Validity of the Nonlinear Schrödinger Approximation for the Two-Dimensional Water Wave Problem With and Without Surface Tension in the Arc Length Formulation

Wolf-Patrick Düll

Abstract

We consider the two-dimensional water wave problem in an infinitely long canal of finite depth both with and without surface tension. In order to describe the evolution of the envelopes of small oscillating wave packet-like solutions to this problem, the Nonlinear Schrödinger equation can be derived as a formal approximation equation. In recent years, the validity of this approximation has been proven by several authors for the case without surface tension. In this paper, we rigorously justify the Nonlinear Schrödinger approximation for the cases with and without surface tension by proving error estimates over a physically relevant timespan in the arc length formulation of the two-dimensional water wave problem. The error estimates are uniform with respect to the strength of the surface tension, as the height of the wave packet and the surface tension go to zero.

1. Introduction

In this paper, we consider the two-dimensional water wave problem with finite depth of water. The two-dimensional 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. In Eulerian coordinates, the two-dimensional water wave problem with finite depth has the following form: the fluid fills a domain $\Omega(t) = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, -h < y < \eta(x, t)\}$ in between the bottom $B = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, y = -h\}$ and the free top surface $\Gamma(t) = \{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, y = \eta(x, t)\}$. The velocity field $V = (v_1, v_2)$ of the fluid is governed by the incompressible Euler’s equations

$$V_t + (V \cdot \nabla)V = -\nabla p + g \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{in } \Omega(t),$$

(1)
\[ \nabla \cdot V = 0 \quad \text{in} \quad \Omega(t), \quad (2) \]

where \( p \) is the pressure and \( g \) the constant of gravity.

Assuming that fluid particles on the top surface remain 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, yields the boundary conditions

\[ \eta_t = V \cdot \begin{pmatrix} -\eta_x \\ 1 \end{pmatrix} \quad \text{at} \quad \Gamma(t), \quad (3) \]

\[ p = -bgh^2 \kappa \quad \text{at} \quad \Gamma(t), \quad (4) \]

\[ v_2 = 0 \quad \text{at} \quad B, \quad (5) \]

where \( b \geq 0 \) is the Bond number, which is proportional to the strength of the surface tension, and \( \kappa \) is the curvature of \( \Gamma(t) \).

If the flow is additionally assumed to be irrotational, the above system can be reduced to a system defined on \( \Gamma(t) \). Due to the irrotationality of the motion there exists a velocity potential \( \phi \) with \( V = \nabla \phi \), which is harmonic in \( \Omega(t) \) with vanishing normal derivative at \( B \). Moreover, the motion of the vertical component of the velocity is uniquely determined by the horizontal one, i.e., there exists an operator \( K = K(\eta) \) such that

\[ \phi_y = K(\eta) \phi_x. \quad (6) \]

By using the potential \( \phi \), the system (1)–(5) can be reduced to

\[ \eta_t = V \cdot \begin{pmatrix} -\eta_x \\ 1 \end{pmatrix} \quad \text{at} \quad \Gamma(t), \quad (7) \]

\[ \phi_t = -\frac{1}{2}((\phi_x)^2 + (K\phi_x)^2) - g\eta + bgh^2 \left( \frac{\eta_x}{\sqrt{1 + \eta_x^2}} \right)_x \quad \text{at} \quad \Gamma(t), \quad (8) \]

or to

\[ \eta_t = K v_1 - v_1 \eta_x \quad \text{at} \quad \Gamma(t), \quad (9) \]

\[ (v_1)_t = -g\eta_x - \frac{1}{2}(v_1^2 + (Kv_1)^2)_x + bgh^2 \left( \frac{\eta_x}{\sqrt{1 + \eta_x^2}} \right)_{xx} \quad \text{at} \quad \Gamma(t). \quad (10) \]

From now on, let space and time in the above system be rescaled in such a way that \( h = 1 \) and \( g = 1 \).

Choosing Eulerian coordinates to formulate the equations for the motion of water waves is natural for describing many physical experiments, but there are also alternative coordinate systems which yield appropriate frameworks for formulating the water wave problem. Each of these coordinate systems has its own advantages concerning applicability and mathematical structure of the resulting equations of the water wave problem. Hence, depending on the problem one intends to solve, one has to find out which coordinate system is the most adapted one.
The most known alternative systems are Lagrangian coordinates, see, for example, [49], holomorphic coordinates, see, for example, [17], the arc length formulation, see, for example, [3], and abstract coordinate independent systems which are based on the fact that the solutions of (1)–(5) can be interpreted as the geodesic flow with respect to the potential energy, the kinetic energy and in case of surface tension also the surface energy on the infinite dimensional Riemannian manifold of volume-preserving homeomorphisms of $\Omega(0)$, see, for example, [43]. In this differential geometric variational framework, the boundary conditions (3)–(5) appear as natural boundary conditions.

For an overview on the local and global well-posedness results for the water wave problem in the various formulations we refer to [10] and the references therein.

Concerning the qualitative behavior of the solutions, the full water wave problem is extremely complicated to analyze. A qualitative understanding of the solutions to the full water wave problem being usable for practical applications does not seem within reach for the near future, neither analytically nor numerically. Therefore, it is important to approximate the full model in different parameter regimes by suitable reduced model equations whose solutions have similar but more easily accessible qualitative properties.

The simplest reduced model equation is the linear wave equation. The most famous nonlinear approximation equations are the Korteweg-de Vries (KdV) equation and the Nonlinear Schrödinger (NLS) equation. By inserting the ansatz

$$\left(\begin{array}{c}
\eta \\
v_1
\end{array}\right)(x, t) = \varepsilon^2 A \left(\varepsilon(x \pm t), \varepsilon^3 t\right) \left(\frac{1}{\mp 1}\right) + O(\varepsilon^3)$$

with $0 < \varepsilon \ll 1$ and $A : \mathbb{R}^2 \rightarrow \mathbb{R}$ into (9)–(10), expanding the operator $K$ with respect to $\varepsilon$ and equating the terms with the lowest power of $\varepsilon$ one obtains that $A$ has to satisfy in lowest order with respect to $\varepsilon$ the KdV equation

$$A_\tau = \pm \left(\frac{1}{6} - \frac{b}{3}\right) A_{\xi\xi\xi} \pm \frac{3}{2} AA_\xi,$$

where $\tau = \varepsilon^3 t$ and $\xi = \varepsilon(x \pm t)$, if $b \neq 1/3$. For further information about the KdV approximation we refer to [10] and the references therein.

The ansatz for the NLS approximation is

$$\left(\begin{array}{c}
\eta \\
v_1
\end{array}\right)(x, t) = \varepsilon A \left(\varepsilon(x - cgt), \varepsilon^2 t\right) e^{i(k_0 x - \omega_0 t)} \varphi(k_0, b) + O(\varepsilon^2) + c.c.,$$

where $0 < \varepsilon \ll 1$. Here $\omega_0 > 0$ is the basic temporal wave number associated via the linear dispersion relation of the two-dimensional water wave problem with finite depth, namely

$$\omega(k) = \omega(k, b) = \pm \sqrt{(k + bk^3) \tanh(k)},$$

to the basic spatial wave number $k_0 > 0$ of the underlying carrier wave $e^{i(k_0 x - \omega_0 t)}$, that means that $\omega_0 = \omega(k_0)$, where the branch of solutions

$$\omega(k) = \omega(k, b) = \text{sgn}(k) \sqrt{(k + bk^3) \tanh(k)}$$

(12)
Fig. 1. An envelope (advancing with the group velocity \(c_g\)) with characteristic length scale of order \(O(1/\varepsilon)\) of an oscillating wave packet \(\eta\) of order \(O(\varepsilon)\) (advancing with the phase velocity \(c_p = \omega_0/k_0\)) is approximately described by the amplitude \(A\) which solves the NLS equation \(\text{(13)}\)

is chosen. Moreover, \(c_g\) is the group velocity, i.e., \(c_g = \omega'(k_0) = \partial_k \omega(k_0, b)\), \(A\) the complex-valued amplitude, \(\varphi(k_0, b) \in \mathbb{C}^2\) and c.c. the complex conjugate. This ansatz leads to waves moving to the right; to obtain waves moving to the left, \(\omega_0\) and \(c_g\) have to be replaced by \(-\omega_0\) and \(-c_g\).

By inserting this ansatz into \((9)-(10)\), one obtains that for an explicitly computable vector \(\varphi(k_0, b)\) the amplitude \(A\) has to satisfy at leading order in \(\varepsilon\) the NLS equation

\[
A \tau = i \frac{\partial^2 \omega(k_0, b)}{2} A_{\xi \xi} + i \nu(k_0, b) A |A|^2, \tag{13}
\]

where \(\tau = \varepsilon^2 t, \xi = \varepsilon(x - c_g t)\) and \(\nu(k_0, b) \in \mathbb{R}\). Hence, the NLS equation \((13)\) approximately describes the dynamics of spatially and temporarily oscillating wave packet-like solutions to the two-dimensional water wave problem; see Fig. 1.

In one space dimension, both the KdV equation and the NLS equation are completely integrable Hamiltonian systems which can be explicitly solved with the help of inverse scattering schemes; see, for example, [1].

The first formal derivation of the NLS approximation for the two-dimensional water wave problem was given by Zakharov [50] in 1968. The NLS approximation is used, for example, in the context of modeling monster waves; see [27]. However, the NLS approximation plays not only an important role for the mathematical description of surface water waves but also in other areas of science and technology, for example, in nonlinear optics to model data transmission via fiber optic cables with the help of light pulses [1, 35], in biology to model waves in DNA [19], in plasma physics [42] or in quantum theory [28]. In numerical simulations, the simulation of the evolution of the envelope with the help of the NLS approximation yields a significant reduction of complexity and consequently an increase of efficiency compared to the simulation of the whole wave packet.

Although the NLS approximation is very successful in many applications, it should not be taken for granted that the NLS approximation always yields correct predictions of the behavior of the original system. Indeed, there are counterexamples where the NLS approximation fails [34, 37]. Hence, it is important to answer the
question of the validity of the NLS approximation for a given system by proving error estimates over a physically relevant timespan. In general, this is a highly nontrivial mathematical problem for a variety of reasons.

Given the general abstract evolutionary problem

$$\partial_t W = \mathcal{L} W + \mathfrak{B}(W, W) + \mathfrak{S}(W),$$  \hspace{1cm} (14)

with \( x, t \in \mathbb{R} \) and \( W = (W_1(x, t) \ W_2(x, t))^T \in \mathbb{R}^2 \), \( \mathcal{L} \) is a linear operator whose symbol is a diagonal matrix of the form

$$\hat{\mathcal{L}}(k) = \text{diag} \ (-i \omega(k), i \omega(k)), \hspace{1cm} (15)$$

where \( k \in \mathbb{R} \) and \( \omega \) is a piecewise smooth real-valued odd function. Furthermore, \( \mathfrak{B} \) is a bilinear operator and \( \mathfrak{S}(W) \) consists of terms being at least cubic in \( W \) or is equal to 0.

The NLS equation (13) can be derived as a formal approximation equation with the help of the ansatz

$$W = \varepsilon \tilde{\Psi},$$

where

$$\varepsilon \tilde{\Psi}(x, t) = \varepsilon \Psi_{NLS}(x, t) \left( \begin{array}{c} 1 \\ 0 \end{array} \right) + \varepsilon^2 \Psi_h(x, t), \hspace{1cm} (16)$$

$$\Psi_{NLS}(x, t) = A(\varepsilon(x - c_g t), \varepsilon^2 t)e^{i(k_0 x - \omega_0 t)} + \text{c.c.}, \hspace{1cm} (17)$$

$$\Psi_h(x, t) = \left( \begin{array}{c} \tilde{A}_{01}(\varepsilon(x - c_g t), \varepsilon^2 t) \\ \tilde{A}_{02}(\varepsilon(x - c_g t), \varepsilon^2 t) \end{array} \right) + \left( \begin{array}{c} \tilde{A}_{21}(\varepsilon(x - c_g t), \varepsilon^2 t) \\ \tilde{A}_{22}(\varepsilon(x - c_g t), \varepsilon^2 t) \end{array} \right)e^{2i(k_0 x - \omega_0 t)} + \text{c.c.}, \hspace{1cm} (18)$$

\( k_0 > 0, \omega_0 = \omega(k_0), c_g = \omega'(k_0), \tilde{A}_{01}, \tilde{A}_{02} \) are real-valued functions and \( \tilde{A}_{21}, \tilde{A}_{22} \) complex-valued functions.

Inserting this ansatz into (14) and equating the coefficients in front of the \( \varepsilon^m e^{i(k_0 x - \omega_0 t)} \) for \( m \in \{1, 2, 3\} \) and \( j \in \{0, 1, 2\} \) to 0 yields the NLS equation

$$A_\tau = i \frac{\omega''(k_0)}{2} A_{\xi \xi} + i \nu(k_0)|A|^2, \hspace{1cm} (19)$$

where \( \tau = \varepsilon^2 t, \xi = \varepsilon(x - c_g t) \) and \( \nu(k_0) \in \mathbb{R} \), if \( \omega \) satisfies

$$\lim_{k \to 0^\pm} \omega(k) \neq 0 \hspace{1cm} (20)$$

or

$$\lim_{k \to 0^\pm} \omega'(k) \neq \omega'(k_0) \hspace{1cm} (21)$$

as well as

$$\pm \omega(2k_0) \neq 2\omega(k_0) \hspace{1cm} (22)$$

and

$$\omega''(k_0) \neq 0. \hspace{1cm} (23)$$

The above ansatz leads to wave packets moving to the right; to obtain wave packets moving to the left, \( \omega_0, c_g \) have to be replaced by \(-\omega_0, -c_g\) and \( (\Psi_{NLS}(x, t) \ 0)^T \) by \( (0 \ \Psi_{NLS}(x, t))^T \).
It is possible to modify $\tilde{\Psi}$ to make it an even more accurate approximation. Indeed, if there exists an integer $M > 2$ such that
\begin{equation}
\pm \omega(mk_0) \neq m\omega(k_0) \quad (24)
\end{equation}
for all integers $m \in [2, M)$, then there exists a function $\Psi$ dependent on $\varepsilon$ such that
\begin{equation}
\lim_{\varepsilon \to 0} \| \Psi(\cdot, t) - \tilde{\Psi}(\cdot, t) \|_{C^0} = 0 \quad (25)
\end{equation}
and
\begin{equation}
\text{Res}(\varepsilon \Psi) := -\partial_t (\varepsilon \Psi) + \varepsilon \mathcal{L}\Psi + \mathcal{B}(\varepsilon \Psi, \varepsilon \Psi) + \mathcal{S}(\varepsilon \Psi) = \mathcal{O}(\varepsilon^M). \quad (26)
\end{equation}

The two-dimensional water wave problem with finite depth can be transformed to an evolutionary system of the form (14) by diagonalizing the linear part of the system (9)–(10). More precisely, if one makes the linear coordinate transform
\begin{equation}
W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} := \frac{1}{2} \begin{pmatrix} \sigma & 1 \\ -\sigma & 1 \end{pmatrix} \begin{pmatrix} y \\ v_1 \end{pmatrix}, \quad (27)
\end{equation}
where $\sigma$ is a linear operator defined by its symbol
\begin{equation}
\sigma(k) = \sigma(k, b) = \sqrt{k + bk^3 \tanh(k)} \quad (28)
\end{equation}
for $k \in \mathbb{R}$, then $W$ satisfies a system of the form (14) with $\omega$ defined by (12).

