and (3.75), and, (3.79) and (3.80), we obtain

$$
\begin{aligned}
\partial_t \check{R}_j =& j\, i\omega \check{R}_j \\
&+ \varepsilon \vartheta^{-1} \sum_{j_2 \in \{\pm 1\}} N_{jj_2}(\partial_t \psi_c + i\omega \psi_c, R_{j_2}) \\
&+ \varepsilon^2 \vartheta^{-1} \Big(\sum_{j_2,j_3,j_4 \in \{\pm 1\}} \mathcal{T}_{jj_2j_3j_4}(\partial_t \psi_{j_4} + i\omega \psi_{j_4}, \psi_{j_4}, R_{j_3}) \\
&\qquad + \sum_{j_2,j_3,j_4 \in \{\pm 1\}} \mathcal{T}_{jj_2j_3j_4}(\psi_{j_4}, \partial_t \psi_{j_4} + i\omega \psi_{j_4}, R_{j_3}) \Big) \\
&+ \frac{\varepsilon^\beta}{2} \vartheta^{-1} \mathcal{G}_j(\vartheta R_{-1}, \vartheta R_1, \vartheta R_{-1}, \vartheta R_1) + \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_j}(\varepsilon \Psi) \\
&+ \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}\big(\psi_c, \frac{\varepsilon^\beta}{2} \vartheta^{-1} \mathcal{G}_{j_2}(\vartheta R_{-1}, \vartheta R_1, \vartheta R_{-1}, \vartheta R_1)\big) \\
&+ \varepsilon \sum_{j_2 \in \{\pm 1\}} \vartheta^{-1} N_{jj_2}\big(\psi_c, \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_{j_2}}(\varepsilon \Psi)\big) \\
&+ \varepsilon^3 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} \vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4}, \psi_{j_4}, \vartheta^{-1} \mathcal{G}_{j_3}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\big) \\
&+ \varepsilon^2 \sum_{j_2,j_3,j_4 \in \{\pm 1\}} \vartheta^{-1} \mathcal{T}_{jj_2j_3j_4}\big(\psi_{j_4}, \psi_{j_4}, \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_{j_3}}(\varepsilon \Psi)\big) \\
&+ \mathcal{O}(\varepsilon^2) \big(\|R_{-1}\|^2_{H^m} + \|R_1\|^2_{H^m} \big)^{1/2}.
\end{aligned}
$$

Due to the bound (3.46) for $\partial_t \psi_{\pm 1} + i\omega \psi_{\pm 1}$, we obtain that the L^2-Norms of the second, third and forth term are $\mathcal{O}(\varepsilon^2)(\|R_{-1}\|^2_{H^1} + \|R_1\|^2_{H^1})^{1/2}$ by using the estimates (2.32), and (3.76), (3.77) and (3.81).
Due to our choice of $\beta = 5/2$ and Ψ, i.e. due to (2.38), and, (2.32) and (3.43), the L^2-Norm of the fifth and sixth term are bounded by $\mathcal{O}(\varepsilon^2)\big(\varepsilon^{1/2}\, (\|R_{-1}\|^2_{H^m} + \|R_1\|^2_{H^m}) + 1\big)$.
Now, we also see, by using the estimates (2.32), (3.76), (3.77) and (3.81) that the L^2-Norms of the last three terms are bounded by $\mathcal{O}(\varepsilon^2)\big(\varepsilon^{1/2}(\|R_{-1}\|^2_{H^m} + \|R_1\|^2_{H^m}) + 1\big)$.

We now obtain

$$
\begin{aligned}
\partial_t E_0 &\le \mathcal{O}\big((\|R_{-1}\|^2_{H^1} + \|R_1\|^2_{H^1})^{1/2} \big) \sum_{j=\pm 1} \|\partial_t \check{R}_j - j i\omega \check{R}_j\|_{L^2} \\
&\le \varepsilon^2\, \mathcal{O}(\varepsilon^{1/2}\, (\|R_{-1}\|^2_{H^m} + \|R_1\|^2_{H^m})^{3/2} + \|R_{-1}\|^2_{H^m} + \|R_1\|^2_{H^m} + 1).
\end{aligned}
$$

(Note $\sqrt{x} \le x + 1$ for $x > 0$.) $\square$

Lemma 3.3.5. *For $j \in \{\pm 1\}$, $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $f_j, f \in H^1(\mathbb{R})$, we have*

$$\int_{\mathbb{R}} f_j \, N_{j-j}(g, f_{-j}) \, dx + \int_{\mathbb{R}} f_{-j} \, N_{-jj}(g, f_j) \, dx \leq \mathcal{O}\big(\|g\|_{L^1}\big) \, \|f_{-1}\|_{L^2} \|f_1\|_{L^2} \,, \quad (3.83)$$

$$\int_{\mathbb{R}} f \, N_{jj}(g, f) \, dx \leq \mathcal{O}\big(\|g\|_{L^1}\big) \, \|f\|_{L^2}^2 \,. \quad (3.84)$$

Proof. By using (2.66), i.e. (2.68), we get

$$\begin{aligned}
&\int_{\mathbb{R}} f_j \, N_{j-j}(g, f_{-j}) \, dx + \int_{\mathbb{R}} f_{-j} \, N_{-jj}(g, f_j) \, dx \\
&\qquad = \int_{\mathbb{R}} f_{-1} \, N_{-11}(g, f_1) \, dx + \int_{\mathbb{R}} f_1 \, N_{1-1}(g, f_{-1}) \, dx \\
&\qquad = \int_{\mathbb{R}} f_{-1} \, \big(N_{-11}(g, f_1) + N^*_{1-1}(g, f_1)\big) \, dx \,,
\end{aligned}$$

where, just as in (2.67), we use the notation

$$\widehat{N}^*_{j_1 j_2}(\psi_c, f)(k) := \int_{\mathbb{R}} n_{j_1 j_2}(-m, k-m, -k) \widehat{\psi}_c(k-m) \widehat{f}(m) \, dm \,.$$

For $|k|, |m| \geq \delta$, we have

$$\begin{aligned}
&n_{-11}(k, k-m, m) + n_{1-1}(-m, k-m, -k) \\
&\qquad = \frac{\rho_{-11}(k, k-m, m) \; \chi_c(k-m)}{\omega(k) + \omega(m) - \omega(k-m)} + \frac{\rho_{1-1}(-m, k-m, -k) \; \chi_c(k-m)}{\omega(-k) + \omega(-m) + \omega(k-m)} \\
&\qquad = \frac{i \; \chi_c(k-m)}{\omega(k) + \omega(m) - \omega(k-m)} \big(c_{-1}(k, k-m, m) + a_{1,s}(-m, k-m, -k)\big) \,.
\end{aligned}$$

By making a similar expansion as in (2.55), we get

$$\begin{aligned}
&\int_{\mathbb{R}} f_{-1} \, \big(N_{-11}(g, f_1) + N^*_{1-1}(g, f_1)\big) \, dx \\
&\qquad = \int_{\mathbb{R}} f_{-1} \, \frac{1}{2 \, i\omega} \big(C_{-1}(g_c, f_1) + A^*_{1,s}(g_c, f_1)\big) \, dx + \mathcal{O}\big(\|g\|_{L^1}\big) \, \|f_{-1}\|_{L^2} \|f_1\|_{L^2}.
\end{aligned}$$

due to (3.54).
Then, by using Cauchy-Schwarz, we obtain (3.83) due to (3.56).

Now, we prove (3.84).
By using (2.66), i.e. (2.68), we get

$$\int_{\mathbb{R}} f \, N_{jj}(g, f) \, dx = \frac{1}{2} \int_{\mathbb{R}} f \, \big(N_{jj}(g, f) + N^*_{jj}(g, f)\big) \, dx \,.$$

Therefore, we have

$$\int_{\mathbb{R}} f\, N_{jj}(g,f)\,dx \le \frac{1}{2}\|f\|_{L^2}\big\|N_{jj}(g,f)+N_{jj}^*(g,f)\big\|_{L^2}\,,$$

where

$$\begin{aligned}&\widehat{N}_{jj}(g,f)+\widehat{N}_{jj}^*(g,f)\\&\qquad=\int_{\mathbb{R}}\big(n_{j_1j_2}(k,k-m,m)+n_{j_1j_2}(-m,k-m,-k)\big)\widehat{g}(k-m)\widehat{f}(m)\,dm\,,\end{aligned}$$

and

$$\begin{aligned}&n_{jj}(k,k-m,m)+n_{jj}(-m,k-m,-k)\\&\qquad=\frac{\rho_{jj}(k,k-m,m)+\rho_{jj}(-m,k-m,-k)}{\omega(k)-\omega(m)+j\,\omega(k-m)}\,\chi_c(k-m)\qquad\text{for } |k|\to\infty.\end{aligned}$$

Since

$$\begin{aligned}\rho_{-1-1}(k,k-m,m)&+\rho_{-1-1}(-m,k-m,-k)\\&=i\,\big(a_{-1}(k,k-m,m)+a_{-1}(-m,k-m,-k)\\&\qquad+a_{-1}(k,m,k-m)+a_{-1}(-m,-k,k-m)\big)\,,\end{aligned}$$

$$\begin{aligned}\rho_{11}(k,k-m,m)&+\rho_{11}(-m,k-m,-k)\\&=-i\big(c_1(k,k-m,m)+c_1(-m,k-m,-k)\big)\,,\end{aligned}$$

we obtain

$$\big\|N_{jj}(g,f)+N_{jj}^*(g,f)\big\|_{L^2}\le\mathcal{O}\big(\|g\|_{L^1}\big)\,\|f\|_{L^2}\,.$$

due to (3.55) (e.g. by exploiting (2.54) and (2.52)). □

Corollary 3.3.6. *Let $\varepsilon<\varepsilon_0$ and ε_0 be sufficiently small.*
For $\ell\ge 1$, the energy $\mathcal{E}_\ell$ is equivalent to $\big(\|R_{-1}\|_{H^\ell}+\|R_1\|_{H^\ell}\big)^2$, i.e. there are constants $C_1,C_2>0$ such that

$$\big(\|R_{-1}\|_{H^\ell}+\|R_1\|_{H^\ell}\big)^2\quad\le\quad C_1\,\mathcal{E}_\ell\quad\le\quad C_2\,\big(\|R_{-1}\|_{H^\ell}+\|R_1\|_{H^\ell}\big)^2.$$

Proof. In particular thanks to (3.83) and (3.84), the proof works analogous to the combined proof of lemma 2.2.14 and corollary 2.2.16. □

Lemma 3.3.7. *For $\ell\ge 1$, we have*

$$\partial_t E_\ell=\varepsilon^2\,V_\ell+\varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell+1),\tag{3.85}$$

where

$$V_\ell = \sum_{j_1\in\{\pm1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi_q}, R_1^{\Psi_q}, \vartheta R_{-1}, \vartheta R_1)\, dx \tag{3.86}$$
$$+ \sum_{j_1,j_2\in\{\pm1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2})\, dx$$
$$+ \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_2}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big),$$

and

$$R_j^{\Psi_q} = \psi_{q_j} + \frac{1}{2}\varepsilon^{\beta-2}(\vartheta R_j) \tag{3.87}$$

with ψ_{q_j} *as in* (3.70).

Remark 3.3.8. *Due to* (2.73) *and* (2.37), *we see that* $\varepsilon^2 V_\ell$ *has the desired* ε*-order.*

Proof. The proof is analogous to the one of lemma 2.2.17.
We have

$$\partial_t E_\ell = \sum_{j_1\in\{\pm1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_t \partial_x^\ell R_{j_1}\, dx$$
$$+ \varepsilon \sum_{j_1,j_2\in\{\pm1\}} \Big(\int_{\mathbb{R}} \partial_t \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2})\, dx$$
$$+ \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} \partial_t N_{j_1 j_2}(\psi_c, R_{j_2})\, dx \Big).$$

Using the error equations (3.66) and (3.67), and exploiting

$$\begin{pmatrix} R_{-1}^{\Psi} \\ R_1^{\Psi} \end{pmatrix} = \begin{pmatrix} \psi_c \\ 0 \end{pmatrix} + \varepsilon \begin{pmatrix} R_{-1}^{\Psi_q} \\ R_1^{\Psi_q} \end{pmatrix},$$

we get

$$
\begin{aligned}
\partial_t E_\ell = & \sum_{j_1 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, i\omega \partial_x^\ell R_{j_1} \, dx \\
& + \varepsilon \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(\psi_c, 0, \vartheta R_{-1}, \vartheta R_1) \, dx \\
& \qquad + \sum_{j_2 \in \{\pm 1\}} \Big(+ j_1 \int_{\mathbb{R}} i\omega \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad\qquad + j_2 \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, i\omega R_{j_2}) \, dx \\
& \qquad\qquad - \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(i\omega \psi_c, R_{j_2}) \, dx \\
& \qquad\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\partial_t \psi_c + i\omega \psi_c, R_{j_2}) \, dx \Big) \Big) \\
& + \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi_q}, R_1^{\Psi_q}, \vartheta R_{-1}, \vartheta R_1) \, dx \\
& + \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \vartheta^{-1} \mathcal{G}_{j_2}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\big) \, dx \Big) \\
& + \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon \Psi) \, dx \\
& + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon \Psi) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_{j_2}}(\varepsilon \Psi)\big) \, dx \Big) .
\end{aligned}
$$

Exploiting the skew symmetry of $i\omega$ in the third integral and then using (3.74) and the definition (3.86), we get

$$
\begin{aligned}
\partial_t E_\ell = {}& \varepsilon \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}\big(\psi_c, 0, (\vartheta - \vartheta_{\varepsilon,\infty}) R_{-1}, (\vartheta - \vartheta_{\varepsilon,\infty}) R_1\big)\, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\partial_t \psi_c + i\omega\psi_c, R_{j_2})\, dx \Big) \\
& + \varepsilon^2 V_\ell \\
& + \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon\Psi)\, dx \\
& + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \varepsilon^{-\beta} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{j_1}}(\varepsilon\Psi)\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}(\psi_c, R_{j_2})\, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_2}\big(\psi_c, \varepsilon^{-\beta} \vartheta^{-1} \mathrm{Res}_{u_{j_2}}(\varepsilon\Psi)\big)\, dx \Big) .
\end{aligned}
$$

We now show that all terms except the term $\varepsilon^2 V_\ell$ can be estimated against $\varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1)$. Thereby we will especially take advantage of corollary 3.3.6 and (3.45).
For the first integral, we can use (2.73), Cauchy-Schwarz and the fact that

$$
\big(\hat{\vartheta}(k) - \hat{\vartheta}_{\varepsilon,\infty}(k)\big) = \begin{cases} \varepsilon + (1-\varepsilon)\frac{|k|}{\delta} & \text{when } 0 \neq \pm\omega(0^+) \neq 2\omega(k_0) \text{ and } |k| \leq \varepsilon\,, \\ 0 & \text{else}\,, \end{cases}
$$

in order to get

$$
\varepsilon \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}\big(\psi_c, 0, (\vartheta - \vartheta_{\varepsilon,\infty}) R_{-1}, (\vartheta - \vartheta_{\varepsilon,\infty}) R_1\big)\, dx \leq \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell).
$$

The second integral in the above evolution equality is $\varepsilon^3 \mathcal{O}(\mathcal{E}_\ell)$ due to the estimate (3.46). We obtain this by first using (2.73) and then exploiting (3.83) and (3.84) in order to estimate without losing regularity.

The last three integrals are $\varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1)$ due to (3.43). To see this, we use first (2.73), then integration by parts to shift some derivatives away from $R_{\pm 1}$, and finally Cauchy-Schwarz together with (3.76) and (3.77). Here, we also exploit the estimate $\sqrt{x} \leq |x| + 1$ after using corollary 3.3.6.

□

3.3.2 Closing the error estimates via energy transformations

We will in the following close our error estimates such that theorem 1.2.1 follows. Apart from some technical details, closing our energy estimates will in some sense

be easy since we are formally in the case $\deg^*(\rho) \leq 1$. For $b > 0$, which corresponds to the for water wave problem unsolved case, all the more so since we have on top of that $\deg(\omega) > \deg^*(\rho)$.
Interestingly, the case $b = 0$ is more difficult due to the fact that in this case we formally have $\deg(\omega) < \deg^*(\rho)$. The arc length formulation of the water wave problem seems to suit the case $b > 0$ better than the case $b = 0$.
Up to this point everything we have proved held for $b \geq 0$. We will in the following close our energy estimates for $b > 0$ in such a way that one also sees, which terms would have to be analyzed further for $b = 0$. In order to close the energy estimates for $b = 0$, one would have to exploit some additional key properties of the system (3.32).

We will now first prove two technical lemmas needed, before we will close our energy estimates.