For $b = 0$, the dispersion relation $\omega$ satisfies (21)–(24) for all $k_0 > 0$. For $b > 0$, one has to choose a basic wave number $k_0 > 0$ for which the conditions (21)–(24) are valid in order to be able to derive a sufficiently accurate NLS approximation. Or, if $k_0 > 0$ is given, then one has to choose those values of $b \geq 0$ for which (21)–(24) are valid in order to be able to derive a sufficiently accurate NLS approximation.

To guarantee that qualitative properties of solutions to the NLS equation (19) are also true for solutions to system (14), it has to be proven that the error
\begin{equation}
\varepsilon^\beta R := W - \varepsilon \Psi \quad (29)
\end{equation}
is of order $\mathcal{O}(\varepsilon^\beta)$ with $\beta > 1$ on the characteristic time scale of the NLS equation (19), this means that there exists a $\tau_0 > 0$ such that $R$ is of order $\mathcal{O}(1)$ for all $t \in [0, \tau_0/\varepsilon^2]$. The rescaled error $R$ satisfies for appropriately chosen $\beta$ and $\Psi$ an evolution equation of the form
\begin{equation}
\partial_t R = \mathcal{L}R + \varepsilon B(\Psi, R) + \mathcal{O}(\varepsilon^2), \quad (30)
\end{equation}
where
\begin{equation}
B(\Psi, R) = \mathcal{B}(\Psi, R) + \mathcal{B}(R, \Psi). \quad (31)
\end{equation}

Since the Fourier transform $\varepsilon \hat{\Psi}_{NLS}$ of $\varepsilon \hat{\Psi}_{NLS}$ is strongly concentrated around the wave numbers $\pm k_0$, the approximation $\varepsilon \Psi$ can be split into
\begin{equation}
\varepsilon \Psi = \varepsilon \Psi_c + \varepsilon^2 \Psi_s = \varepsilon \left( \begin{pmatrix} \psi_1 \\ 0 \end{pmatrix} \right) + \varepsilon \left( \begin{pmatrix} \psi_{-1} \\ 0 \end{pmatrix} \right) + \varepsilon^2 \Psi_s, \quad (32)
\end{equation}
where the supports of $\hat{\psi}_{\pm 1}$ satisfy
\[ \text{supp} \hat{\psi}_{\pm 1} = \{ k \in \mathbb{R} : |k \mp k_0| \leq \delta \} \] (33)
for a $\delta \in (0, k_0)$ sufficiently small, but independent of $\varepsilon$, and $\Psi_s$ is of order $\mathcal{O}(1)$. Consequently, we have
\[ \partial_t R = \mathcal{L} R + \varepsilon B(\Psi_c, R) + \mathcal{O}(\varepsilon^2). \] (34)
Hence, the main difficulty is to control the quadratic term $\varepsilon B(\Psi_c, R)$ over a timespan of order $\mathcal{O}(\varepsilon^{-2})$.

A classical strategy is to eliminate the quadratic term with the help of a so-called normal-form transform
\[ \tilde{R} := R + \varepsilon N(\Psi_c, R), \] (35)
where $N$ is an appropriate bilinear mapping, which can be constructed with the help of the Fourier transform. More precisely, let
\[ \hat{B}_{ji}(\Psi_c, R)(k) = \int_{\mathbb{R}} \sum_{\ell \in \{\pm 1\}, j_2 \in \{1, 2\}} \hat{b}_{ji j_2}(k, k - m, m) \hat{\psi}_\ell(k - m) \hat{R}_{j_2}(m) dm \] (36)
and
\[ \hat{N}_{ji}(\Psi_c, R)(k) = \int_{\mathbb{R}} \sum_{\ell \in \{\pm 1\}, j_2 \in \{1, 2\}} \hat{n}_{ji j_2 \ell}(k, k - m, m) \hat{\psi}_\ell(k - m) \hat{R}_{j_2}(m) dm, \] (37)
where $j_1, j_2 \in \{1, 2\}$ denote the components of the vectors $\hat{B}, \hat{N}$ and $\hat{R}$. Then, by inserting (35) into (34), one obtains that $\tilde{R}$ solves an evolution equation of the form
\[ \partial_t \tilde{R} = \mathcal{L} \tilde{R} + \varepsilon^2 g(\Psi_c, \tilde{R}) + \mathcal{O}(\varepsilon^2), \] (38)
where $g$ is of order $\mathcal{O}(1)$, if
\[ \hat{n}_{ji j_2 \ell}(k, k - m, m) = \frac{\hat{b}_{ji j_2}(k, k - m, m)}{i(j_1 \omega(k) + \omega(k - m) - j_2 \omega(m))} \] (39)
and if the normal-form transform $R \mapsto \tilde{R}$ is invertible. Furthermore, due to (33), it turns out that it is even possible to simplify the kernels $\hat{n}_{ji j_2 \ell}$ to
\[ \hat{n}_{ji j_2 \ell}(k) = \frac{\hat{b}_{ji j_2}(k, \ell k_0, k - \ell k_0)}{i(j_1 \omega(k) + \omega(\ell k_0) - j_2 \omega(k - \ell k_0))}. \] (40)

The strategy of using normal-form transforms to eliminate semilinear quadratic terms in hyperbolic systems was introduced in [41]. In the context of justifying NLS approximations, it was first applied in [29].

However, there are serious difficulties. The first one is the possible occurrence of resonances. This means that the denominator of the fraction in (40) may have zeros, the so-called resonances or resonant wave numbers (to the wave number $\ell k_0$). Since $\omega$ is odd, any resonance implies further resonances. Namely, if $k$ is
resonant to $\ell k_0$, then $-k$ is resonant to $-\ell k_0$. Moreover, if $k$ is resonant to $\ell k_0$ and $j_1 = j_2$, then $\pm (k - \ell k_0)$ is resonant to $\mp \ell k_0$.

In the case of the two-dimensional water wave problem with finite depth, there is a resonance at $k = 0$, but the numerator of the fraction in (40) also vanishes at $k = 0$ and the singularity is removable. Such a resonance is called trivial. Otherwise it is called non-trivial. The resonance at $k = 0$ implies resonances at $k = \pm k_0$, which are non-trivial. Moreover, for all basic wave numbers $k_0 > 0$ there exist some $b \in (0, 1/3)$ such that there are additional non-trivial resonances for $j_1 = j_2 = -1$.

In the context of the justification of the NLS approximation for an evolutionary system of the form (14) with resonances at 0 and $\ell k_0$, it is relevant if the wave numbers $\ell k_0$ are stable, this means that for any wave number $\ell k_1 \in \mathbb{R} \setminus \{0, \ell k_0\}$ being a non-trivial resonance with respect to $\ell k_0$ for $j_1 = j_2 = -1$ the NLS subspace in the Three Wave Interaction (TWI) system associated to the wave numbers $k_{0\ell} := -\ell k_0, k_{1\ell} := \ell k_1$ and $k_{2\ell} := -\ell (k_1 - k_0)$, which then satisfy

$$k_{0\ell} + k_{1\ell} + k_{2\ell} = 0, \quad \omega(k_{0\ell}) + \omega(k_{1\ell}) + \omega(k_{2\ell}) = 0, \quad (41)$$

is stable. More precisely, inserting the ansatz

$$W_{\ell}(x, t) := \sum_{j=0}^{2} A_{j\ell}(\varepsilon t) e^{i(k_{j\ell} x - \omega(k_{j\ell}) t)} \left( \begin{array}{c} 1 \\ 0 \end{array} \right) + c.c., \quad (42)$$

where $0 < \varepsilon \ll 1$, in (14) and equating the coefficients of $\varepsilon^2 e^{i(k_{j\ell} x - \omega(k_{j\ell}) t)}$ for $j \in \{0, 1, 2\}$ to zero yields the so-called TWI system

$$\begin{align*}
\partial_x A_{0\ell} &= \hat{b}_{11}(-\ell k_0, -\ell k_1, \ell k_1, \ell k_0) A_{1\ell} A_{2\ell}, \\
\partial_x A_{1\ell} &= \hat{b}_{11}(\ell k_1, \ell k_0, \ell k_1, \ell k_0) A_{0\ell} A_{2\ell}, \\
\partial_x A_{2\ell} &= \hat{b}_{11}(-\ell k_1 + \ell k_0, \ell k_0, 0) A_{0\ell} A_{1\ell} \quad (43)
\end{align*}$$

with $\tau = \varepsilon t$ and $\hat{b}_{11}$ as in (36). This system has three invariant subspaces consisting of fixed points, namely $M_{0\ell} = \{A_{1\ell} = A_{2\ell} = 0\}$, $M_{1\ell} = \{A_{0\ell} = A_{2\ell} = 0\}$ and $M_{2\ell} = \{A_{0\ell} = A_{1\ell} = 0\}$. The so-called NLS subspace $M_{0\ell}$ is stable if and only if

$$\hat{b}_{11}(\ell k_1, \ell k_0, \ell k_1, \ell k_0) < 0, \quad (44)$$

and then

$$E := |A_{1\ell}|^2 - \frac{\hat{b}_{11}(\ell k_1, \ell k_0, \ell k_1, \ell k_0)}{\hat{b}_{11}(-\ell k_1 + \ell k_0, \ell k_0, 0)} |A_{2\ell}|^2 \quad (45)$$

is a non-negative conserved quantity of the system (43); see [34]. Since (14) is a real-valued system, the wave number $k_0$ is stable if and only if $-k_0$ is stable. If $k_0$ is unstable, then the corresponding NLS approximation can fail under certain conditions; see [34,37].

For the two-dimensional water wave problem with finite depth, the values of the coefficients in the corresponding TWI system (43) can be computed explicitly with the help of the method from [37]. It turns out that $k_0$ is stable if and only if
Fig. 2. The graph of $k \mapsto \omega(k, b) - \omega(k - k_0, b) - \omega(k_0, b)$ for $k \geq 0$, $k_0 = 2$ and different values of $b$. From top left to the bottom right: Panel (i) $b = 1/3$, (ii) $b = 1/3.5$, (iii) $b = 1/4.15$, (iv) $b = b_1(2) = 0.2396825654$, (v) $b = 1/4.25$, (vi) $b = b_0(2) = 0.2240838469$, (vii) $b = 1/5$, (viii) $b = 1/200$, (ix) $b = 0$

$k_0 < \max \{k_1, k_0 - k_1\}$ for all $k_1 \in \mathbb{R} \setminus \{0, k_0\}$ being a non-trivial resonance with respect to $k_0$ for $j_1 = j_2 = -1$.

For any $k_0 > 0$ all additional non-trivial resonances and all values of $b$ for which $k_0$ is stable can be determined by analyzing the zeros of the function $\tilde{\omega}(k, b)$ with $\tilde{\omega}(k, b) = \omega(k, b) - \omega(k - k_0, b) - \omega(k_0, b)$ on $[k_0/2, \infty) \times \mathbb{R}_0^+$ and using the symmetry of $\omega$ as discussed above. It turns out that for all $k_0 > 0$ there exist a smallest $b_1 = b_1(k_0) \in (0, 1/3)$, a largest $b_0 = b_0(k_0) \in (0, b_1)$ and a strictly monotonically decreasing function $k_1 \in C^0((0, b_0))$ with $k_1(b) > k_0$ for all $b \in (0, b_0)$ and $k_1(b) \to \infty$ for $b \to 0$ such that $\tilde{\omega}$ has no other zeros than $(k_0, b)$ if $b \in [0] \cup (b_1, \infty)$ and exactly two zeros $(k_0, b)$ and $(k_1(b), b)$ if $b \in (0, b_0)$; see Fig. 2.

Another difficulty is the fact that the normal-form transform $R \mapsto R + \varepsilon N(\Psi, R)$ may lose regularity, this means that it maps the Sobolev space $H^n(\mathbb{R}, \mathbb{C})$ into $H^{n-j}(\mathbb{R}, \mathbb{C})$ for $j \in (0, n]$. A loss of regularity happens, for example, if $B(W, W)$
contains quasilinear terms. A normal-form transform losing regularity may not be invertible. But even if it was invertible, the mapping $\tilde{R} \mapsto \varepsilon^2 g(\Psi, \tilde{R})$ would lose even more regularity such that it would not be possible in general to derive estimates for $\tilde{R}$ directly from equation (38) - for example, by applying the variation of constants formula and Gronwall’s inequality.

For these reasons the validity of the NLS approximation for systems with quasilinear quadratic terms is a highly non-trivial problem, which remained unsolved in general for more than four decades. The first validity theorems for the NLS approximation in the literature were proven only for systems with special structural properties.

In [29], the NLS approximation was justified for quasilinear systems without quasilinear quadratic terms. Moreover, semilinear quadratic terms were only admitted if they cause no resonances or trivial resonances. In this situation, the method of normal-form transforms discussed above can be successfully used.

In [13, 32–34], the method of normal-form transforms was further developed to make it applicable to systems with non-trivial resonances at $k = \pm k_0$, additional non-trivial resonances with the property that the NLS subspace is stable with respect to those resonances or in case of analytic initial data also additional non-trivial resonances with the property that the NLS subspace is unstable with respect to those resonances.