Lemma 3.3.9. *Let $N \in \mathbb{N}$ and $\ell \geq 2N + 1$.*
By introducing the notation

$$\tilde{R}_j^{\Psi} := \Psi_j + \varepsilon^{\beta-1}\vartheta R_j \,, \tag{3.88}$$

we obtain

$$\begin{aligned} &\partial_x^\ell \mathcal{G}_j(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1) \\ &\quad= \sum_{n=0}^{N} \binom{\ell}{n} \mathcal{G}_j(\partial_x^n \tilde{R}_{-1}^\Psi, \partial_x^n \tilde{R}_1^\Psi, \partial_x^{\ell-n}\vartheta R_{-1}, \partial_x^{\ell-n}\vartheta R_1) \\ &\qquad+ \sum_{n=N+1}^{\ell-N-1} \binom{\ell}{n} \mathcal{G}_j(\partial_x^n R_{-1}^\Psi, \partial_x^n R_1^\Psi, \partial_x^{\ell-n}\vartheta R_{-1}, \partial_x^{\ell-n}\vartheta R_1) \\ &\qquad+ \sum_{n=\ell-N}^{\ell} \binom{\ell}{n} \mathcal{G}_j(\partial_x^n \Psi_{-1}, \partial_x^n \Psi_1, \partial_x^{\ell-n}\vartheta R_{-1}, \partial_x^{\ell-n}\vartheta R_1) \,. \end{aligned} \tag{3.89}$$

Proof. According to (3.68), we have

$$\begin{aligned} \partial_x^\ell \mathcal{G}_j(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1) &= \partial_x^\ell \big(\mathcal{A}_j(R_{-1}^\Psi, \vartheta R_{-1}) + \mathcal{A}_j(\vartheta R_{-1}, R_{-1}^\Psi)\big) \\ &\quad+ \partial_x^\ell \big(\mathcal{B}_j(R_1^\Psi, \vartheta R_1) + \mathcal{B}_j(\vartheta R_1, R_1^\Psi)\big) \\ &\quad+ \partial_x^\ell \big(\mathcal{C}_j(R_{-1}^\Psi, \vartheta R_1) + \mathcal{C}_j(\vartheta R_{-1}, R_1^\Psi)\big) \,. \end{aligned}$$

Leibniz's rule and the definition of $R_j^\Psi := \Psi_j + \frac{1}{2}\varepsilon^{\beta-1}\vartheta R_j$ yield that for continuous bilinear operators $\mathcal{Z}$:

$$\begin{aligned}\partial_x^\ell \mathcal{Z}(R_{j_1}^\Psi, \vartheta R_{j_2}) &= \sum_{n=0}^{N} \binom{\ell}{n} \mathcal{Z}(\partial_x^n R_{j_1}^\Psi, \partial_x^{\ell-n} \vartheta R_{j_2}) \\ &\quad + \sum_{n=N+1}^{\ell-N-1} \binom{\ell}{n} \mathcal{Z}(\partial_x^n R_{j_1}^\Psi, \partial_x^{\ell-n} \vartheta R_{j_2}) \\ &\quad + \sum_{n=\ell-N}^{\ell} \binom{\ell}{n} \mathcal{Z}(\partial_x^n \Psi_{j_1}, \partial_x^{\ell-n} \vartheta R_{j_2}) \\ &\quad + \frac{1}{2}\varepsilon^{\beta-1} \sum_{n=\ell-N}^{\ell} \binom{\ell}{n} \mathcal{Z}(\partial_x^n \vartheta R_{j_1}, \partial_x^{\ell-n} \vartheta R_{j_2}) \,.\end{aligned}$$

Proceeding analogously for $\mathcal{Z}(\vartheta R_{j_2}, R_{j_1}^\Psi)$ and then combining all three equations, we now obtain (3.89) due to the fact that

$$\sum_{n=\ell-N}^{\ell} \binom{\ell}{n} \mathcal{Z}(\partial_x^n f, \partial_x^{\ell-n} g) = \sum_{\tilde{n}=0}^{N} \binom{\ell}{\tilde{n}} \mathcal{Z}(\partial_x^{\ell-\tilde{n}} f, \partial_x^{\tilde{n}} g) \,.$$

□

Lemma 3.3.10. *Let*

$$\widehat{\mathcal{Z}}_j(f,g) := \int_{\mathbb{R}} z_j(k, k-m, m) \, \widehat{f}(k-m) \, \widehat{g}(m) \, dm$$

where

$$z_j(k, k-m, m) = \sum_{i=0}^{I_j} z_{j,i}^1(k) z_{j,i}^2(k-m) z_{j,i}^3(m)$$

with $I_j < \infty$. *The functions* $z_{j,i}^1$, $z_{j,i}^2$ *and* $z_{j,i}^3$ *shall be sufficiently smooth, fulfill*

$$\deg^*(z_{j,i}^1), \deg^*(z_{j,i}^2), \deg^*(z_{j,i}^3) < \infty \,,$$

and have the property (1.10).
Let f, h *and* g *be sufficiently regular.*
For $A := \mathcal{Z}_1(f, \cdot)$, $B := \mathcal{Z}_2(h, \cdot)$ *there is some* $C(A,B,f,g) > 0$ *such that for all* $g \in H^{s+1}$, *we have*

$$\|ABg - BAg\|_{L^2} \leq C(A,B,f,g) \, \|g\|_{H^s} \,, \tag{3.90}$$

where

$s :=$

$$\max\Big\{ \max_{i \in \{0,\dots,I_1\}} \big(\deg^*(z_{1,i}^1) + \deg^*(z_{1,i}^3) \big) + \max_{i \in \{0,\dots,I_2\}} \big(\deg^*(z_{2,i}^1) + \deg^*(z_{2,i}^3) \big) - 1 \,, 0 \Big\}.$$

Remark 3.3.11. *By proceeding similarly as in the proof for lemma 2.2.20, one can make a more precise estimate and explicitly calculate the required regularity for f and h.*

Before, we prove this rather abstractly formulated lemma, let us first bring an example to illustrate the idea behind it.

Example 3.3.12. Let $g := \partial_x^\ell R$.
Moreover, let

$$A g = \partial_x (f \partial_x^{-1} g)\,, \qquad\qquad B g = \partial_x h \partial_x g\,.$$

We have

$$\|Ag\|_{L^2} \leq \|f\|_{H^2} \|\partial_x^{-1} g\|_{H^1}\,, \qquad\qquad \|Bg\|_{L^2} \leq \|h\|_{H^2} \|g\|_{H^1}\,.$$

We get

$$\begin{aligned} ABg - BAg &= \partial_x\big(f\partial_x^{-1}(\partial_x h \partial_x g)\big) - \partial_x h \partial_x^2 (f \partial_x^{-1} g) \\ &= f\partial_x h \partial_x g - \partial_x h f \partial_x g \\ &\quad + \partial_x f \partial_x^{-1}(\partial_x h \partial_x g) - \partial_x h (2\partial_x f g + \partial_x^2 f \partial_x^{-1} g) \\ &= \partial_x f \partial_x^{-1}(\partial_x h \partial_x g) - \partial_x h (2\partial_x f g + \partial_x^2 f \partial_x^{-1} g) \end{aligned}$$

and

$$\begin{aligned} \|\partial_x f \partial_x^{-1}(\partial_x h \partial_x g)\|_{L^2} &= \|\partial_x f\big(\partial_x h g - \partial_x^{-1}(\partial_x^2 h g)\big)\|_{L^2} \\ &\leq \|\partial_x f\|_\infty \|\partial_x h\|_\infty \|g\|_{L^2} + \|\partial_x f\|_{L^2} \Big\| \int_{\mathbb{R}} \partial_x^2 h g \, dx \, \Big\|_\infty \\ &\leq \|f\|_{H^2} \|h\|_{H^2} \|g\|_{L^2}\,, \end{aligned}$$

$$\|\partial_x h (2\partial_x f g + \partial_x^2 f \partial_x^{-1} g)\|_{L^2} \leq \|f\|_{H^3} \|h\|_{H^2} \|\partial_x^{-1} g\|_{H^1}\,.$$

Thus, we obtain

$$\|AB\partial_x^\ell R - BA\partial_x^\ell R\|_{L^2} \leq \mathcal{O}\big(\|f\|_{H^3}\|h\|_{H^2}\big)\, \|\partial_x^{l-1} R\|_{H^1}\,.$$

This is a much better result than the one we could obtain by estimating the L^2-norms of $AB\partial_x^\ell R$ and $BA\partial_x^\ell R$ separately, what would involve an estimate like

$$\|f \partial_x h \partial_x g\|_{L^2} = \|f\partial_x h \partial_x^{\ell+1} R\|_{L^2} \leq \|f\|_{H^1} \|h\|_{H^2} \|\partial_x^{\ell+1} R\|_{L^2}\,.$$

Proof of lemma 3.3.10. Without a loss of generality, let

$$\begin{aligned} A(f,g) &:= \int_{\mathbb{R}} a^1(k) \tilde{a}^2(k-m) a^3(m)\, \widehat{g}(m)\, dm\,, \\ B(h,g) &:= \int_{\mathbb{R}} b^1(k) \tilde{b}^2(k-m) b^3(m)\, \widehat{g}(m)\, dm\,, \end{aligned}$$

$\tilde{a}^2(k-m) := a^2(k-m)f(k-m)$ and $\tilde{b}^2(k-m) := b^2(k-m)h(k-m)$.
By using the Taylor expansions

$$\begin{aligned} a^1(k) &= a^1(n) + (a^1)'(n)\,(k-n) + \dots \\ &= a^1(n) + (a^1)'(n)\,(k-m) + (a^1)'(n)\,(m-n) + \dots\,, \end{aligned}$$

$$a^3(m) = a^3(n) + (a^3)'(n)\,(m-n) + \dots\,,$$

$$b^1(m) = b^1(n) + (b^1)'(n)\,(m-n) + \dots\,,$$

we have

$$\begin{aligned} \widehat{AB}(g) &= \int_{\mathbb{R}}\int_{\mathbb{R}} a^1(k)\tilde{a}^2(k-m)a^3(m)\,b^1(m)\tilde{b}^2(m-n)b^3(n)\,\widehat{g}(n)\,dm\,dn \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{a}^2(k-m)\tilde{b}^2(m-n)\,a^1(n)a^3(n)b^1(n)b^3(n)\,\widehat{g}(n)\,dm\,dn \\ &\quad + \mathcal{R}_1\,, \end{aligned}$$

where in $\mathcal{R}_1$ we just collected all the other integrals that emerged.
Via analogous Taylor expansions, we can get

$$\begin{aligned} \widehat{BA}(g) &= \int_{\mathbb{R}}\int_{\mathbb{R}} b^1(k)\tilde{b}^2(k-m)b^3(m)\,a^1(m)\tilde{a}^2(m-n)a^3(n)\,\widehat{g}(n)\,dm\,dn \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{b}^2(k-m)\tilde{a}^2(m-n)\,a^1(n)a^3(n)b^1(n)b^3(n)\,\widehat{g}(n)\,dm\,dn \\ &\quad + \mathcal{R}_2\,, \end{aligned}$$

where in $\mathcal{R}_2$ we just collected all the other integrals that emerged.
Since the convolution of functions is a commutative operation, we now get

$$\begin{aligned} \widehat{AB}(g) - \widehat{BA}(g) &= \int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{a}^2(k-m)\tilde{b}^2(m-n)\,a^1(n)a^3(n)b^1(n)b^3(n)\,\widehat{g}(n)\,dm\,dn \\ &\quad - \int_{\mathbb{R}}\int_{\mathbb{R}} \tilde{b}^2(k-m)\tilde{a}^2(m-n)\,a^1(n)a^3(n)b^1(n)b^3(n)\,\widehat{g}(n)\,dm\,dn \\ &\quad + \mathcal{R}_1 - \mathcal{R}_2 \\ &= \mathcal{R}_1 - \mathcal{R}_2. \end{aligned}$$

Due to the Taylor expansions and Plancherel, the L^2-norm of an integral from $\mathcal{R}_1$ and $\mathcal{R}_2$ can be easily estimated, e.g.

$$\begin{aligned} &\Big\| \int_{\mathbb{R}}\int_{\mathbb{R}} \nu(k-m)\tilde{a}^2(k-m)\sigma(m-n)\tilde{b}^2(m-n)\,\gamma(n)\widehat{g}(n)\,dm\,dn\Big\|_{L^2} \\ &\qquad = \mathcal{O}(1)\,\|\mathcal{F}^{-1}[\nu\tilde{a}^2]\,\mathcal{F}^{-1}[\sigma\tilde{b}^2]\mathcal{F}^{-1}[\gamma\widehat{g}]\|_{L^2} \\ &\qquad \le \mathcal{O}\big(\|\mathcal{F}^{-1}[\nu\tilde{a}^2]\|_{H^1}\|\mathcal{F}^{-1}[\sigma\tilde{b}^2]\|_{H^1}\big)\,\|\mathcal{F}^{-1}[\gamma\widehat{g}]\|_{L^2}\,. \end{aligned}$$

For the integrals involving the remainder this of course has to be modified by proceeding similarly as in the proof for lemma 2.2.20.
Since the property (1.10) is assumed, the L^2-norm of $\mathcal{R}_1$ and $\mathcal{R}_2$ can be estimated by terms only involving the H^s-norm of g, where

$$s = \max\left\{ \deg^*(a^1) + \deg^*(a^3) + \deg^*(b^1) + \deg^*(b^3) - 1\,, \quad 0 \right\}.$$

$\square$

Remark 3.3.13. *As in the last section, we will from now on assume that* $\varepsilon\mathcal{E}_\ell \leq 1$ *for* $\varepsilon < \varepsilon_0$.

Corollary 3.3.14. *Let* $b > 0$ *and* $\ell \geq 4$.
For $\varepsilon < \varepsilon_0$ *and* ε_0 *sufficiently small, there exists an energy* $\tilde{\mathcal{E}}_\ell$ *and some constants* $c, C > 0$ *such that*

$$\left(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\right)^2 \leq c\,\tilde{\mathcal{E}}_\ell \leq C\left(\|R_{-1}\|_{H^\ell} + \|R_1\|_{H^\ell}\right)^2 \tag{3.91}$$

and

$$\partial_t\tilde{\mathcal{E}}_\ell \leq \varepsilon^2\,\mathcal{O}\big(\tilde{\mathcal{E}}_\ell + 1\big)\,.$$

Remark 3.3.15. *Up to this point we have proved everything for* $b \geq 0$. *While we only prove corollary 3.3.14 for the case* $b > 0$, *we will present the proof in such a way that it is clear which terms have to be further analyzed for the case* $b = 0$.

Proof. According to the definition of $\mathcal{E}_\ell$ in (3.73) and due to lemma 3.3.4, we have

$$\partial_t\mathcal{E}_\ell = \partial_t E_0 + \partial_t E_\ell = \partial_t E_\ell + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,,$$

where, due to lemma 3.3.7,

$$
\begin{aligned}
\partial_t E_\ell &= \varepsilon^2 V_\ell + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1) \\
&= \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi_q}, R_1^{\Psi_q}, \vartheta R_{-1}, \vartheta R_1) \, dx \\
&\quad + \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1) \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_1}(\psi_c, R_{j_1}) \, dx \\
&\quad + \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 j_1}\big(\psi_c, \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\big) \, dx \\
&\quad + \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1) \, \partial_x^\ell \vartheta^{-1} N_{j_1 -j_1}(\psi_c, R_{-j_1}) \, dx \\
&\quad + \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell R_{j_1} \, \partial_x^\ell \vartheta^{-1} N_{j_1 -j_1}\big(\psi_c, \vartheta^{-1} \mathcal{G}_{-j_1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\big) \, dx \\
&\quad + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1) \\
&=: \sum_{i=0}^{4} I_i + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1) \,.
\end{aligned}
$$

First, we analyze the term I_0.
Using (2.73), we get

$$
\begin{aligned}
I_0 = \varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{-1} \, \partial_x^\ell \mathcal{G}_{-1}(R_{-1}^{\Psi_q}, R_1^{\Psi_q}, \vartheta R_{-1}, \vartheta R_1) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_1 \, \partial_x^\ell \mathcal{G}_1(R_{-1}^{\Psi_q}, R_1^{\Psi_q}, \vartheta R_{-1}, \vartheta R_1) \, dx + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-2} \mathcal{E}_\ell^{3/2}) \,.
\end{aligned} \tag{3.92}
$$

By proceeding analogously as in (3.89) and setting

$$
\tilde{R}_j^{\Psi_q} = \psi_{q_j} + \varepsilon^{\beta-2} (\vartheta R_j) \,, \tag{3.93}
$$

we obtain

$$
\begin{aligned}
I_0 = \varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{-1} \, \mathcal{G}_{-1}(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}, \partial_x^\ell \vartheta R_1) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_1 \, \mathcal{G}_1(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}, \partial_x^\ell \vartheta R_1) \, dx + \varepsilon^2 \mathcal{O}(\mathcal{E}_\ell + 1) \,,
\end{aligned}
$$

due to (3.54).