In [40], the validity of the NLS approximation was obtained for a quasilinear reduced model of the two-dimensional water wave problem with finite depth and without surface tension. This reduced model shares with the Lagrangian formulation of the two-dimensional water wave problem some of the principal difficulties which have to be overcome for a validity proof for the NLS approximation, for example the fact that the quadratic nonlinearity loses regularity of half a derivative. In this case the elimination of the quadratic terms is possible with the help of normal-form transforms. The cubic nonlinearity of the transformed system then lose one derivative and can be handled by using a Cauchy-Kowalevskaya argument.

For the quasilinear KdV equation, the NLS approximation was justified by simply applying a Miura transform [36]. Another approach to address the problem of the validity of the NLS approximation for a dispersive equation can be found in [31]. In [4], the NLS approximation of time oscillatory long wave solutions to wave equations with quasilinear quadratic terms was justified. Because of the scaling behavior of the long wave solutions it is not necessary to eliminate the quadratic terms such that a normal-form transform is not needed. In [11], it was proven that analytic solutions of a two-dimensional wave equation with a quadratic nonlinearity can be approximated with the help of a two-dimensional NLS equation if the set of resonances is separated from the set of integer multiples of the basic wave vector $k_0 \in \mathbb{R}^2$ of the underlying carrier wave.

In [9], the NLS approximation was justified for a nonlinear Klein-Gordon equation with a quasilinear quadratic term, which is the first validity proof of the NLS approximation of a nonlinear hyperbolic equation with a quasilinear quadratic term losing regularity of more than half a derivative by error estimates in Sobolev spaces. The linear dispersion relation of the Klein-Gordon equation causes no resonances. The loss of regularity is overcome by using the so-called modified energy method.
The main idea of this method is as follows. Instead of performing the normal-form transform (35) explicitly and estimating the transformed error $\tilde{R}$ the normal-form transform is only used to construct an energy $E_s$ which is an appropriate adaptation of

$$E_s = \sum_{l=0}^{s} \sum_{j=1}^{2} \left( \frac{1}{2} \| \partial_x^l R_j \|_{L^2}^2 + \epsilon \int_R \partial_x^l R_j \partial_x^l N_j(\Psi_c, R) \, dx \right)$$

for a sufficiently large $s > 0$. Since $E_s$ differs from $\| \tilde{R} \|_{H^s}^2$ only by terms of order $O(\epsilon^2)$, the evolution equations of $E_s$ and $\| \tilde{R} \|_{H^s}^2$ share the property that their right-hand sides are of order $O(\epsilon^2)$. The energy $E_s$ has the advantage that in the case of a normal-form transform which loses regularity the right-hand side of the evolution equation of $E_s$ has better regularity properties than the right-hand side of the evolution equation of $\| \tilde{R} \|_{H^s}^2$.

An early version of a modified energy can be found in [8] as an ingredient to simplify and generalize the proof of the error estimates for the KdV approximation of the water wave problem compared with the alternative proofs in [38, 39].

The first modified energy which was used to overcome regularity problems in quasilinear equations was constructed in [21]. The modified energy from this article was further developed in [17, 20, 22, 23] to apply it to prove large time and global existence results for the water wave problem in holomorphic coordinates.

A similar modified energy as in [9] was constructed in [12] to justify the NLS approximation for a quasilinear equation whose linear dispersion relation causes resonances. In [6], another modified energy was introduced to improve the NLS approximation result from [40].

In [18], the modified energies from [9, 12] were combined and extended to prove the validity of the NLS approximation for two further quasilinear quadratic dispersive systems. One system is a reduced model of the two-dimensional water wave problem with finite depth and $b \geq 1/3$, which shares with the arc length formulation of the two-dimensional water wave problem some of the principal difficulties which have to be overcome for a validity proof for the NLS approximation, for example the fact that the nonlinearity loses regularity of one derivative. The other system is the first dispersive system containing a quasilinear quadratic nonlinearity that loses regularity of $m$ derivatives with an arbitrary $m > 0$ for which the NLS approximation was justified.

For the water wave problem, all justification results for the NLS approximation in the previous literature are restricted to the case without surface tension. For the two-dimensional water wave problem with infinite depth and without surface tension in Lagrangian coordinates, the NLS approximation was justified in [45] by finding an alternative kind of a transform adapted to the special structure of that problem. For the three-dimensional water wave problem with infinite depth and without surface tension in Lagrangian coordinates, the two-dimensional NLS approximation was justified in [44] in an analogous way.

In [14], the validity of the NLS approximation was proven for the two-dimensional water wave problem with finite depth and without surface tension in Lagrangian coordinates. In these coordinates, the evolutionary system has a quasilinear quadratic nonlinearity losing regularity of only half a derivative in the case without surface
tension. The occurring resonances are handled with the help of the same strategy as in [13]. Despite the loss of regularity the normal-form transform can be inverted, which is proven by interpreting the normal-form transform as a system of differential equations whose solvability is obtained with the help of appropriate a priori estimates in Sobolev spaces. The loss of one derivative in the evolutionary system for the transformed error can be handled with the help of the same Cauchy-Kowalevskaya argument as in [40].

In [24], the NLS approximation for the two-dimensional water wave problem with infinite depth and without surface tension in holomorphic coordinates was justified by using the modified energy method.

In the present paper, we solve the open problem of justifying the NLS approximation for the full two-dimensional water wave problem with finite depth and with surface tension. Our approximation result is valid both for the case without surface tension and for the case with surface tension if there are no other non-trivial resonances than \( \pm k_0 \) or \( k_0 \) is stable. Our error estimates are uniform with respect to the strength of the surface tension as the height of the wave packet and the surface tension go to zero. We prove the following:

**Theorem 1.1.** Let \( \omega \) be the dispersion relation (12) of the two-dimensional water wave problem (9)–(10). Moreover, let \( k_0 > 0 \) and \( s \geq 10 \). Then there exist \( b_0, b_1 \in \mathbb{R} \) with \( 0 < b_0 < b_1 < 1/3 \) such that the following holds: for all \( \tau_0, C_0 > 0 \) there exist an \( \varepsilon_0 > 0 \) and a function \( C \in C^0(\mathcal{B}, \mathbb{R}^+) \), where \( \mathcal{B} \) is the set of all \( b \in \mathbb{R}_0^+ \setminus [b_0, b_1] \) for which \( k \mapsto \omega(k, b) \) satisfies (21), (23) and (24) with \( M = 6 \), such that for all \( b \in \mathcal{B} \), all solutions \( A \in C^0([0, \tau_0], H^s(\mathbb{R}, \mathbb{C})) \) of the NLS equation (13) with

\[
\sup_{\tau \in [0, \tau_0]} \| A(\cdot, \tau) \|_{H^s(\mathbb{R}, \mathbb{C})} \leq C_0
\]

and all \( \varepsilon \in (0, \varepsilon_0) \), there exists a solution

\[
(\eta, v_1) \in C^0([0, \tau_0\varepsilon^{-2}], (H^s(\mathbb{R}, \mathbb{R}))^2)
\]

of (9)–(10) which satisfies

\[
\sup_{t \in [0, \tau_0\varepsilon^{-2}]} \left\| \begin{pmatrix} \eta \\ v_1 \end{pmatrix}(\cdot, t) - \varepsilon \Psi_{NLS}(\cdot, t) \varphi(k_0, b) \right\|_{(H^s(\mathbb{R}, \mathbb{R}))^2} \leq C(b) \varepsilon^{3/2}, \quad (47)
\]

where

\[
\Psi_{NLS}(x, t) = A(\varepsilon(x - \partial_k \omega(k_0, b)t), \varepsilon^2 t)e^{i(k_0 x - \omega(k_0, b)t)} + c.c.,
\]

and \( \varphi(k_0, b) \in \mathbb{R}^2 \) is an explicitly computable vector. In particular, the error estimate (47) is uniform with respect to \( b \) as \( b \) and \( \varepsilon \) go to zero.

The error of order \( \mathcal{O}(\varepsilon^{3/2}) \) is small compared with the solution \( (\eta, v_1) \) and the approximation \( \varepsilon \Psi_{NLS} \), which are both of order \( \mathcal{O}(\varepsilon) \) in \( L^\infty \) such that the dynamics of the NLS equation can be found in the two-dimensional water wave problem, too. Our theorem guarantees that, for instance, parts of the dynamics of time-periodic
solitary wave solutions present in the NLS equation for $\partial^2_k \omega(k_0, b)$ and $\nu(k_0, b)$ having the same sign can be found approximately in the water wave problem. For a discussion of the values of $\nu(k_0, b)$ in (13), see also [1, Figure 4.15, p. 321].

It should be noted that the smoothness in our error bound is equal to the assumed smoothness of the amplitude. This can be achieved by using a modified approximation which has compact support in Fourier space but differs only slightly from $\epsilon \Psi_{NLS}$. Such an approximation can be constructed because the Fourier transform of $\epsilon \Psi_{NLS}$ is sufficiently strongly concentrated around the wave numbers $\pm k_0$.

The constants $b_0$ and $b_1$ from Theorem 1.1 can be chosen in such a way that the following holds: $b_1$ is the smallest number such that for all $b \in (b_1, \infty)$ there are no other non-trivial resonances than $\pm k_0$ and $b_0$ is the largest number such that $k_0$ is stable for all $b \in (0, b_0)$. For $k_0 = 2$ this choice of $b_0$ and $b_1$ is presented in Fig. 2. One can see that the length of the interval $[b_0, b_1]$, which contains all values of $b$ for which $k_0 = 2$ is unstable and therefore the validity of the NLS approximation cannot be expected for all sufficiently small initial data in the Sobolev space $H^s(\mathbb{R}, \mathbb{C})$, is very small. The same is true for the corresponding interval for any other $k_0 > 0$. Moreover, for all $k_0 > 0$ the number of values of $b$ for which the corresponding dispersion relation $k \mapsto \omega(k, b)$ does not satisfy (21), (23) and (24) with $M = 6$ is finite.

Now, we explain the main ideas of the proof of Theorem 1.1 and the plan of the paper. Like in many other proofs of related estimates in the literature we will assume in our proof that $s$ is an integer in order to simplify the analysis by using Leibniz’s rule, but our proofs can be generalized to be valid for all $s \geq 10$.

We perform our proof in the arc length formulation of the two-dimensional water wave problem. The main advantage of this formulation is that in the corresponding evolutionary system the surface tension dependent term with the most derivatives is linear, which allows us to prove the desired uniform error estimates. Transferring the estimates into Eulerian coordinates, we do not lose powers of $\epsilon$ since in the scaling regime of the NLS equation, the coordinates of the free surface in arc length parametrization are very close to Eulerian coordinates. The same advantages have already been used in the proof of the validity of the KdV approximation for the two-dimensional water wave problem in the arc length formulation in [8].

In Section 2 we review the arc length formulation and identify the linear terms, the quadratic terms and the terms losing regularity in the corresponding evolutionary system. Then we diagonalize the linear part of the system to obtain a system which has the structure of (14). In Section 3 we present the formal derivation of the NLS approximation for this system. Section 4 is devoted to the error estimates.

In order to perform the error estimates we use the modified energy method. The modified energy we construct is a subtle generalization of the energies in [9,12]. The normal-form transform behind our energy is an extension of a normal-form transform of the form (35) and (40) in order to handle the non-trivial resonances.

The problems with the resonances at $\pm k_0$ are circumvented by rescaling the error in Fourier space as in [13,14,40] with the help of the weight function

$$\hat{\vartheta}(k) = \begin{cases} 1, & |k| > \delta_0, \\ \epsilon + (1 - \epsilon)|k|/\delta_0, & |k| \leq \delta_0. \end{cases}$$
with a $\delta_0 = \delta_0(b) \in (0, k_0/20)$ sufficiently small, but independent of $\varepsilon$. The choice of the weight function makes sense because the quadratic terms in the evolutionary system of the two-dimensional water wave problem in the arc length formulation vanish at $k = 0$ such that the Fourier transform of the error can grow only slowly for $|k| \ll 1$. But since $\widehat{\varrho}(k) = O(\varepsilon^{-1})$ for $|k| \leq \delta_0$, the normal-form transform has to be extended by an appropriately chosen trilinear mapping.

To control the additional non-trivial resonances the terms of the form (46) in our energy are slightly modified by weight functions and correction functions similar to those in the final energy in [13], which are motivated by the conserved quantity (45).

Due to the structure of the evolutionary system of the two-dimensional water wave problem in the arc length formulation all terms on the right-hand side of the evolution equation of our energy can either directly be estimated by the energy or be identified as time derivatives of time dependent integrals. By adding these integrals to the energy we obtain our final energy, which can be bounded with the help of Gronwall’s inequality over the desired timespan of order $O(\varepsilon^{-2})$. Since this energy controls a Sobolev norm of the error, we finally obtain our approximation result.

The methods of proof developed in the present paper can also be used to prove the validity of the NLS approximation for other dispersive systems with quasilinear quadratic terms.

**Notation.** We denote the Fourier transform of a function $u \in L^2(\mathbb{R}, \mathbb{K})$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ by

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

Let $H^s(\mathbb{R}, \mathbb{K})$ be the space of functions mapping from $\mathbb{R}$ into $\mathbb{K}$ for which the norm

$$\|u\|_{H^s(\mathbb{R}, \mathbb{K})} = \left(\int_{\mathbb{R}} |\hat{u}(k)|^2 (1 + |k|^2)^s \, dk\right)^{1/2}$$

is finite. We also write $L^2$ and $H^s$ instead of $L^2(\mathbb{R}, \mathbb{R})$ and $H^s(\mathbb{R}, \mathbb{R})$. Moreover, we use the space $L^p(m)(\mathbb{R}, \mathbb{K})$ defined by $u \in L^p(m)(\mathbb{R}, \mathbb{K}) \iff u\sigma^m \in L^p(\mathbb{R}, \mathbb{K})$, where $\sigma(x) = (1 + x^2)^{1/2}$.

Furthermore, we write $A \lesssim B$ if $A \leq CB$ for a constant $C > 0$ which does not depend on $A$ and $B$, as well as $A = O(B)$ if $|A| \lesssim B$.