Plugging in the definition of (3.68), we get

$$
\begin{aligned}
I_0 = \varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{-1} \big(\mathcal{A}_{-1,s}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_{-1}(\partial_x^\ell \vartheta R_{-1}, \tilde{R}_1^{\Psi_q}) \big) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1} \big(\mathcal{B}_{-1,s}(\tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta R_1) + \mathcal{C}_{-1}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_1) \big) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_1 \big(\mathcal{A}_{1,s}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_1(\partial_x^\ell \vartheta R_{-1}, \tilde{R}_1^{\Psi_q}) \big) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_1 \big(\mathcal{B}_{1,s}(\tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta R_1) + \mathcal{C}_1(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_1) \big) \, dx \\
& + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \; .
\end{aligned}
$$

By using lemma 2.2.11 and (2.72) together with the notation (3.58), we have

$$
\begin{aligned}
\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_2} \, \mathcal{Z}(\tilde{R}_{j_1}^{\Psi_q}, \partial_x^\ell \vartheta R_{j_2}) \, dx &= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell \vartheta R_{j_2} \, \mathcal{Z}^*(\tilde{R}_{j_1}^{\Psi_q}, \partial_x^\ell R_{j_2}) \, dx \\
&= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_2} \, \mathcal{Z}^*(\tilde{R}_{j_1}^{\Psi_q}, \partial_x^\ell R_{j_2}) \, dx + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \, ,
\end{aligned}
$$

and thus

$$
\begin{aligned}
\varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{j_2} \, \mathcal{Z}(\tilde{R}_{j_1}^{\Psi_q}, \partial_x^\ell \vartheta R_{j_2}) \, dx \\
& = \frac{1}{2} \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_2} \, \big[\mathcal{Z} + \mathcal{Z}^* \big] (\tilde{R}_{j_1}^{\Psi_q}, \partial_x^\ell R_{j_2}) \, dx + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1)
\end{aligned} \tag{3.94}
$$

for $\mathcal{Z} = \mathcal{A}_{-1,s}$.
Now, we obtain

$$
\begin{aligned}
\int_{\mathbb{R}} & \partial_x^\ell R_{-1} \, \mathcal{A}_{-1,s}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}) \, dx \\
& = \frac{1}{2} \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1} \, \big[\mathcal{A}_{-1,s} + \mathcal{A}_{-1,s}^* \big] (\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell R_{-1}) \, dx + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \\
& = \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \, ,
\end{aligned}
$$

due to (3.55).
We can proceed analogous for $\mathcal{B}_{1,s}, \mathcal{C}_{-1}$ and similarly for $\mathcal{C}_1$, such that (3.55) yields

$$
\begin{aligned}
I_0 = \varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{-1} \big(\mathcal{B}_{-1,s}(\tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta R_1) + \mathcal{C}_{-1}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_1) \big) \, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_1 \big(\mathcal{A}_{1,s}(\tilde{R}_{-1}^{\Psi_q}, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_1(\partial_x^\ell \vartheta R_{-1}, \tilde{R}_1^{\Psi_q}) \big) \, dx \\
& + \varepsilon^2 \, \mathcal{O}(\mathcal{E}_\ell + 1) \; .
\end{aligned}
$$

Using lemma 2.2.11, (2.72) and the notation (3.58) again, we get

$$
\begin{aligned}
I_0 = \varepsilon^2 & \int_{\mathbb{R}} \partial_x^\ell R_{-1} \left(\mathcal{A}^*_{1,s}(\tilde{R}^{\Psi_q}_{-1}, \partial_x^\ell R_1) + \mathcal{C}_{-1}(\tilde{R}^{\Psi_q}_{-1}, \partial_x^\ell R_1)\right) dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1} \left(\mathcal{B}_{-1,s}(\tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1) + \mathcal{C}_{1,*}(\partial_x^\ell R_1, \tilde{R}^{\Psi_q}_{1})\right) dx \\
& + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1) \; .
\end{aligned}
$$

According to (3.56), the mapping $F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \cdot)$, which we define by

$$
F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \cdot) := \mathcal{A}^*_{1,s}(\tilde{R}^{\Psi_q}_{-1}, \cdot) + \mathcal{C}_{-1}(\tilde{R}^{\Psi_q}_{-1}, \cdot) + \mathcal{B}_{-1,s}(\tilde{R}^{\Psi_q}_{1}, \cdot) + \mathcal{C}_{1,*}(\cdot, \tilde{R}^{\Psi_q}_{1}) \, , \quad (3.95)
$$

maps $H^{1/2}(\mathbb{R})$ onto $L^2(\mathbb{R})$. By exploiting (3.66), (3.67), the skew symmetry of $i\omega$ and the properties $F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \cdot)$ inherited from (3.95), we have

$$
\begin{aligned}
& \frac{1}{2}\varepsilon^2\, \partial_t \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1)\, dx \\
& \qquad = \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, i\omega \partial_x^\ell R_1)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{-1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\, \frac{1}{i\omega} F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} F\big(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell \vartheta^{-1} \mathcal{G}_1(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} \Big(F(\partial_t \tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1) + F(\tilde{R}^{\Psi_q}_{-1}, \partial_t \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1) \Big)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^2 \varepsilon^{-\beta} \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_{-1}}(\varepsilon\Psi)\, \frac{1}{i\omega} F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell R_1)\, dx \\
& \qquad + \frac{1}{2}\varepsilon^2 \varepsilon^{-\beta} \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} F(\tilde{R}^{\Psi_q}_{-1}, \tilde{R}^{\Psi_q}_{1}, \partial_x^\ell \vartheta^{-1} \mathrm{Res}_{u_1}(\varepsilon\Psi))\, dx \, .
\end{aligned}
$$

The last three integrals are $\varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1)$ in particular due to lemma 3.2.2 and the assumption that $\varepsilon \mathcal{E}_\ell \leq 1$.

By adding a zero, we obtain

$$
\begin{aligned}
\frac{1}{2}\varepsilon^2\,\partial_t \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx & \\
= I_0 & \\
-\frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega}\Big(i\omega F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1) - F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, i\omega\partial_x^\ell R_1)\Big)\,dx & \\
+\frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1}\mathcal{G}_{-1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx & \\
+\frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F\big(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta^{-1}\mathcal{G}_1(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\big)\,dx & \\
+\varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,. &
\end{aligned}
$$

Using Plancherel for the second integral, we have

$$
\begin{aligned}
-\frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega}\Big(i\omega F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1) - F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, i\omega\partial_x^\ell R_1)\Big)\,dx & \\
\eqsim \varepsilon^2 \int_{\mathbb{R}} \overline{\widehat{\partial_x^\ell R_{-1}}(k)} \int_{\mathbb{R}} \tilde{f}(k, k-m, m)\frac{\omega(k)-\omega(m)}{\omega(k)}\widehat{\partial_x^\ell R_1}(m)\,dm\,dk\,, &
\end{aligned}
$$

where, according to (3.95), the function $\tilde{f}(k, k-m, m)$ can be explicitly given by

$$
\begin{aligned}
\tilde{f}(k, k-m, m) = {} & \big(a_1^s(-m, k-m, -k) + c_{-1}(k, k-m, m)\big)\widehat{\tilde{R}_{-1}^{\Psi_q}}(k-m) \\
& + \big(b_{-1}^s(k, k-m, m) + c_1(-m, -k, k-m)\big)\widehat{\tilde{R}_1^{\Psi_q}}(k-m)\,.
\end{aligned}
$$

For some $C > 0$, we have

$$
\Big|\Big(\frac{k^2}{|k|^2+1}\Big)\,\frac{\omega(k)-\omega(m)}{\omega(k)}\Big| \le C\big((|k|^2+1)^{-1/2}\,(|k-m|^2+1)\big)\,.
$$

This can be shown by using Taylor, exactly like in the proof of lemma 2.2.20. By exploiting (3.56), we therefore obtain

$$
\begin{aligned}
-\frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega}\Big(i\omega F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1) - F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, i\omega\partial_x^\ell R_1)\Big)\,dx & \\
\le \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,. &
\end{aligned}
$$

We now arrive at

$$
\begin{aligned}
I_0 = {} & \frac{1}{2}\varepsilon^2\,\partial_t \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx \\
& - \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{-1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx \\
& - \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F\big(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta^{-1} \mathcal{G}_1(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\,dx \\
& + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}
$$

In the case $b \neq 0$, we have $\deg(\omega) = 3/2$. Therefore we can make the estimates

$$
\begin{aligned}
& \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{-1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx \\
& \quad \leq \mathcal{O}(\varepsilon^3) \big\| \partial_x^{\ell-1} \vartheta^{-1} \mathcal{G}_{-1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1) \big\|_{L^2} \Big\| \frac{\partial_x}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1) \Big\|_{L^2} \\
& \quad \leq \varepsilon^3\,\mathcal{O}(\mathcal{E}_\ell + 1)
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{1}{2}\varepsilon^3 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F\big(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta^{-1} \mathcal{G}_1(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\,dx \\
& \quad \leq \mathcal{O}(\varepsilon^3) \|\partial_x^\ell R_{-1}\|_{L^2} \Big\| \frac{1}{i\omega} F\big(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell \vartheta^{-1} \mathcal{G}_1(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big) \Big\|_{L^2} \\
& \quad \leq \varepsilon^3\,\mathcal{O}(\mathcal{E}_\ell + 1)\,.
\end{aligned}
$$

Thus, we obtain

$$
I_0 = \frac{1}{2}\varepsilon^2\,\partial_t \int_{\mathbb{R}} \partial_x^\ell R_{-1}\,\frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\,dx + \varepsilon^2\,\mathcal{O}(\mathcal{E}_\ell + 1)\,,
$$

due to (3.56) and (3.54).