**2. The Water Wave Problem in the Arc Length Formulation**

In what follows we review the essential points of the arc length formulation of the two-dimensional water wave problem with finite depth. Let $P(t) : \mathbb{R} \rightarrow \Gamma(t), \alpha \mapsto P(\alpha, t) = (x(\alpha, t), y(\alpha, t))$ be a parametrization of the free top surface $\Gamma(t)$ by arc length, that means we have

$$(x_\alpha^2 + y_\alpha^2)^{1/2} = 1.$$  

(48)
Let \( U \) and \( T \) be the normal and the tangential velocity on the free top surface measured in the coordinates of the arc length parametrization, meaning that

\[
(x, y)_t(\alpha, t) = U(\alpha, t) \hat{n}(\alpha, t) + T(\alpha, t) \hat{t}(\alpha, t),
\]

where \( \hat{n} = (-\sin \theta, \cos \theta) \) and \( \hat{t} = (\cos \theta, \sin \theta) \) are the upward unit normal vectors and the unit tangential vectors to the free top surface and \( \theta = \arctan(y_\alpha/x_\alpha) \) are the tangent angles on the free top surface. Because of (48), \( T \) satisfies

\[
T_\alpha - \theta_\alpha U = 0.
\]

Integrating this relation determines \( T \) depending on \( \theta \) and \( U \) up to a constant. Since arc length parametrizations are invariant under translations, this constant can be set to 0 without loss of generality. This implies

\[
T(\alpha, t) = \int_{-\infty}^\alpha \theta_\alpha(\beta, t) U(\beta, t) \, d\beta.
\]

The normal velocity \( U \) is governed by the incompressible Euler’s equations (1)–(2), the boundary conditions (3)–(5) and the form of the free top surface.

From now on, we consider irrotational flows. Then the normal velocity \( U \) can be expressed in terms of the free top surface and the physical tangential velocity \( v \) of the fluid particles on the free top surface, where the evolution of \( v \) is determined by (1)–(5) and the form of the free top surface. Moreover, as long as \( y(\cdot, t), \theta(\cdot, t) \) and \( v(\cdot, t) \) are sufficiently regular and localized, for example, \( y(\cdot, t), v(\cdot, t) \in L^2 \) and \( \theta(\cdot, t) \in H^2 \), then, due to (48), the evolution of \( x \) is completely determined by the evolution of \( \theta \) and therefore \( U(\cdot, t) \) can be represented as a function of \( y(\cdot, t), \theta(\cdot, t) \) and \( v(\cdot, t) \).

Finally, using all the above information, one can derive the following evolutionary system:

\[
y_t = U \cos \theta + Ty_\alpha, \tag{52}
\]

\[
v_t = -y_\alpha + b \theta_\alpha - \delta \delta_\alpha + U \theta_t, \tag{53}
\]

\[
\theta_t = U_\alpha + T \theta_\alpha, \tag{54}
\]

\[
\delta_{\alpha t} = -c \theta_\alpha - b \theta_\alpha \alpha_\alpha - (\delta \delta_\alpha)_\alpha + (U_\alpha + v \theta_\alpha)^2, \tag{55}
\]

\[
y_\alpha = \sin \theta, \tag{56}
\]

\[
\delta = v - T, \tag{57}
\]

where

\[
c = U_t + v \theta_t + \delta U_\alpha + \delta v \theta_\alpha + \cos \theta. \tag{58}
\]

For further details of the derivation of this system, an explicit formula for \( U \) and the local well-posedness of the system in Sobolev spaces, we refer to [3,8].

The evolution equations (54) and (55) are included because they have better regularity properties than the evolution equations for the spatial derivatives of \( y \) and \( v \). The evolutionary system (52)–(58) could be posed entirely in terms of the variables \( \theta \) and \( \delta_\alpha \) as it is done in [3]. But since \( y \) and \( v \) are the physically relevant
variables for which we would like to derive and justify the NLS approximation, we keep these variables. Consequently, we need the equations (56)–(57) as consistency conditions between the systems (52)–(53) and (54)–(55).

The main advantage of system (52)–(58) is that in the case of surface tension, i.e., for \( b > 0 \), the term with the most derivatives in (52)–(58) is linear.

In order to derive the NLS approximation and to prove the error estimates we need to extract the linear and the quadratic components of system (52)–(58). In this context, the linear operator \( K_0 \) defined by its symbol

\[
\hat{K}_0(k) = -i \tanh(k)
\]

for all \( k \in \mathbb{R} \) plays an important role. The operator \( K_0 \) is the linearization of the operator \( K \) from (6) around the trivial solution \((\eta, \phi_x) = (0, 0)\). We present some properties of \( K_0 \) which we will need below. We now have the following:

**Lemma 2.1.** Let \( s \geq 0 \) and \( q > \frac{1}{2} \). Then we have

\[
\| K_0 f \|_{H^s} \lesssim \| f \|_{H^s},
\]

(60)

\[
\| [K_0, g] f \|_{H^s} \lesssim \| g \|_{H^{s+q}} \| f \|_{H^q},
\]

(61)

\[
\| [K_0, g] f \|_{H^s} \lesssim \| g \|_{H^q} \| f \|_{H^s},
\]

(62)

\[
\| (1 + K_0^2) f \|_{H^s} \lesssim \| f \|_{H^0}.
\]

(63)

**Proof.** The lemma is a special case of Lemma 3.7 and Lemma 3.8 in [8].\( \square \)

With the help of \( K_0 \) one obtains the following expansion of the system (52)–(58).

**Lemma 2.2.**

\[
y_t = K_0 v + (K_0[K_0, y])v - (1 + K_0^2)(yv) + m_I,
\]

(64)

\[
v_t = -y_a + by_{a\alpha a} - \frac{1}{2}(v^2)_\alpha + \frac{1}{2}((K_0 v)^2)_\alpha + m_{II},
\]

(65)

\[
\kappa_t = K_0 \delta_{\alpha a} - (\delta \kappa)_\alpha + (K_0[K_0, y] \delta_\alpha - (1 + K_0^2)(y \delta_\alpha))_{a\alpha} + (K_0[K_0, \theta] \delta_\alpha - (1 + K_0^2)(\theta \delta_\alpha))_\alpha + (m_{III})_a,
\]

(66)

\[
\delta_{\alpha a t} = -\kappa_a + b \kappa_{a\alpha a} + ((K_0 \theta - b K_0 \kappa_a + c_0) \kappa)_a
\]

\[
-((\delta \kappa_{a\alpha})_a - ((\delta \kappa_a)^2)_a + ((K_0 \delta_a)^2)_a + (m_{IV})_a,
\]

(67)

\[
y_\alpha = \theta + m_V,
\]

(68)

\[
\theta(\alpha, t) = \int_{-\infty}^\alpha \kappa(\beta, t) d\beta,
\]

(69)

\[
\delta(\alpha, t) = v(\alpha, t) - \int_{-\infty}^\alpha ((K_0 v) \kappa)(\beta, t) d\beta + m_{V I}(\alpha, t),
\]

(70)

where

\[
\| m_I \|_{H^s} \lesssim (\| y \|_{L^2}^2 + \| \theta \|_{H^s}^2)(\| v \|_{L^2}^2 + \| \delta_\alpha \|_{L^2})
\]

(71)
for $s \geq 1$, as long as $\|y\|_{L^2}, \|\theta\|_{H^s} \lesssim 1$,
\[ \|m_{II}\|_{H^s} \lesssim (\|y\|_{L^2} + \|\theta\|_{H^{s+1}})(\|v\|_{L^2}^2 + \|\delta\|_{H^2}^2) \]  
(72)

for $s \geq 2$, as long as $\|y\|_{L^2}, \|\theta\|_{H^{s+1}} \lesssim 1$,
\[ \|(m_{II})_\alpha\|_{H^s} \lesssim (\|y\|_{L^2}^2 + \|\theta\|_{H^{s+1}}^2)(\|v\|_{L^2} + \|\delta\|_{H^2}) \]  
(73)

for $s \geq 2$, as long as $\|y\|_{L^2}, \|\theta\|_{H^{s+1}} \lesssim 1$,
\[ \|(m_{IV})_\alpha\|_{H^s} \lesssim (\|y\|_{L^2} + \|\theta\|_{H^{s+1}})(\|v\|_{L^2}^2 + \|\delta\|_{H^{s+1}}^2) \]  
(74)

for $s \geq 2$, as long as $\|y\|_{L^2}, \|\theta\|_{H^{s+1}} \lesssim 1$,
\[ \|m_{VI}\|_{H^s} \lesssim \|\theta\|_{H^s}^3 \]  
(75)

for $s \geq 1$, as long as $\|\theta\|_{H^s} \lesssim 1$,
\[ \|m_{VI}C_0 + (m_{VI})_\alpha\|_{H^{s-2}} \lesssim (\|y\|_{L^2}^2 + \|\theta\|_{H^s}^2)(\|v\|_{L^2} + \|\delta\|_{H^{s-2}}) \]  
(76)

for $s \geq 2$, as long as $\|y\|_{L^2}, \|\theta\|_{H^s} \lesssim 1$, and
\[ \|c_0\|_{H^s} \lesssim \|y\|_{L^2}^2 + \|\theta\|_{H^s}^2 + b\|\theta\|_{H^{s+1}}^2 + \|v\|_{L^2}^2 + \|\delta\|_{H^s}^2 \]  
(77)

for $s \geq 6$, as long as $\|y\|_{L^2}, \|\theta\|_{H^s}, \sqrt{b}\|\theta\|_{H^{s+1}}, \|v\|_{L^2}, \|\delta\|_{H^s} \lesssim 1$. Moreover, we have
\[ \|\theta\|_{H^s} \lesssim \|y\|_{L^2} + \|\kappa\|_{H^{s-1}} \]  
(78)

for $s \geq 1$, as long as $\|y\|_{L^2} + \|\kappa\|_{L^2} \ll 1$, and
\[ \|\delta\|_{H^s} \lesssim (1 + \|y\|_{L^2} + \|\kappa\|_{H^1})(\|v\|_{L^2} + \|\delta\|_{H^{s-1}}) \]  
(79)

for $s \geq 1$, as long as $\|y\|_{L^2} + \|\kappa\|_{H^1} \ll 1$.

All bounds are uniform with respect to $b \lesssim 1$.

**Proof:** The expansions (64)–(70) follow directly from Lemma 3.9 in [8]. The bounds (71)–(79) follow directly from the bounds in the Lemmas 3.1–3.9 in [8], which are also uniform with respect to $b \lesssim 1$, and the well-known interpolation inequality
\[ \|f\|_{L^2} \lesssim \mu\|f\|_{L^2} + \mu^{-1}\|f\|_{L^2} \]  
(80)
for all $f \in H^2(\mathbb{R})$ and all $\mu > 0$. \hfill \Box

In Lemma 2.2 we replaced the evolution equations for $\theta$ and $\delta\alpha$ by the evolution equations for the respective spatial derivative, where we used that the spatial derivative of the tangent angle $\theta$ is the curvature $\kappa$. Due to this replacement, the Fourier transform of all quadratic terms in the resulting evolutionary system (64)–(70) vanishes at $k = 0$ such that we can handle the resonances at $k = 0$ and $k = \pm k_0$ as explained in the introduction.

To emphasize the geometric meaning of the terms in system (64)–(70) we included both $\theta$ and $\kappa$ in the system. Therefore, we need the additional equation...
(69) as a consistency condition. The expansion of (58) is incorporated directly into
(64)–(70).

In contrast to the justification of the NLS approximation, the occurring res-
onances need not to be taken into account in the context of proving local well-
posedness of the water wave problem. Hence, local well-posedness can be shown
by analyzing the evolution equations for $\theta$ and $\delta_\alpha$, whereas for proving the validity
of the NLS approximation we will need the evolution equations for $\kappa$ and $\delta_{\alpha\alpha}$.

We diagonalize system (64)–(70) by

$$\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & -\sigma^{-1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix},$$

(81)

$$\begin{pmatrix} \kappa \\ \delta_{\alpha\alpha} \end{pmatrix} = \begin{pmatrix} \sigma^{-1} & -\sigma^{-1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_{-2} \\ u_2 \end{pmatrix},$$

(82)

where $\sigma^{-1}$ is the inverse of the linear operator $\sigma$ with the symbol

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

(83)

Then we have

$$\begin{pmatrix} u_{-1} \\ u_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sigma & 1 \\ -\sigma & 1 \end{pmatrix} \begin{pmatrix} y \\ v \end{pmatrix},$$

(84)

$$\begin{pmatrix} u_{-2} \\ u_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sigma & 1 \\ -\sigma & 1 \end{pmatrix} \begin{pmatrix} \kappa \\ \delta_{\alpha\alpha} \end{pmatrix},$$

(85)

and Lemma 2.2 yields

$$(u_{-1})_t = -i \omega u_{-1} - \frac{1}{4}((u_{-1} + u_1)^2)_\alpha + \frac{1}{4}((K_0(u_{-1} + u_1))^2)_\alpha$$

$$+ \frac{1}{2}(\sigma K_0[\sigma^{-1}(u_{-1} - u_1)](u_{-1} + u_1))_\alpha$$

$$- \frac{1}{2}(\sigma(1 + K_0^2)\sigma^{-1}(u_{-1} - u_1)(u_{-1} + u_1))_\alpha$$

$$+ m_{-1},$$

(86)

$$(u_1)_t = i \omega u_1 - \frac{1}{4}((u_{-1} + u_1)^2)_\alpha + \frac{1}{4}((K_0(u_{-1} + u_1))^2)_\alpha$$

$$- \frac{1}{2}(\sigma K_0[\sigma^{-1}(u_{-1} - u_1)](u_{-1} + u_1))_\alpha$$

$$+ \frac{1}{2}(\sigma(1 + K_0^2)\sigma^{-1}(u_{-1} - u_1)(u_{-1} + u_1))_\alpha$$

$$+ m_1,$$

(87)

$$(u_{-2})_t = -i \omega u_{-2} - (\partial_\alpha^{-2}(u_{-2} + u_2)u_{-2})_\alpha$$

$$- \frac{1}{2}([\sigma, \partial_\alpha^{-2}(u_{-2} + u_2)]\sigma^{-1}(u_{-2} - u_2))_\alpha$$

$$+ \frac{1}{2}(K_0\partial_\alpha^{-1}\sigma^{-1}(u_{-2} - u_2)\sigma^{-1}(u_{-2} - u_2))_\alpha.$$
\[
\begin{aligned}
\frac{1}{2} b \sigma^{-1}(u_{-2} - u_2) K_0 \sigma^{-1}(u_{-2} - u_2)_\alpha \\
\frac{1}{2} ((\partial_{\alpha}^{-1}(u_{-2} + u_2))^2)_\alpha + \frac{1}{2} ((K_0 \partial_{\alpha}^{-1}(u_{-2} + u_2))^2)_\alpha \\
+ \frac{1}{2} (\sigma K_0[K_0, \sigma^{-1}(u_{-1} - u_1)] \partial_{\alpha}^{-1}(u_{-2} + u_2))_\alpha \\
\frac{1}{2} (\sigma (1 + K_0^2)(\sigma^{-1}(u_{-1} - u_1) \partial_{\alpha}^{-1}(u_{-2} + u_2)))_\alpha \\
+ \frac{1}{2} (\sigma K_0[K_0, \partial_{\alpha}^{-1}(u_{-2} - u_2)] \partial_{\alpha}^{-1}(u_{-2} + u_2))_\alpha \\
\frac{1}{2} (\sigma (1 + K_0^2)(\partial_{\alpha}^{-1}(u_{-2} - u_2) \partial_{\alpha}^{-1}(u_{-2} + u_2)))_\alpha \\
+ \frac{1}{2} (c_1 \sigma^{-1}(u_{-2} - u_2))_\alpha + (m_2)_\alpha,
\end{aligned}
\]