Now, we analyze the term $I_1 + I_2$.

Using (2.73) we get

$$
\begin{aligned}
& I_1 + I_2 && (3.96)\\
& \quad := \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_1}(\psi_c, R_{j_1})\, dx \\
& \qquad\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell \vartheta^{-1} N_{j_1 j_1}\big(\psi_c, \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big) \\
& \quad = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^\ell N_{j_1 j_1}(\psi_c, R_{j_1})\, dx \\
& \qquad\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, \partial_x^\ell N_{j_1 j_1}\big(\psi_c, \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big) \\
& \qquad\qquad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2}) \, .
\end{aligned}
$$

Exploiting Leibniz's rule, we get

$$
\begin{aligned}
& I_1 + I_2 \\
& \quad = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& \qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 j_1}(\partial_x \psi_c, \partial_x^{\ell-1} R_{j_1})\, dx \\
& \qquad\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 j_1}\big(\psi_c, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \\
& \qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell R_{j_1}\, N_{j_1 j_1}\big(\partial_x \psi_c, \partial_x^{\ell-1} \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big) \\
& \qquad\qquad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2}) \, ,
\end{aligned}
$$

due to (3.54) and (3.76), and, also (3.53) and (2.37).

By using (2.66), we get

$$
\begin{aligned}
& I_1 + I_2 \\
& \quad = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, \big[N_{j_1 j_1} + N_{j_1 j_1}^*\big](\psi_c, \partial_x^\ell R_{j_1})\, dx \\
& \qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 j_1}(\partial_x \psi_c, \partial_x^{\ell-1} R_{j_1})\, dx \\
& \qquad\qquad - \ell \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^{-1} N_{j_1 j_1}^*(\partial_x \psi_c, \partial_x^\ell R_{j_1})\, dx \Big) \\
& \qquad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2}) \, .
\end{aligned}
$$

By setting

$$\mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_1}) \tag{3.97}$$
$$:= \big[N_{j_1 j_1} + N^*_{j_1 j_1}\big](\psi_c, \partial_x^\ell R_{j_1}) + \ell\,\big(N_{j_1 j_1}(\partial_x \psi_c, \partial_x^{\ell-1} R_{j_1}) - \partial_x^{-1} N^*_{j_1 j_1}(\psi_c, \partial_x^\ell R_{j_1})\big)$$

and using (2.73), we get

$$I_1 + I_2 = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_x^\ell \mathcal{G}_{j_1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_1})\, dx$$
$$+\ \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2})\,.$$

Moreover, we have

$$\|\mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_1})\|_{L^2} \leq \mathcal{O}\big(\|R_{j_1}\|_{H^\ell}\big) \tag{3.98}$$

according to the proofs of (3.84) and (3.76).

Now, we are in almost the same situation as before for I_0. The only difference is that we here have the terms $\mathcal{N}_\ell(\psi_c, R_{j_1})$ inside the integrals instead of the terms $\partial_x^\ell R_{j_1}$. This makes everything a bit more complicated.

The definition (3.97) implies

$$\mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_1}) = \mathcal{N}_\ell^*(\psi_c, \partial_x^\ell R_{j_1}). \tag{3.99}$$

Moreover, according to (3.90), we have

$$\|\mathcal{Z}^*\big(f, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_1})\big) - \mathcal{N}_\ell\big(\psi_c, \mathcal{Z}^*(f, \partial_x^\ell R_{j_1})\big)\|_{L^2} \leq \mathcal{O}\big(\|R_{j_1}\|_{H^\ell}\big) \tag{3.100}$$

for $\mathcal{Z} \in \{\mathcal{A}_{-1,s}, \mathcal{A}_{1,s}, \mathcal{B}_{-1,s}, \mathcal{B}_{1,s}, \mathcal{C}_{-1}, \tilde{\mathcal{C}}_1\}$, with $\tilde{\mathcal{C}}_1(f,g) := \mathcal{C}_1(g,f)$.
These two facts will now allow us to proceed for $I_1 + I_2$ analogously as for I_0. Using (3.89), exploiting (3.54) and (3.98), and then plugging in the definition of (3.68), we get

$$I_1 + I_2 = \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{A}_{-1,s}(\tilde{R}^\Psi_{-1}, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_{-1}(\partial_x^\ell \vartheta R_{-1}, \tilde{R}^\Psi_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{-1})\, dx$$
$$+ \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{B}_{-1,s}(\tilde{R}^\Psi_1, \partial_x^\ell \vartheta R_1) + \mathcal{C}_{-1}(\tilde{R}^\Psi_{-1}, \partial_x^\ell \vartheta R_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{-1})\, dx$$
$$+ \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{A}_{1,s}(\tilde{R}^\Psi_{-1}, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_1(\partial_x^\ell \vartheta R_{-1}, \tilde{R}^\Psi_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_1)\, dx$$
$$+ \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{B}_{1,s}(\tilde{R}^\Psi_1, \partial_x^\ell \vartheta R_1) + \mathcal{C}_1(\tilde{R}^\Psi_{-1}, \partial_x^\ell \vartheta R_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_1)\, dx$$
$$+\ \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2})\,.$$

By using (2.66), (2.72), (3.100) and (3.99) we have

$$
\begin{aligned}
\varepsilon^2 \int_{\mathbb{R}} & \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_2})\, \mathcal{Z}(\tilde{R}_{j_1}^\Psi, \partial_x^\ell \vartheta R_{j_3})\, dx \\
&= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell \vartheta R_{j_3}\, \mathcal{Z}^*\big(\tilde{R}_{j_1}^\Psi, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_2})\big)\, dx \\
&= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_3}\, \mathcal{Z}^*\big(\tilde{R}_{j_1}^\Psi, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_2})\big)\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2}) \\
&= \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{j_3}\, \mathcal{N}_\ell\big(\psi_c, \mathcal{Z}^*(\tilde{R}_{j_1}^\Psi, \partial_x^\ell R_{j_2})\big)\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2}) \\
&= \varepsilon^2 \int_{\mathbb{R}} \mathcal{N}_\ell^*(\psi_c, \partial_x^\ell R_{j_3})\, \mathcal{Z}^*(\tilde{R}_{j_1}^\Psi, \partial_x^\ell R_{j_2})\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2}) \\
&= \varepsilon^2 \int_{\mathbb{R}} \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_3})\, \mathcal{Z}^*(\tilde{R}_{j_1}^\Psi, \partial_x^\ell R_{j_2})\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
\end{aligned} \tag{3.101}
$$

Due to (3.55), we can use (3.101) to get

$$
\begin{aligned}
\varepsilon^2 \int_{\mathbb{R}} & \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_2})\, \mathcal{Z}(\tilde{R}_{j_1}^\Psi, \partial_x^\ell \vartheta R_{j_2})\, dx \\
&= \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{j_2})\, \big[\mathcal{Z} + \mathcal{Z}^*\big](\tilde{R}_{j_1}^\Psi, \partial_x^\ell R_{j_2})\, dx + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + 1) \\
&= \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,,
\end{aligned}
$$

for $\mathcal{Z} \in \{\mathcal{A}_{-1,s}, \mathcal{B}_{1,s}, \mathcal{C}_{-1}, \tilde{\mathcal{C}}_1\}$ with $\tilde{\mathcal{C}}_1(f,g) := \mathcal{C}_1(g,f)$.
We arrive at

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 & \int_{\mathbb{R}} \big(\mathcal{B}_{-1,s}(\tilde{R}_1^\Psi, \partial_x^\ell \vartheta R_1) + \mathcal{C}_{-1}(\tilde{R}_{-1}^\Psi, \partial_x^\ell \vartheta R_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{-1})\, dx \\
&+ \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{A}_{1,s}(\tilde{R}_{-1}^\Psi, \partial_x^\ell \vartheta R_{-1}) + \mathcal{C}_1(\partial_x^\ell \vartheta R_{-1}, \tilde{R}_1^\Psi)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_1)\, dx \\
&+ \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
\end{aligned}
$$

By using (3.101) and (2.72), we get

$$
\begin{aligned}
I_1 + I_2 = \varepsilon^2 & \int_{\mathbb{R}} \big(\mathcal{A}_{1,s}^*(\tilde{R}_{-1}^\Psi, \partial_x^\ell R_1) + \mathcal{C}_{-1}(\tilde{R}_{-1}^\Psi, \partial_x^\ell R_1)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{-1})\, dx \\
&+ \varepsilon^2 \int_{\mathbb{R}} \big(\mathcal{B}_{-1,s}(\tilde{R}_1^\Psi, \partial_x^\ell R_1) + \mathcal{C}_{1,*}(\partial_x^\ell R_1, \tilde{R}_1^\Psi)\big)\, \mathcal{N}_\ell(\psi_c, \partial_x^\ell R_{-1})\, dx \\
&+ \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,,
\end{aligned}
$$

such that we are now in position to exploit (3.56).
When $b \neq 0$, we have an even better estimate for $\mathcal{N}_\ell(\psi_c, \cdot)$ than (3.98). In this

case, we have $\deg(\omega') = 1/2$ such that (3.97), (2.52), (3.54) and (3.55) yield

$$\|\mathcal{N}_\ell(\psi_c, f)\|_{H^{1/2}} \leq \mathcal{O}\big(\|f\|_{L^2}\big). \tag{3.102}$$

Due to (3.56), we therefore obtain

$$\begin{aligned}
I_1 + I_2 = {} & \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \mathcal{N}_\ell\Big(\psi_c, \big(\mathcal{A}^*_{1,s}(\tilde{R}^\Psi_{-1}, \partial_x^\ell R_1) + \mathcal{C}_{-1}(\tilde{R}^\Psi_{-1}, \partial_x^\ell R_1)\big)\Big)\, dx \\
& + \varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \mathcal{N}_\ell\Big(\psi_c, \big(\mathcal{B}_{-1,s}(\tilde{R}^\Psi_{1}, \partial_x^\ell R_1) + \mathcal{C}_{1,*}(\partial_x^\ell R_1, \tilde{R}^\Psi_{1})\big)\Big)\, dx \\
& + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2}) \\
= {} & \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
\end{aligned}$$

Now, we analyze the term $I_3 + I_4$.
Using (2.73) we get

$$\begin{aligned}
& I_3 + I_4 \\
& := \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^\ell \vartheta^{-1} N_{j_1 - j_1}(\psi_c, R_{-j_1})\, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{-j_1}\, \partial_x^\ell \vartheta^{-1} N_{-j_1 j_1}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\big)\, dx\ \Big) \\
& = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\, \partial_x^\ell N_{j_1 - j_1}(\psi_c, R_{-j_1})\, dx \\
& \qquad + \int_{\mathbb{R}} \partial_x^\ell R_{-j_1}\, \partial_x^\ell N_{-j_1 j_1}\big(\psi_c, \vartheta^{-1}\mathcal{G}_{j_1}(R^\Psi_{-1}, R^\Psi_1, \vartheta R_{-1}, \vartheta R_1)\big)\, dx\ \Big) \\
& \quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
\end{aligned} \tag{3.103}$$

Exploiting Leibniz's rule, we get

$$
\begin{aligned}
& I_3 + I_4 \\
&\quad = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 - j_1}(\psi_c, \partial_x^\ell R_{-j_1})\, dx \\
&\qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 - j_1}(\partial_x \psi_c, \partial_x^{\ell-1} R_{-j_1})\, dx \\
&\qquad\qquad + \int_{\mathbb{R}} \partial_x^\ell R_{-j_1}\, N_{-j_1 j_1}\big(\psi_c, \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \\
&\qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell R_{-j_1}\, N_{-j_1 j_1}\big(\partial_x \psi_c, \partial_x^{\ell-1} \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big) \\
&\qquad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2}),
\end{aligned}
$$

due to (3.54) and (3.77), and, also (3.53) and (2.37).
By using (2.66), we get

$$
\begin{aligned}
& I_3 + I_4 \\
&\quad = \varepsilon^2 \sum_{j_1 \in \{\pm 1\}} \Big(\int_{\mathbb{R}} \partial_x^\ell \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, \big[N_{j_1 - j_1} + N_{-j_1 j_1}^*\big](\psi_c, \partial_x^\ell R_{-j_1})\, dx \\
&\qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\, N_{j_1 - j_1}(\partial_x \psi_c, \partial_x^{\ell-1} R_{-j_1})\, dx \\
&\qquad\qquad + \ell \int_{\mathbb{R}} \partial_x^\ell R_{-j_1}\, N_{-j_1 j_1}\big(\partial_x \psi_c, \partial_x^{\ell-1} \vartheta^{-1} \mathcal{G}_{j_1}(R_{-1}^\Psi, R_1^\Psi, \vartheta R_{-1}, \vartheta R_1)\big)\, dx \Big) \\
&\qquad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1} \mathcal{E}_\ell^{3/2}).
\end{aligned}
$$

For $|k|, |m| \geq \delta$, we have

$$
\begin{aligned}
& n_{-11}(k, k-m, m) + n_{1-1}(-m, k-m, -k) \\
&\qquad = \frac{i\, \chi_c(k-m)}{\omega(k) + \omega(m) - \omega(k-m)} \big(c_{-1}(k, k-m, m) + a_{1,s}(-m, k-m, -k)\big).
\end{aligned}
$$

In the case $b \neq 0$, we have $\deg(\omega) = 3/2$ such that by making a similar expansion as in (2.55), we get that for $|k| \to \infty$:

$$
\begin{aligned}
& n_{-11}(k, k-m, m) + n_{1-1}(-m, k-m, -k) \\
&\quad = \Big(\frac{c_{-1}(k, k-m, m) + a_{1,s}(-m, k-m, -k)}{-2i\omega(k)} + \mathcal{O}(|k|^{-3/2}) \Big) \chi_c(k-m).
\end{aligned}
\tag{3.104}
$$

In the case $b \neq 0$, we can also improve the estimate (3.77) to

$$
\|N_{j-j}(\psi_c, g)\|_{H^{1/2}} \leq C \|g\|_{L^2} \tag{3.105}
$$

by making a similar expansion to (2.55).
We thus get

$$
\begin{aligned}
& I_3 + I_4 \\
& \quad = -\frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell \mathcal{G}_{-1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\, \frac{1}{i\omega}\big[C_{-1} + A_{1,s}^*\big](\psi_c, \partial_x^\ell R_1)\, dx \\
& \quad\quad - \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell \mathcal{G}_{1}(R_{-1}^{\Psi}, R_1^{\Psi}, \vartheta R_{-1}, \vartheta R_1)\, \frac{1}{i\omega}\big[C_{-1}^* + A_{1,s}\big](\psi_c, \partial_x^\ell R_{-1})\, dx \\
& \quad\quad + \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
\end{aligned}
$$

Due to (3.56) and $\deg(\omega) = 3/2$, we then obtain

$$
I_3 + I_4 = \varepsilon^2\, \mathcal{O}(\mathcal{E}_\ell + \varepsilon^{\beta-1}\mathcal{E}_\ell^{3/2})\,.
$$

Hence, by choosing ε_0 small enough and summing up our results for I_0-I_4, we can define a modified energy

$$
\tilde{\mathcal{E}}_\ell = \mathcal{E}_\ell - \varepsilon^2 \mathcal{D}_0\,,
$$

with

$$
\mathcal{D}_0 := \frac{1}{2}\varepsilon^2 \int_{\mathbb{R}} \partial_x^\ell R_{-1}\, \frac{1}{i\omega} F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)\, dx = \mathcal{O}(\mathcal{E}_\ell)\,,
$$

and $F(\tilde{R}_{-1}^{\Psi_q}, \tilde{R}_1^{\Psi_q}, \partial_x^\ell R_1)$ as in (3.95), such that

$$
\partial_t \tilde{\mathcal{E}}_\ell \lesssim \varepsilon^2\big(1 + \mathcal{E}_\ell\big)\,.
$$

□

Corollary 3.3.14 now allows us to prove theorem 1.2.1, in the same fashion we proved theorem 1.1.1 with corollary 2.2.30 in the last section.

Outlook

In this thesis, we only considered systems with quadratic quasilinear terms. It should be easy to extend our result to be valid for systems that can also have quasilinear terms of higher orders. A more difficult task would be the extension of our result to quasilinear dispersive systems that are more complicated, in the sense that their nonlinearities are of a more general form. An example for this is the class of systems (1.30)-(1.31) from chapter 3. As long as some conditions, cf. remark 2.2.19, are fulfilled one should also be able to look at such systems with arbitrarily large $\deg^*(\rho)$. Moreover, we expect that for such systems one can soften (1.9) and allow certain nonlinear terms to be stronger than the linear part of the system.

As we stated earlier, we expect that our techniques will also be useful for much more complicated quasilinear systems like the water wave problem. Especially our approach for handling quasilinear terms with arbitrarily large $\deg^*(\rho)$.