(88)

\[
(u_2)_t = i \omega u_2 - (\partial_{\alpha}^{-2}(u_{-2} + u_2)u_2)_\alpha \\
+ \frac{1}{2} ([\sigma, \partial_{\alpha}^{-2}(u_{-2} + u_2)] \sigma^{-1}(u_{-2} - u_2))_\alpha \\
+ \frac{1}{2} (K_0 \partial_{\alpha}^{-1}(u_{-2} - u_2) \sigma^{-1}(u_{-2} - u_2))_\alpha \\
- \frac{1}{2} b \sigma^{-1}(u_{-2} - u_2) K_0 \sigma^{-1}(u_{-2} - u_2)_\alpha \\
- \frac{1}{2} ((\partial_{\alpha}^{-1}(u_{-2} + u_2))^2)_\alpha + \frac{1}{2} ((K_0 \partial_{\alpha}^{-1}(u_{-2} + u_2))^2)_\alpha \\
\frac{1}{2} (\sigma K_0[K_0, \sigma^{-1}(u_{-1} - u_1)] \partial_{\alpha}^{-1}(u_{-2} + u_2))_\alpha \\
+ \frac{1}{2} (\sigma (1 + K_0^2)(\sigma^{-1}(u_{-1} - u_1) \partial_{\alpha}^{-1}(u_{-2} + u_2)))_\alpha \\
+ \frac{1}{2} (\sigma K_0[K_0, \partial_{\alpha}^{-1}(u_{-2} - u_2)] \partial_{\alpha}^{-1}(u_{-2} + u_2))_\alpha \\
+ \frac{1}{2} (\sigma (1 + K_0^2)(\partial_{\alpha}^{-1}(u_{-2} - u_2) \partial_{\alpha}^{-1}(u_{-2} + u_2)))_\alpha \\
+ \frac{1}{2} (c_1 \sigma^{-1}(u_{-2} - u_2))_\alpha + (m_2)_\alpha
\]

(89)

as well as

\[
\partial_{\alpha}^{-1} \sigma^{-1}(u_{-2} - u_2) = \sigma^{-1}(u_{-1} - u_1)_\alpha + m_3,
\]

(90)

\[
\partial_{\alpha}^{-2}(u_{-2} + u_2) = (u_{-1} + u_1) - \partial_{\alpha}^{-1}(K_0(u_{-1} + u_1) \sigma^{-1}(u_{-2} - u_2)) + m_3,
\]

(91)

where \( \omega \) is the linear operator with the symbol

\[
\omega(k) = \omega(k, b) = \text{sgn}(k) \sqrt{(k + bk^3) \text{tanh}(k)},
\]

(92)
\frac{\partial_{\alpha}^{-1}}{\partial_{-i k^{-1}}} \text{defined by the symbol } -i k^{-1} \text{ and and}
\|m_{-1}\|_{H^s} \lesssim \|u_{-1}\|_{L^2} + \|u_1\|_{L^2}^3 + \|u_{-2}\|_{H^{s-1/2}}^3 + \|u_2\|_{H^{s-1/2}}^3
\quad \text{ (93)}
for \(s \geq 2\), as long as \(\|u_{\pm 1}\|_{L^2}, \|u_{\pm 2}\|_{H^{s-1/2}} \ll 1\),
\|m_{1}\|_{H^s} \lesssim \|u_{-1}\|_{L^2}^3 + \|u_1\|_{L^2}^3 + \|u_{-2}\|_{H^{s-1/2}}^3 + \|u_2\|_{H^{s-1/2}}^3
\quad \text{ (94)}
for \(s \geq 2\), as long as \(\|u_{\pm 1}\|_{L^2}, \|u_{\pm 2}\|_{H^{s-1/2}} \ll 1\),
\|(m_{-2})_{\alpha}\|_{H^s} \lesssim \|u_{-1}\|_{L^2}^3 + \|u_1\|_{L^2}^3 + \|u_{-2}\|_{H^{s-1/2}}^3 + \|u_2\|_{H^{s-1/2}}^3
\quad \text{ (95)}
for \(s \geq 2\), as long as \(\|u_{\pm 1}\|_{L^2}, \|u_{\pm 2}\|_{H^s} \ll 1\),
\|m_{-3}\|_{H^s} \lesssim \|u_{-1}\|_{L^2}^3 + \|u_1\|_{L^2}^3 + \|\sigma^{-1}(u_{-2} - u_2)\|_{H^{s-1}}^3
\quad \text{ (97)}
for \(s \geq 2\), as long as \(\|u_{\pm 1}\|_{L^2}, \|\sigma^{-1}(u_{-2} - u_2)\|_{H^{s-1}} \ll 1\),
\|m_3\|_{C^0} + \|(m_3)_{\alpha}\|_{H^{s-1}} \lesssim \|u_{-1}\|_{L^2}^3 + \|u_1\|_{L^2}^3
\quad + \|\sigma^{-1}(u_{-2} - u_2)\|_{H^{s-1}}^3 + \|u_2\|_{H^{s-3}}^3
\quad \text{ (98)}
for \(s \geq 3\), as long as \(\|u_{\pm 1}\|_{L^2}, \|\sigma^{-1}(u_{-2} - u_2)\|_{H^{s-1}}, \|u_{-2}\|_{H^{s-3}} \ll 1\), as well
\|c_1\|_{H^s} \lesssim \|u_{-1}\|_{L^2}^2 + \|u_1\|_{L^2}^2 + \|u_{-2}\|_{H^{s-1}}^2 + \|u_2\|_{H^{s-1}}^2
\quad \text{ (99)}
for \(s \geq 6\), as long as \(\|u_{\pm 1}\|_{L^2}, \|u_{\pm 2}\|_{H^{s-1}} \ll 1\). All bounds are uniform with respect to \(b \lesssim 1\).

In the diagonalized system (86)-(91) both \(\theta\) and \(\kappa\) are expressed in terms of the variables \(u_{-2}\) and \(u_2\) such that the consistency condition (69) is not needed anymore.

We close this section by collecting some properties of the operator \(\sigma\), which will be useful for our further argumentation. We have the identities
\[\sigma K_0[K_0, \sigma^{-1} f]g - \sigma(1 + K_0^2)(\sigma^{-1} fg) = -gf - K_0g K_0 f - [\sigma, g] \sigma^{-1} f - [K_0 \sigma, K_0 g] \sigma^{-1} f\]
\quad \text{ (100)}
as well as
\[[\sigma, f] \sigma^{-1} g = \sigma^{-1} g \sigma f - gf + [\sigma, \sigma^{-1} g] f.\]
\quad \text{ (101)}
Moreover, a direct computation using
\[\sigma(k) - \sigma(l) = (\sigma(k) + \sigma(l))^{-1}(\sigma^2(k) - \sigma^2(l))\]
\quad \text{ (102)}
for all \(k, l \in \mathbb{R}\) and the mean value theorem yields
\[\|[\sigma, g] f\|_{H^s} \lesssim \|\sigma \sigma_{\alpha}\|_{H^{s-1}} \|f\|_{H^{s-1}} + \|g_{\alpha}\|_{H^{s-1}} \|\sigma f\|_{H^{s-1}}\]
\quad \text{ (103)}
for \(s > 3/2\).
3. The Derivation of the NLS Approximation

In order to derive the NLS approximation for system (86)–(91), we introduce the vector-valued function

\[ \mathcal{U} := \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \] with \( \mathcal{U}_j := \begin{pmatrix} u_{j-1} \\ u_j \end{pmatrix} \) for \( j = 1, 2, \)

and make the ansatz

\[ \mathcal{U} = \varepsilon \tilde{\Psi} = \varepsilon \begin{pmatrix} \tilde{\Psi}_1 \\ \tilde{\Psi}_2 \end{pmatrix} \] (104)

with

\[ \varepsilon \tilde{\Psi}_j = \varepsilon \tilde{\Psi}_j^0 + \varepsilon^2 \tilde{\Psi}_j^{0,1} + \varepsilon^2 \tilde{\Psi}_j^{0,2} + \varepsilon^2 \tilde{\Psi}_j^{0,3}, \]

\[ \varepsilon \tilde{\Psi}_j^{0,1}(\alpha, t) = \varepsilon \tilde{A}_{-j+1}^0 (\alpha - c_g t, \varepsilon^2 t) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \]

\[ \varepsilon^2 \tilde{\Psi}_j^{0,2}(\alpha, t) = \begin{pmatrix} \varepsilon^2 \tilde{A}_{-j+1}^0 (\alpha - c_g t, \varepsilon^2 t) \\ \varepsilon^2 \tilde{A}_{j-1}^0 (\alpha - c_g t, \varepsilon^2 t) \end{pmatrix}, \]

\[ \varepsilon^2 \tilde{\Psi}_j^{0,3}(\alpha, t) = \begin{pmatrix} \varepsilon^2 \tilde{A}_{-j+1}^0 (\alpha - c_g t, \varepsilon^2 t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \varepsilon^2 \tilde{A}_{j-1}^0 (\alpha - c_g t, \varepsilon^2 t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix}, \]

where \( 0 < \varepsilon \ll 1, j \in \{1, 2\}, \mathcal{E} = e^{i(k_0 \alpha - \omega_0 t)}, k_0 > 0, \omega_0 = \omega(k_0, b), c_g = \partial_k \omega(k_0, b) \text{ and } \tilde{A}_{m-\ell}^0 = \tilde{A}_{m\ell}^0. \)

Our ansatz leads to waves moving to the right. For waves moving to the left one has to replace in the above ansatz the vector \((1, 0)^T\) by \((0, 1)^T\) as well as \(-\omega_0\) by \(\omega_0\) and \(c_g\) by \(-c_g\).

First, we insert the ansatz (104) for \(\mathcal{U}_1\) into (86)–(87). Then we replace the dispersion relation \(k \mapsto \omega(k, b)\) in all terms of the form \(\omega \tilde{A}_{m\ell}^0 \mathcal{E}^\ell\) by their Taylor expansions around \(k = \ell k_0\). (Details of these expansions are contained in Lemma 25 of [40], for example.) After that, we equate the coefficients of the \(\varepsilon^0 \mathcal{E}^\ell\) to zero.

We find that the coefficients of \(\varepsilon^1 \mathcal{E}\) and \(\varepsilon^2 \mathcal{E}\) vanish identically due to the definition of \(\omega\) and \(c_g\). For \(\varepsilon^3 \mathcal{E}\) we obtain

\[ \partial_\tau \tilde{A}_{-11}^0 = \frac{1}{2} i \partial_\alpha^2 \omega(k_0, b) \partial_\alpha^2 \tilde{A}_{-11}^0 + g_1, \]

where \(\tau = \varepsilon^2 t, \alpha = \varepsilon (\alpha - c_g t)\) and \(g_1\) is a sum of multiples of \(\tilde{A}_{-11}^0 |\tilde{A}_{-11}^0|^2, \tilde{A}_{-11}^0 \tilde{A}_{00}^0\) and \(\tilde{A}_{01}^0 \tilde{A}_{m0}^0\) with \(m \in \{\pm 1\}\). In the next steps we obtain algebraic relations such that the \(\tilde{A}_{m2}^0\) can be expressed in terms of \((\tilde{A}_{-11}^0)^2\) and the \(\tilde{A}_{m0}^0\) in terms of \(|\tilde{A}_{-11}^0|^2\), respectively.

For \(\varepsilon^2 \mathcal{E}\) we obtain

\[ (-2\omega_0 + \omega(2k_0, b)) \tilde{A}_{-12}^0 = \gamma_{-12} |\tilde{A}_{-11}^0|^2, \]

\[ (-2\omega_0 - \omega(2k_0, b)) \tilde{A}_{12}^0 = \gamma_{12} |\tilde{A}_{-11}^0|^2 \]
with coefficients $\gamma_{m2} \in \mathbb{C}$. For all $b \geq 0$ with
\[ -2\omega_0 \pm \omega(2k_0, b) \neq 0, \tag{105} \]
the $\tilde{A}_{m2}^0$ are well-defined in terms of $(\tilde{A}_{-11}^0)^2$. All terms vanish identically for $\varepsilon^2 \mathbf{E}^0$. This is obvious for the linear terms. For the quadratic terms the calculations are analogous to those of Appendix A of [40] (see specifically equation (94)). The nonlinear terms in $\varepsilon^3 \mathbf{E}^0$ must be perfect derivatives with respect to $\alpha$ since no other combination of terms in the approximation (104) leads to terms proportional to $\varepsilon^3 \mathbf{E}^0$. Thus we find
\[ -c_g \partial_\alpha \tilde{A}_{-10}^0 = -\partial_k \omega(0, b) \partial_\alpha \tilde{A}_{-10}^0 + \gamma_{-10} \partial_\alpha (\tilde{A}_{-11}^0 \tilde{A}_{-11}^0), \]
\[ -c_g \partial_\alpha \tilde{A}_{10}^0 = \partial_k \omega(0, b) \partial_\alpha \tilde{A}_{10}^0 + \gamma_{10} \partial_\alpha (\tilde{A}_{-11}^0 \tilde{A}_{-11}^0), \]
where we now have $\gamma_{m0} \in \mathbb{R}$, on account of to the fact that we consider a real-valued problem. For all $b \geq 0$ with
\[ c_g \neq \pm \partial_k \omega(0, b), \tag{106} \]
we can divide the equations for $\varepsilon^3 \mathbf{E}^0$ by $\partial_\alpha$ and can express the $\tilde{A}_{m0}^0$ in terms of $|\tilde{A}_{-11}^0|^2$.

As mentioned above the nonlinear terms in the equation for $\varepsilon^3 \mathbf{E}^1$ include $\tilde{A}_{-11}^0 |\tilde{A}_{-11}^0|^2$ as well as terms consisting of combinations of $\tilde{A}_{-11}^0$ with the $\tilde{A}_{m0}^0$ and of $\tilde{A}_{-11}^0$ with the $\tilde{A}_{m2}^0$. Eliminating $\tilde{A}_{m0}^0$ and $\tilde{A}_{m2}^0$ by the algebraic relations obtained for $\varepsilon^3 \mathbf{E}^0$ and $\varepsilon^2 \mathbf{E}^2$ finally gives the NLS equation
\[ \partial_\tau \tilde{A}_{-11}^0 = i \frac{\omega^2(k_0, b)}{2} \partial_\alpha^2 \tilde{A}_{-11}^0 + i \nu_2(k_0, b) \tilde{A}_{-11}^0 |\tilde{A}_{-11}^0|^2, \tag{107} \]
with a $\nu_2(k_0, b) \in \mathbb{R}$.

An explicit formula for $\nu_2$ can be found in [5, p. 504]. It can be seen with the help of that formula if the NLS equation (107) is defocusing or focusing for a given basic wave number $k_0 > 0$. Since we will consider solutions of (107) on time intervals $[0, \tau_0]$ with $\tau_0 \sim \mathcal{O}(1)$, this will not affect our analysis.

The approximation function $\varepsilon \tilde{\Psi}_2$ is determined by inserting (104) into (90)–(91), using the formulas for $\varepsilon \tilde{\Psi}_1$ derived above and equating the coefficients of the $\varepsilon^p \mathbf{E}^\ell$ to zero. It turns out that $\varepsilon \tilde{\Psi}_2$ can be expressed in terms of the components of $\varepsilon \tilde{\Psi}_1$ and its derivatives. In particular, we have
\[ \tilde{\Psi}_{2 \pm 1}^0 = \partial_\alpha^2 \tilde{\Psi}_{1 \pm 1}^0. \]

To prove the approximation property of the NLS equation (107) it will be helpful to extend the approximation $\varepsilon \tilde{\Psi}$ by higher order correction terms in order to make the residual of the resulting approximation of the equations (86)–(91) smaller in Sobolev norms. The residual of an approximation $\alpha$ of an algebraic or differential equation
\[ A(f) = B(f), \tag{108} \]
where $A$, $B$ are functions depending on the function $f$ and in the case of a differential equation also on derivatives of $f$, is defined by

$$\text{Res}(a) := B(a) - A(a).$$

Hence, $\text{Res}(a)$ contains all terms that do not cancel after inserting the ansatz $f = a$ in (108) and quantifies how much $a$ fails to be a solution of (108). There holds $\text{Res}(a) = 0$ if and only if $a$ is an exact solution of (108).

We introduce the notation

$$\text{Res}(\epsilon \tilde{\Psi}) = \begin{pmatrix} \text{Res}_1(\epsilon \tilde{\Psi}) \\ \text{Res}_2(\epsilon \tilde{\Psi}) \\ \text{Res}_3(\epsilon \tilde{\Psi}) \end{pmatrix}$$

with $\text{Res}_j(\epsilon \tilde{\Psi}) = \begin{pmatrix} \text{res}_{-j}(\epsilon \tilde{\Psi}) \\ \text{res}_j(\epsilon \tilde{\Psi}) \end{pmatrix}$ for $j = 1, 2, 3$, where $\text{res}_{-1}(\epsilon \tilde{\Psi})$ is the residual of $\epsilon \tilde{\Psi}$ of equation (86), $\text{res}_1(\epsilon \tilde{\Psi})$ the residual of $\epsilon \tilde{\Psi}$ of equation (87), $\text{res}_2(\epsilon \tilde{\Psi})$ the residual of $\epsilon \tilde{\Psi}$ of equation (88), $\text{res}_2(\epsilon \tilde{\Psi})$ the residual of $\epsilon \tilde{\Psi}$ of equation (89), $\text{res}_3(\epsilon \tilde{\Psi})$ the residual of $\epsilon \tilde{\Psi}$ of equation (90) and $\text{res}_3(\epsilon \tilde{\Psi})$ the residual of $\epsilon \tilde{\Psi}$ of equation (91).

In order to replace $\epsilon \tilde{\Psi}$ in $\text{Res}(\epsilon \tilde{\Psi})$ by a better approximation $\epsilon \Psi$ we proceed analogously as in Section 2 of [14]. In a first step we construct an extended approximation

$$\mathcal{U} = \epsilon \tilde{\Psi}^{\text{ext}} = \epsilon \left( \tilde{\Psi}^{\text{ext}}_1 \right)$$

with

$$\epsilon \tilde{\Psi}^{\text{ext}}_j = \epsilon \tilde{\Psi}_j + \epsilon^2 \tilde{\Psi}_j^{\text{add}},$$

where $\epsilon \tilde{\Psi}_j$ is as above and $\epsilon^2 \tilde{\Psi}_j^{\text{add}}$ is of the form

$$\begin{align*}
\epsilon^2 \tilde{\Psi}_j^{\text{add}} &= \sum_{\ell \in \{\pm 1\}} \sum_{n=1}^{4} \left( \epsilon^{1+n} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right) \\
&\quad + \sum_{n=1}^{3} \left( \epsilon^{2+n} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right) \\
&\quad + \sum_{\ell \in \{\pm 2\}} \sum_{n=1}^{3} \left( \epsilon^{2+n} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right) \\
&\quad + \sum_{\ell \in \{\pm 3\}} \sum_{n=0}^{2} \left( \epsilon^{3+n} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right) \\
&\quad + \sum_{\ell \in \{\pm 4\}} \sum_{n=0}^{1} \left( \epsilon^{4+n} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right) \\
&\quad + \sum_{\ell \in \{\pm 5\}} \left( \epsilon^{5} \tilde{A}_{\ell j}^n (\epsilon (\alpha - c_g t), \epsilon^2 t) \mathbf{E}^\ell \right)
\end{align*}
with $\tilde{A}_{\mp j-\ell}^n = \tilde{A}_{\mp j\ell}^n$. Then we have

$$\varepsilon\tilde{\Psi}^0_{j} = \sum_{|\ell| \leq 5} \sum_{\beta(\ell, n) \leq 5} \varepsilon^\beta(\ell, n) \tilde{\Psi}^n_{j\ell},$$

where $j \in \{1, 2\}$, $\ell \in \mathbb{Z}$, $n \in \mathbb{N}_0$ and $\beta(\ell, n) = 1 + ||\ell| - 1| + n$ as well as

$$\varepsilon\tilde{\Psi}^{0\mp}_{1\pm 1} = \varepsilon\tilde{\psi}^{0\mp}_{1\pm 1}(1 0),$$

$$\varepsilon\tilde{\Psi}^{0\mp}_{2\pm 1} = \varepsilon\tilde{\psi}^{0\mp}_{2\pm 1}(1 0),$$

$$\varepsilon^\beta(\ell, n) \tilde{\psi}^n_{j\ell} = \varepsilon^\beta(\ell, n) \begin{pmatrix} \tilde{\psi}^n_{j\ell} \\ \tilde{\psi}^n_{j\ell} \end{pmatrix} \text{ for } (\ell, n) \neq (\pm 1, 0),$$

$$\tilde{\psi}^n_{j\ell}(\alpha, t) = \tilde{A}_{\mp j\ell}^n(\varepsilon(\alpha - \varepsilon g t), \varepsilon^2 t) E^\ell.$$ 

The functions $\tilde{A}_{\mp j\ell}^n$ are computed in a similar procedure as $\tilde{A}_{\mp 1\pm 1}, \tilde{A}_{\mp j\ell}^0$ and $\tilde{A}_{\mp j2}^0$ in $\varepsilon \tilde{\Psi}^0_j$. More precisely, inserting $\varepsilon\tilde{\psi}^0_{1\ell}$ into (86)–(87) and equating the coefficients in front of the $\varepsilon^\beta(\ell, n) E^\ell$ to zero yields a system of algebraic equations and inhomogeneous linear Schrödinger equations that can be solved recursively. For all $b \geq 0$ with $-\ell \omega_0 \pm \omega(\ell k_0, b) \neq 0$ for $\ell \in \{2, 3, 4, 5\}$ and $c_g \neq \pm \delta \omega(0, b)$ the functions $\tilde{A}_{\pm j\ell}^n$ with $|p| = 1$ and $(p, |\ell|) \neq (-1, 1)$ are uniquely determined by the algebraic equations. The functions $\tilde{A}_{-1\pm 1}^n$ satisfy the inhomogeneous linear Schrödinger equations. Moreover, since the functions $\tilde{A}_{-1\pm 1}^4$ do not appear in the equations for any other $\tilde{A}_{-1\pm 1}^n$, we can set $\tilde{A}_{-1\pm 1}^4 = 0$.

The approximation function $\varepsilon\tilde{\psi}^{ext}_1$ is determined by inserting $\varepsilon\tilde{\psi}^{ext}_1$ into (90)–(91), using the formulas for $\varepsilon\tilde{\psi}^{ext}_1$ derived above and equating the coefficients of the $\varepsilon^\beta(\ell, n) E^\ell$ to zero. It turns out that $\varepsilon\tilde{\psi}^{ext}_2$ can be expressed in terms of the components of $\varepsilon\tilde{\psi}^{ext}_1$ and their first three spatial derivatives.

In a second step we apply a Fourier truncation procedure to the approximation $\varepsilon\tilde{\psi}^{ext}_1$. More precisely, we define

$$\psi^n_{\pm j\ell} := \mathcal{F}^{-1}(\chi_{[-\delta_0, \delta_0]} \mathcal{F}(\tilde{\psi}^n_{\pm j\ell} E^{-\ell})) E^\ell,$$

where $\chi_{[-\delta_0, \delta_0]}$ is the characteristic function on $[-\delta_0, \delta_0]$ and $\delta_0 = \delta_0(b) \in (0, k_0/20)$ will be determined in Section 4. Then the functions $\psi^n_{\pm j\ell}$ have the compact support

$$\{k \in \mathbb{R} : |k - \ell k_0| \leq \delta_0 < k_0/20\}$$

in Fourier space for all $0 < \varepsilon \ll 1$ and they are of the form

$$\psi^n_{\pm j\ell}(\alpha, t) = A^n_{\mp j\ell}(\varepsilon(\alpha - c_g t), \varepsilon^2 t) E^\ell,$$

with $A^n_{\pm j-\ell} = A^n_{\mp j\ell}$. By using these functions, we construct our final approximation

$$\mathcal{U} = \varepsilon \Psi = \varepsilon \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix},$$
with
\[ \varepsilon \Psi_j = \sum_{|\ell| \leq 5} \sum_{\beta(\ell, n) \leq 5} \varepsilon^{\beta(\ell, n)} \psi_j^n, \] (114)

where \( j \in \{1, 2\}, \ell \in \mathbb{Z}, n \in \mathbb{N}_0 \) and \( \beta(\ell, n) = 1 + |\ell| - 1 + n \) as well as

\[ \varepsilon \Psi_{1 \pm 1}^0 = \varepsilon \psi_{-1 \pm 1}^0 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \] (115)
\[ \varepsilon \Psi_{2 \pm 1}^0 = \varepsilon \partial_{\alpha}^2 \psi_{-1 \pm 1}^0 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \] (116)
\[ \varepsilon^{\beta(\ell, n)} \psi_j^n = \varepsilon^{\beta(\ell, n)} \begin{pmatrix} \psi_j^n \\ \psi_j^n \end{pmatrix} \] for \((\ell, n) \neq (\pm 1, 0)\). (117)

For later purposes, we set
\[ \psi_{\pm 1} := \psi_{-1 \pm 1}^0, \] (118)
\[ \varepsilon^2 \psi_j^h = \varepsilon^2 \begin{pmatrix} \psi_j^h \\ \psi_j^h \end{pmatrix} := \varepsilon \psi_j - (\varepsilon \Psi_{0 j}^0 + \varepsilon \Psi_{0 j-1}^0). \] (119)

Since the Fourier transform of the functions \( \tilde{\psi}_{\pm j}^n \) in \( \varepsilon \tilde{\Psi}^{ext} \) are strongly concentrated around the wave numbers \( \ell k_0 \) if the functions \( \tilde{A}_{\pm j}^n \) are sufficiently regular, the approximation \( \varepsilon \tilde{\Psi}^{ext} \) is only changed slightly by the Fourier truncation procedure. This fact is a consequence of the estimate
\[ \| (\chi_{[-\delta_0, \delta_0]} - 1) \varepsilon^{-1} \hat{f}(\varepsilon^{-1} \cdot) \|_{L^2(m)} \leq C(\delta_0) \varepsilon^{m+M-1/2} \| f \|_{H^{m+M}} \] (120)

for all \( M, m \geq 0 \). The Fourier truncation procedure will give us a simpler control of the error and makes our final approximation \( \varepsilon \Psi \) an analytic function. For related strategies compare [14,15,33,40].