Some of the techniques we use for our error estimates here may could also be transfered in order to show the existence of long time solutions for quasilinear systems since the methods of proofs resemble each other, see [DH18]. One can most likely not handle nontrivial resonances with these techniques, but the modified energy and the energy transformations could may help one to study systems with arbitrarily large $\deg^*(\rho)$.

The techniques, we introduced in order be able to handle quasilinear nonlinearities with arbitrarily large $\deg^*(\rho)$ could maybe also be useful for the justification of other approximations, like for example the Whitham approximation.

Last but not least, we expect that in particular our energy transformations from section 2.2.3 could also be interesting for proving the local existence of solutions to quasilinear dispersive systems.

Bibliography

[A01] Agrawal, G. P.: Solitons and nonlinear wave equations. Academic Press (2001).

[A03] Ambrose, D. M.: Well-posedness of vortex sheets with surface tension. SIAM J. Math. Anal. **35**(1), 211-244 (2003).

[AS81] Ablowitz, M.J., Segur, H.: Solitons and the inverse scattering transform. In: SIAM Studies in Applied Mathematics, vol. 4. SIAM (1981).

[C17] Cummings, P.: Nonlinear Schrödinger approximations for partial differential equations with quadratic and quasilinear terms. PhD thesis, OpenBU: `https://open.bu.edu/handle/2144/24090` (2017).

[C87] Craig, W.: Nonstrictly hyperbolic nonlinear systems. Math. Ann., **277**(2), 213-232 (1987).

[CDS15] Chirilus-Bruckner, M., Düll, W.-P., Schneider, G.: NLS approximation of time oscillatory long waves for equations with quasilinear quadratic terms. Math. Nachr. **288**(2-3), 158-166 (2015).

[CS13] Chong, C., Schneider, G.: Numerical evidence for the validity of the NLS approximation in systems with a quasilinear quadratic nonlinearity. ZAMM Z. Angew. Math. Mech. **93**(9), 688-696 (2013).

[CW17] Cummings, P., Wayne, C. E.: Modified energy functionals and the NLS approximation. Discrete Contin. Dyn. Syst. **37**(3), 1295-1321 (2017).

[D12] Düll, W.-P.: Validity of the Korteweg-de Vries Approximation for the Two-Dimensional Water Wave Problem in the Arc Length Formulation. Comm. Pure Appl. Math. **65**(3), 381-429 (2012).

[D17] Düll, W.-P.: Justification of the nonlinear Schrödinger approximation for a quasilinear Klein-Gordon equation. Comm. Math. Phys. **355**(3), 1189-1207 (2017).

[D18] Düll, W.-P.: On the Mathematical Description of Time-Dependent Surface Water Waves. Jahresber. Dtsch. Math. Ver. **120** (2), 117-141 (2018).

[D19] Düll, W.-P.: Validity of the nonlinear Schrödinger approximation for the two-dimensional water wave problem with and without surface tension in the arc length formulation. arXiv:1907.03014v1 (2019).

[DH18] Düll, W.-P.; Heß, M.: Existence of long time solutions and validity of the nonlinear Schrödinger approximation for a quasilinear dispersive equation. J. Differential Equations **264**(4), 2598-2632 (2018).

[DHSZ16] Düll, W.-P., Hermann, A., Schneider, G., Zimmermann, D.: Justification of the 2D NLS equation – Quadratic resonances do not matter in case of analytic initial conditions. J. Math. Anal. Appl. **436**, 847-867 (2016).

[DS06] Düll, W.-P., Schneider, G.: Justification of the Nonlinear Schrödinger equation for a resonant Boussinesq model. Indiana Univ. Math. J. **55**(6), 1813-1834 (2006).

[DSW16] Düll, W.-P., Schneider, G., Wayne, C.E.: Justification of the Nonlinear Schrödinger equation for the evolution of gravity driven 2D surface water waves in a canal of finite depth. Arch. Rat. Mech. Anal. **220**(2), 543-602 (2016).

[HIT16] J.K.Hunter, M.Ifrim, D.Tataru, Two dimensional water waves in holomorphic coordinates, Comm. Math. Phys. **346**(2), 483-552 (2016).

[HITW15] J.K.Hunter, M.Ifrim, D.Tataru, T.K.Wong, Long Time Solutions for a Burgers-Hilbert Equation via a Modified Energy Method, Proc. Amer. Math. Soc. **143**(8), 3407-3412 (2015).

[I15] Il'ichev, A. T.: Envelope solitary waves and dark solitons at a water-ice interface. Proc. Steklov Inst. Math. **289**(1) , 152-166 (2015).

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

[K88] Kalyakin, L.A.: Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. Sb. Math. **60**, 457-483 (1988).

[K75a] Kato, T.: The Cauchy problem for quasi-linear symmetric hyperbolic systems, Arch. Rat. Mech. Anal. 58, 181-205 (1975).

[K75b] Kato, T.:Quasi-linear equations of evolution, with applications to partial differential equations. In: Everitt W.N. (eds) Spectral Theory and Differential Equations. Lecture Notes in Mathematics **448**, Springer, Berlin, Heidelberg, 25-70 (1975).

[KV19] Kolkovska, N., Vucheva, V.: Invariant preserving schemes for double dispersion equations. Adv. Difference Equ. **216** 1-16 (2019).

[KSM92] Kirrmann, P., Schneider, G., Mielke, A.: The validity of modulation equations for extended systems with cubic nonlinearities. Proc. Roy. Soc. Edinburgh Sect. A **122**, 85-91 (1992).

[LG19] Li, Y., Gao, Y.: The method of lower and upper solutions for the cantilever beam equations with fully nonlinear terms. J. Inequal. Appl. **136** 1-16 (2019).

[MN13] Masmoudi, N., Nakanishi, K.: Multifrequency NLS scaling for a model equation of gravity-capillary waves. Commun. Pure Appl. Math. **66**(8), 1202-1240 (2013).

[P11] Pelinovsky, D. E.: Localization in periodic potentials. From Schrödinger operators to the Gross-Pitaevskii equation. London Mathematical Society Lecture Note Series **390**. Cambridge: Cambridge University Press (2011).

[S98a] Schneider, G.: Justification of modulation equations for hyperbolic systems via normal forms. NoDEA Nonlinear Differential Equations Appl. **5**, 69-82 (1998).

[S98b] Schneider, G.: Approximation of the Korteweg-de Vries equation by the Nonlinear Schrödinger equation. J. Differential Equations **147**, 333-354 (1998).

[S05] Schneider, G.: Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. J. Differential Equations **216**, 354-386 (2005).

[S11a] Schneider, G.: The role of the Nonlinear Schrödinger equation in nonlinear optics. In: Oberwolfach seminars 42: Photonic Crystals: Mathematical Analysis and Numerical Approximation by Dörfler, W., Lechleiter, A., Plum, M., Schneider, G. and Wieners, C.. Birkhäuser (2011).

[S11b] Schneider, G.: Justification of the NLS approximation for the KdV equation using the Miura transformation. Advances in Mathematical Physics (2011) 854719.

[SH94] Shi, Y., Hearst, J. E.: The Kirchhoff elastic rod, the nonlinear Schrödinger equation, and DNA supercoiling. J. Chem. Phys. **101**(6), 5186-5200 (1994).

[SSZ15] Schneider, G., Sunny, D.A., Zimmermann, D.: The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions. J. Dynam. Differential Equations **27**(3), 1077-1099 (2015).

[SU17] Schneider, G., Uecker, H.: Nonlinear PDEs. A dynamical systems approach. Graduate Studies in Mathematics **182**, American Mathematical Society, Providence, RI, ISBN: 978-1-4704-3613-1 (2017)

[SW00] Schneider, G., Wayne, C.E.: The long wave limit for the water wave problem. I. the case of zero surface tension. Comm. Pure Appl. Math. **53**(12), 1475-1535 (2000).

[SW02] Schneider, G., Wayne, C.E.: The rigorous approximation of long-wavelength capillary-gravity waves. Arch. Rat. Mech. Anal. **162**, 247-285 (2002).

[SW11] Schneider, G., Wayne, C.E.: Justification of the NLS approximation for a quasilinear water wave model. J. Differential Equations **251**, 238-269 (2011).

[T15] Totz, N.: A justification of the modulation approximation to the 3d full water wave problem. Comm. Math. Phys. **335**(1), 369-443 (2015).

[TW12] Totz, N., Wu, S.: A rigorous justification of the modulation approximation to the 2D full water wave problem. Comm. Math. Phys. **310**(3), 817-883 (2012).

[WC06] Wang, S., Chen, G.: Cauchy problem of the generalized double dispersion equation. Nonlinear Anal. **64**(1), 159-173 (2006).

[Z68] Zakharov, V. E.: Stability of periodic waves of finite amplitude on the surface of a deep fluid. Journal of Applied Mechanics and Technical Physics, **9**(2), 190-194 (1968).