As in Section 2 of [14], the following estimates for the modified residual hold:

**Lemma 3.1.** Let \( s_A \geq 10 \) and \( \tilde{A}_{-11}^0 \in C^0([0, \tau_0], H^{s_A}(\mathbb{R}, \mathbb{C})) \) be a solution of the NLS equation (107) with

\[ \sup_{\tau \in [0, \tau_0]} \| \tilde{A}_{-11}^0 \|_{H^{s_A}} \leq C_A. \]

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 \( \varepsilon \Psi \) satisfies

\[ \sup_{\tau \in [0, \tau_0/\varepsilon^2]} \| \text{Res}(\varepsilon \Psi) \|_{(H^s)^6} \leq C_{Res} \varepsilon^{11/2}, \] (121)
\[ \sup_{\tau \in [0, \tau_0/\varepsilon^2]} \| \varepsilon \Psi_{11} - (\varepsilon \tilde{\Psi}_{11}^0 + \varepsilon \tilde{\Psi}_{1-1}^0) \|_{(H^{s_A})^2} \leq C\psi \varepsilon^{3/2}, \] (122)
\[ \sup_{\tau \in [0, \tau_0/\varepsilon^2]} (\| \tilde{\Psi}_{j+1}^0 \|_{L^1(s+1)(\mathbb{R}, \mathbb{C})}^2 + \| \tilde{\Psi}_j^h \|_{L^1(s+1)(\mathbb{R}, \mathbb{C})}^2) \leq C\psi \] (123)

for \( j \in \{1, 2\} \).
Proof. The first extended approximation $\tilde{\psi}^{ext}$ is constructed in a way that formally we have $\text{Res}(\varepsilon \tilde{\psi}^{ext}) = O(\varepsilon^5)$ and $\varepsilon \tilde{\psi}^{ext}_1 - (\varepsilon \tilde{\psi}^{ext}_{11} + \varepsilon \tilde{\psi}^{ext}_{1-1}) = O(\varepsilon^2)$ on the time interval $[0, \tau_0/\varepsilon^2]$ if $\tilde{A}^{\pm}_{-11}$ is a solution of the NLS equation (107) for $T \in [0, \tau_0]$.

It can be shown analogously as in the proof of Theorem 2.5 in [14] that the regularity condition $\tilde{A}^{\pm}_{-11} \in C^0([0, \tau_0], H_s^s(\mathbb{R}, \mathbb{C}))$ with $s_A \geq 8$ implies $\tilde{A}^{\pm}_{-j+1} \in C^0([0, \tau_0], H_s^s(\mathbb{R}, \mathbb{C}))$ for $n \in \{1, 2, 3\}$ and $\tilde{A}^{\pm}_{p\ell} \in C^0([0, \tau_0], H_s^{s-3}(\mathbb{R}, \mathbb{C}))$ if $(p, \ell) \neq (-j, 1)$, where the respective Sobolev norms are uniformly bounded by the $H_s^s$-norm of $\tilde{A}^{\pm}_{-11}$.

Therefore, by taking into account that $\|f(\varepsilon \cdot)\|_{L^2} = \varepsilon^{-1/2} \|f\|_{L^2}$, we conclude that there exist constants $C_1, C_2 > 0$ depending on $C_A$ such that

$$\sup_{\tau \in [0, \tau_0/\varepsilon^2]} \|\text{Res}(\varepsilon \tilde{\psi}^{ext})\|_{(H_s^{s-10})^2} \leq C_1 \varepsilon^{11/2},$$

(124)

$$\sup_{\tau \in [0, \tau_0/\varepsilon^2]} \|\varepsilon \tilde{\psi}^{ext}_1 - (\varepsilon \tilde{\psi}^{ext}_{11} + \varepsilon \tilde{\psi}^{ext}_{1-1})\|_{(H_s^{s-8})^2} \leq C_2 \varepsilon^{3/2},$$

(125)

if we have $\tilde{A}^{\pm}_{-11} \in C^0([0, \tau_0], H_s^s(\mathbb{R}, \mathbb{C}))$ with $s_A \geq 10$ (because two additional spatial derivatives of $\tilde{A}^{\pm}_{-11}$ are needed to bound $\text{Res}(\varepsilon \tilde{\psi}^{ext})$).

Since the Fourier transform of the final approximation $\psi$ has a compact support whose size depends on $k_0$, there exists a constant $C = C(k_0) > 0$ such that $\|\Psi\|_{L^2} \leq C \|\Psi\|_{L^2}$ and $\|\tilde{\Psi}\|_{L^1(s)} \leq C \|	ilde{\Psi}\|_{L^1}$ for all $s \geq 0$. Hence, by using the above $L^2$-estimates for $\varepsilon \tilde{\psi}^{ext}$ as well as the estimate (120) for $f = \tilde{A}^{\pm}_{n\ell}$ for each $\tilde{A}^{\pm}_{n\ell}$ with $m = 0$, $M = M(\ell, n)$ determined by the maximal Sobolev regularity of the respective $\tilde{A}^{\pm}_{n\ell}$ and $\delta_0$ as above, we obtain (121) and

$$\sup_{\tau \in [0, \tau_0/\varepsilon^2]} \|\varepsilon \tilde{\psi}_{11} - (\varepsilon \tilde{\psi}_{11}^{0} + \varepsilon \tilde{\psi}_{1-1}^{0})\|_{(H_s^{s+8})^2} \leq C_3 \varepsilon^{3/2}$$

(126)

for a constant $C_3 = C_3(C_A) > 0$ if we have $s_A \geq 10$, which yields $\beta(\ell, n) + M(\ell, n) \geq 6$. By combining (126) and (120) for $f = \varepsilon \tilde{\psi}_{11}$, $m = s_A$, $M = 0$ and $\delta_0$ as above, we obtain (122).

Finally, since $\|\varepsilon^{-1} \tilde{f}(\varepsilon^{-1} \cdot)\|_{L^1} = \|\tilde{f}\|_{L^1}$, estimate (123) follows by construction of $\tilde{\psi}_{j \pm 1}^0$ and $\tilde{\psi}_j^0$.

Remark 3.2. Due to the estimate (121) of the residual obtained with the Fourier truncation procedure, the Sobolev index of the error estimate (47) in Theorem 1.1 is equal to the Sobolev index of the solution to the NLS equation. If we did not apply the Fourier truncation procedure, we could only use estimate (124) in the proof of the error estimates in Section 4. Then, for a given solution of the NLS equation in $H_s^s(\mathbb{R}, \mathbb{C})$, the resulting error could only be estimated in $H_s^{s-10}(\mathbb{R}, \mathbb{R})$.

Remark 3.3. The bound (123) will be used for instance to estimate

$$\|\psi f\|_{H_s^s} \leq C \|\psi\|_{C^2} \|f\|_{H_s^s} \leq C \|	ilde{\psi}\|_{L^1(s)(\mathbb{R}, \mathbb{C})} \|f\|_{H_s^s},$$

without loss of powers in $\varepsilon$ as would be the case with $\|	ilde{\psi}\|_{L^2(s)(\mathbb{R}, \mathbb{C})}$.
Moreover, by an analogous argumentation as in the proof of Lemma 3.3 in [14] we obtain the fact that \( \partial_t \psi_{\pm 1} \) can be approximated by \(-i\omega \psi_{\pm 1} \). More precisely, we obtain

Lemma 3.4. For all \( s \geq 0 \) there exists a constant \( C_\psi > 0 \) such that

\[
\| \partial_t \hat{\psi}_{\pm 1} + i \omega \hat{\psi}_{\pm 1} \|_{L^1(\mathbb{R})} \leq C_\psi \varepsilon^2.
\]  

(127)

4. The Error Estimates

In this section, we justify the NLS approximation for system (86)–(91).

4.1. The Structure of the Evolutionary System for the Error and the Approximation Result in the Arc Length Formulation

Our first step in justifying the NLS approximation is writing the exact solution \( U \) of (86)–(91) as the sum of the NLS approximation and the error. To avoid problems arising from the non-trivial resonances at \( k = \pm k_0 \), we rescale the error with the help of the weight function

\[
\hat{\vartheta}(k) = \begin{cases} 
1, & |k| > \delta_0, \\
\varepsilon + (1 - \varepsilon)|k|/\delta_0, & |k| \leq \delta_0,
\end{cases}
\]

where \( 0 < \varepsilon \ll 1 \) and \( \delta_0 = \delta_0(b) \in (0, k_0/20) \) will be defined below. That means, we write

\[
U = \varepsilon \Psi + \varepsilon^\beta \hat{\vartheta} \mathcal{R},
\]  

(128)

where \( \beta = 5/2 \) and

\[
\hat{\vartheta} \mathcal{R} := \left( \begin{array}{c}
\hat{\vartheta} \mathcal{R}_1 \\
\hat{\vartheta} \mathcal{R}_2
\end{array} \right) \text{ with } \hat{\vartheta} \mathcal{R}_j := \left( \begin{array}{c}
\hat{\vartheta} \mathcal{R}_{-j} \\
\hat{\vartheta} R_j
\end{array} \right) \text{ for } j = 1, 2,
\]

and \( \hat{\vartheta} \mathcal{R}_{\mp j} \) is defined by \( \hat{\vartheta} \mathcal{R}_{\mp j} = \hat{\vartheta} \mathcal{R}_{\mp j} \). By this choice \( \hat{\vartheta} \mathcal{R}_{\mp j}(k) \) is small at the wave numbers close to zero reflecting the fact that the quadratic terms of the evolutionary system of \( \mathcal{U} \) vanish at \( k = 0 \). Hence, we have

\[
\begin{align*}
u_{-1} &= \varepsilon \psi_c + \varepsilon^2 \psi_{-1}^h + \varepsilon^{5/2} \hat{\vartheta} R_{-1}, \\
u_1 &= \varepsilon^2 \psi_1^h + \varepsilon^{5/2} \hat{\vartheta} R_1, \\
u_{-2} &= \varepsilon^2 \psi_{-2}^h + \varepsilon^2 \psi_{-2} + \varepsilon^{5/2} \hat{\vartheta} R_{-2}, \\
u_2 &= \varepsilon^2 \psi_2^h + \varepsilon^{5/2} \hat{\vartheta} R_2,
\end{align*}
\]

where \( \psi_c = \psi_{-1} + \psi_1 \).

The definition of \( \hat{\vartheta} \) directly implies

\[
\sup_{k \in \mathbb{R}} |\hat{\vartheta}^{-1}(k)| = \varepsilon^{-1},
\]  

(129)

\[
\sup_{k \in \mathbb{R}} |(1 - \chi_{[-\delta_0, \delta_0]})(k) \hat{\vartheta}^{-1}(k)| = 1,
\]  

(130)
where the operator $\vartheta^{-1}$ is defined by its symbol $\widehat{\vartheta^{-1}}(k) = \widehat{\vartheta^{-1}}(k) = (\widehat{\vartheta(k)})^{-1}$.

Moreover, we have

$$|k \widehat{\vartheta^{-1}}(k)| = \begin{cases} |k| & \text{for } |k| > \delta_0, \\ \frac{|k|}{\varepsilon + (1 - \varepsilon) \frac{|k|}{\delta_0}} & \text{for } |k| \leq \delta_0. \end{cases}$$

Since

$$\frac{|k|}{\varepsilon + (1 - \varepsilon) \frac{|k|}{\delta_0}} = \frac{1}{\frac{\varepsilon}{|k|} + \frac{(1 - \varepsilon)}{\delta_0}} \leq \frac{1}{\frac{\varepsilon}{\delta_0} + \frac{(1 - \varepsilon)}{\delta_0}} = \delta_0$$

for $0 \neq |k| \leq \delta_0$, we obtain

$$\sup_{k \in \mathbb{R}} |k \widehat{\vartheta^{-1}}(k)| = \max\{\delta_0, |k|\}. \quad (131)$$

Furthermore, we have

$$\widehat{\vartheta^{-1}}(k) \widehat{\vartheta}(m) \chi_c(k - m) = O(1) \quad (132)$$

for $|k| \to \infty$ uniformly with respect to $m \in \mathbb{R}$, where $\chi_c$ is the characteristic function on $\text{supp } \widehat{\psi}_c$. Using (60)–(63), (86)–(99), (100) for $f = \vartheta(R_{-1} - R_1)$ and $g = \psi_c$, (103), (123) and (129)–(132), we obtain

$$\partial_t R_{\mp 1} = \mp i \omega R_{\mp 1}$$

$$\begin{align*}
-\varepsilon \vartheta^{-1} \partial_\alpha (\psi_c \vartheta R_{\mp 1}) &+ \varepsilon \vartheta^{-1} \partial_\alpha (K_0 \psi_c \vartheta K_0 R_{\pm 1}) \\
+ \varepsilon \vartheta^{-1} \partial_\alpha B_{\mp 1-1}(\psi_c, \vartheta R_{-1}) &+ \varepsilon \vartheta^{-1} \partial_\alpha B_{\mp 11}(\psi_c, \vartheta R_1) \\
+ \varepsilon^2 \vartheta^{-1} C_{\mp 1-1-1}(\psi_1, \psi_1 - \vartheta R_{-1}) &+ \varepsilon^2 \vartheta^{-1} C_{\mp 11-1}(\psi_1, \psi_1, \vartheta R_1) \\
+ \varepsilon^2 \vartheta^{-1} C_{\mp 111}(\psi_1, \psi_1, \vartheta R_1) &+ \varepsilon^2 \vartheta^{-1} C_{\mp 1111}(\psi_1, \psi_1, \vartheta R_1) \\
+ \varepsilon^2 M_{\mp 1}(\psi, \mathcal{R}) &+ \varepsilon^{-5/2} \vartheta^{-1} \text{res}_{\mp 1}(\varepsilon \psi),
\end{align*} \quad (133)$$

$$\begin{align*}
\partial_t R_{\mp 2} = \mp i \omega R_{\mp 2} \\
-\varepsilon \vartheta^{-1} \partial_\alpha (\psi_c \vartheta R_{\mp 2}) &+ \varepsilon \vartheta^{-1} \partial_\alpha (K_0 \sigma^{-1} \partial_\alpha \psi_c) \sigma^{-1} \vartheta (R_{-2} - R_2) \\
+ \varepsilon \vartheta^{-1} \partial_\alpha (\sigma^{-1} \partial_\alpha \psi_c) K_0 \sigma^{-1} \partial_\alpha \vartheta (R_{-2} - R_2) &+ \varepsilon \vartheta^{-1} \partial_\alpha (\sigma^{-1} \partial_\alpha \psi_c) \\
+ \varepsilon \vartheta^{-1} \partial_\alpha (\sigma^{-1} \partial_\alpha \psi_c) &+ \varepsilon \vartheta^{-1} \partial_\alpha (\sigma^{-1} \partial_\alpha \psi_c) \\
- \varepsilon^2 \vartheta^{-1} \partial_\alpha (\partial_\alpha^{-2} g_{\psi, \mathcal{R}_2} \vartheta R_{\mp 2}) &+ \varepsilon \vartheta^{-1} \partial_\alpha (\partial_\alpha^{-2} g_{\psi, \mathcal{R}_2} \vartheta R_{\mp 2}) \\
+ \varepsilon^2 \vartheta^{-1} \partial_\alpha ((K_0 \sigma^{-1} \partial_\alpha^{-1} g_{\psi, \mathcal{R}_2} + c(\psi, \mathcal{R})) \sigma^{-1} \vartheta (R_{-2} - R_2)) &+ \varepsilon \vartheta^{-1} \partial_\alpha ((K_0 \sigma^{-1} \partial_\alpha^{-1} g_{\psi, \mathcal{R}_2} + c(\psi, \mathcal{R})) \sigma^{-1} \vartheta (R_{-2} - R_2))
\end{align*} \quad (134)$$
\[ -\frac{\varepsilon^2}{2} b \partial^{-1} \partial_\alpha (\sigma^{-1} g_-(\Psi_2^h, \mathcal{R}_2) K_0 \sigma^{-1} \partial_\alpha (R_2 - R_2)) \]
\[ + \frac{\varepsilon^2}{2} \partial^{-1} \partial_\alpha (\sigma, \partial_\alpha^{-2} \partial (R_2 - R_2) \sigma^{-1} g_-(\Psi_2^h, \mathcal{R}_2)) \]
\[ + \varepsilon^2 \mathcal{M}_{\mp 2}(\Psi, \mathcal{R}) + \varepsilon^{-5/2} \partial^{-1} \text{res}_{\mp 2}(\varepsilon \Psi) \] (134)

as well as
\[ \partial_\alpha^{-1} \sigma^{-1} (R_2 - R_2) = \sigma^{-1} (R_2 - R_2) + \varepsilon \mathcal{M}_{-3}(\Psi, \mathcal{R}) + \varepsilon^{-5/2} \partial^{-1} \text{res}_{-3}(\varepsilon \Psi), \] (135)
\[ \partial_\alpha^{-2} (R_2 - R_2) = R_2 + 1 \]
\[ = -\varepsilon \partial^{-1} \partial_\alpha^{-1} ((\sigma^{-1} \partial_\alpha^2 \Psi c) K_0 \partial (R_2 - R_2)) \]
\[ - \varepsilon \partial^{-1} \partial_\alpha^{-1} ((K_0 \Psi c) \sigma^{-1} \partial (R_2 - R_2)) \]
\[ + \varepsilon \mathcal{M}_3(\Psi, \mathcal{R}) + \varepsilon^{-5/2} \partial^{-1} \text{res}_3(\varepsilon \Psi), \] (136)

where
\[ g_{\pm}(\Psi_2^h, \mathcal{R}_2) = \psi_{\pm}^h \pm \psi_{\pm}^h + \varepsilon^{1/2} \partial (R_\pm - R_\pm), \] (137)

\[ B_{j_1 j_2} \text{ with } j_1 \in \{ \pm 1, \pm 2 \} \text{ and } j_2 \in \{ \pm j_1 \} \text{ are bilinear real-valued mappings,} \]
\[ C_{j_1 m n j_2} \text{ with } j_1, m, n, j_2 \in \{ \pm 1 \} \text{ trilinear real-valued mappings as well as} \]
\[ M_j \text{ with } j \in \{ \pm 1, \pm 2, \pm 3 \} \text{ and } c \text{ nonlinear real-valued functions which satisfy} \]
\[ \| B_{\mp 1 j_2}(\Psi_c, \partial R_{j_2}) \|_{H^1} \lesssim \| R_{j_2} \|_{L^2}, \] (138)

as long as \( \varepsilon^{5/2} \| R_{j_2} \|_{L^2} \lesssim 1, \)
\[ \| \partial \partial B_{\mp 2 j_2}(\psi_c, \partial R_{j_2}) \|_{H^s} \lesssim \| R_{-1} \|_{L^2} + \| R_1 \|_{L^2} + \| R_2 \|_{H^s} + \| R_2 \|_{H^s} \] (139)
for \( s \geq 2, \) as long as \( \varepsilon^{5/2} \| R_{\mp 1} \|_{L^2}, \varepsilon^{5/2} \| R_{\mp 2} \|_{H^s} \lesssim 1, \)
\[ \| C_{j_1 m n j_2}(\psi_m, \psi_n, \partial R_{j_2}) \|_{H^1} \lesssim \| R_{j_2} \|_{L^2} + \| R_{j_2} \|_{H^{3/2}}, \] (140)

as long as \( \varepsilon^{5/2} \| R_{j_2} \|_{L^2}, \varepsilon^{5/2} \| R_{j_2} \|_{H^{3/2}} \lesssim 1, \)
\[ \| M_{\mp 1}(\Psi, \mathcal{R}) \|_{H^1} \lesssim \| R_{-1} \|_{L^2} + \| R_1 \|_{L^2} + \| R_2 \|_{H^{3/2}} + \| R_2 \|_{H^{3/2}}, \] (141)

as long as \( \varepsilon^{5/2} \| R_{\mp 1} \|_{L^2}, \varepsilon^{5/2} \| R_{\mp 2} \|_{H^{3/2}} \lesssim 1, \)
\[ \| M_{\mp 2}(\Psi, \mathcal{R}) \|_{H^1} \lesssim \| R_{-1} \|_{L^2} + \| R_1 \|_{L^2} + \| R_2 \|_{H^{3/2}} + \| R_2 \|_{H^{3/2}} \] (142)

for \( s \geq 2, \) as long as \( \varepsilon^{5/2} \| R_{\mp 1} \|_{L^2}, \varepsilon^{5/2} \| R_{\mp 2} \|_{H^s} \lesssim 1, \)
\[ \| M_{-3}(\Psi, \mathcal{R}) \|_{H^s} \lesssim \| R_{-1} \|_{L^2} + \| R_1 \|_{L^2} + \| R_2 \|_{H^{s-1}} \] (143)

for \( s \geq 2, \) as long as \( \varepsilon^{5/2} \| R_{\pm 1} \|_{L^2}, \varepsilon^{5/2} \| \sigma^{-1} (R_2 - R_2) \|_{H^{s-1}} \lesssim 1, \)
\[ \| \partial \partial M_3(\Psi, \mathcal{R}) \|_{C^0} + \| \partial \partial M_3(\Psi, \mathcal{R}) \|_{H^s} \lesssim \| R_{-1} \|_{L^2} + \| R_1 \|_{L^2} + \| \sigma^{-1} (R_2 - R_2) \|_{H^{s-2}} \] (144)
for $s \geq 2$, as long as $\varepsilon^{5/2} \|R_{\pm 1}\|_{L^2}, \varepsilon^{5/2} \|\sigma^{-1}(R_{-2} - R_2)\|_{H^{s-1}}, \varepsilon^{5/2} \|R_{-2} + R_2\|_{H^{s-2}} \ll 1$ and

$$\|c(\Psi, R)\|_{H^s} \leq C(\|R_{-1}\|_{L^2}, \|R_{-2}\|_{H^{s-1}}, \|R_2\|_{H^{s-1}})$$

(145)

for $s \geq 6$. All bounds are uniform with respect to $b \lesssim 1$ and $\varepsilon \ll 1$.

Moreover, (135)–(136), (143)–(144) and Lemma 3.1 imply

$$\|R_{-1}\|_{H^s} + \|R_1\|_{H^s} \lesssim \|R_{-1}\|_{L^2} + \|R_1\|_{L^2} + \|R_{-2}\|_{H^{s-2}} + \|R_2\|_{H^{s-2}}$$

$$+ \varepsilon \|\partial_s^{-1} R_{-2}\|_{L^2} + \varepsilon \|\partial_s^{-1} R_2\|_{L^2} + \varepsilon^2$$

(146)

for $s \geq 3$, as long as $\varepsilon^{5/2} \|R_{\pm 1}\|_{L^2}, \varepsilon^{5/2} \|R_{\pm 2}\|_{H^{s-1}} \ll 1,$

$$\|\partial_s^{-1} (R_{-2} - R_2)\|_{H^s} \lesssim \|R_{-1}\|_{L^2} + \|R_1\|_{L^2} + \|R_{-2}\|_{H^{s-1}} + \|R_2\|_{H^{s-1}} + \varepsilon^2$$

(147)

for $s \geq 3$, as long as $\varepsilon^{5/2} \|R_{\pm 1}\|_{L^2}, \varepsilon^{5/2} \|R_{\pm 2}\|_{H^{s-1}} \ll 1,$ and

$$\|\partial_s^{-2} \partial (R_{-2} + R_2)\|_{C^0} + \|\partial_s^{-1} (R_{-2} + R_2)\|_{H^s} \lesssim \|R_{-1}\|_{L^2} + \|R_1\|_{L^2}$$

$$+ \|R_{-2}\|_{H^{s-1}} + \|R_2\|_{H^{s-1}} + \varepsilon^2$$

(148)

for $s \geq 3$, as long as $\varepsilon^{5/2} \|R_{\pm 1}\|_{L^2}, \varepsilon^{5/2} \|R_{\pm 2}\|_{H^{s-1}} \ll 1$. These bounds are also uniform with respect to $b \lesssim 1$ and $\varepsilon \ll 1$.

Local existence and uniqueness of solutions $R$ to (133)–(136) in $(L^2(\mathbb{R}, \mathbb{R}))^2 \times (H^s(\mathbb{R}, \mathbb{R}))^2$ with $s + 2 = s_A \geq 10$ follows directly from the local existence and uniqueness results in Sobolev spaces for the arc length formulation of the two-dimensional water wave problem (52)–(58) and the NLS equation.

Now, we discuss the structure of the above evolution equations for the error $R$. These equations are of the form

$$\partial_t R_j = L R_j + \varepsilon Q_j(\psi_c) R_j + \varepsilon^2 W_j(\Psi, R) + \varepsilon^{-\beta} \partial^{-1} \text{Res}_j(\varepsilon \Psi)$$

(149)

for $j \in \{1, 2\}$, with linear operators $L$ and $Q_j(\psi_c)$ and nonlinear functions $W_j$ having the following properties. $L$ can be represented by the diagonal matrix

$$L = \text{diag}(-i\omega, i\omega).$$

(150)

The operators $Q_j(\psi_c)$ are of the form

$$Q_1(\psi_c) R_1 = \left( \begin{array}{cc} Q_{-1-1}(\psi_c) & Q_{-11}(\psi_c) \\ Q_{1-1}(\psi_c) & Q_{11}(\psi_c) \end{array} \right) \left( \begin{array}{c} R_{-1} \\ R_1 \end{array} \right)$$

$$+ \left( \begin{array}{cc} C_{-1-1}(\psi_c, \psi_c) & C_{-11}(\psi_c, \psi_c) \\ C_{1-1}(\psi_c, \psi_c) & C_{11}(\psi_c, \psi_c) \end{array} \right) \left( \begin{array}{c} R_{-1} \\ R_1 \end{array} \right)$$

(151)

and

$$Q_2(\psi_c) R_2 = \left( \begin{array}{cc} Q_{-2-2}(\psi_c) & Q_{-22}(\psi_c) \\ Q_{2-2}(\psi_c) & Q_{22}(\psi_c) \end{array} \right) \left( \begin{array}{c} R_{-2} \\ R_2 \end{array} \right)$$

(152)
respectively, with

\[ Q_{j_1 j_2}(g) f = \sum_{\mu=1}^{2|j_1|+1} Q_{j_1 j_2}^\mu(g) f, \]

\[ (\hat{Q}_{j_1 j_2}^\mu(g) f)(k) = \int_{\mathbb{R}} \hat{\vartheta}^{-1}(k) \hat{q}_{j_1 j_2}^{[j_1,\mu]}(k, k - m, m) \hat{\vartheta}(k - m) \hat{\vartheta}(m) \hat{f}(m) dm, \]

\[ \hat{q}_{j_1 j_2}^{1,1}(k, k - m, m) = -\delta_{j_1 j_2} i k, \]

\[ \hat{q}_{j_1 j_2}^{1,2}(k, k - m, m) = \delta_{j_1 j_2} i k \hat{K}_0(k - m) \hat{K}_0(m), \]

\[ \hat{q}_{j_1 j_2}^{1,3}(k, l, m) \chi_c(l) \chi_c(k - m) = \begin{cases} \mathcal{O}(|k|) & \text{for } |k| \to 0, \\ \mathcal{O}(1) & \text{for } |k| \to \infty, \end{cases} \]

\[ \hat{q}_{j_1 j_2}^{2,1}(k, k - m, m) = -\delta_{j_1 j_2} i k, \]

\[ \hat{q}_{j_1 j_2}^{2,2}(k, k - m, m) = -\frac{1}{2} \operatorname{sgn}(j_2) i k \hat{K}_0(k - m) \sigma^{-1}(k - m) i (k - m) \sigma^{-1}(m), \]

\[ \hat{q}_{j_1 j_2}^{2,3}(k, k - m, m) = -\frac{b}{2} \operatorname{sgn}(j_2) i k \sigma^{-1}(k - m)(k - m)^2 \hat{K}_0(m) \sigma^{-1}(m) i m, \]

\[ \hat{q}_{j_1 j_2}^{2,4}(k, k - m, m) = \frac{1}{2} \operatorname{sgn}(j_1) i k i \frac{\sigma(k) - \sigma(k - m)}{k - (k - m)} \sigma^{-1}(k - m)(k - m)^2 (im)^{-1}, \]

\[ \hat{q}_{j_1 j_2}^{2,5}(k, l, m) \chi_c(l) \chi_c(k - m) = \begin{cases} \mathcal{O}(|k|) & \text{for } |k| \to 0, \\ \mathcal{O}(1) & \text{for } |k| \to \infty, \end{cases} \]

\[ \partial_n \hat{q}_{j_1 j_2}^{2,5}(k, l, m) \chi_c(l) \chi_c(k - m) = \mathcal{O}(|k|^{-1}) \text{ for } |k| \to \infty \text{ and } n \in \{1, 3\}, \]

where the bounds are uniform with respect to \( m \in \mathbb{R} \) and \( b \lesssim 1 \), and

\[ C_{j_1 j_2}(g, h) f = \sum_{m,n \in \{\pm 1\}} \varepsilon \vartheta^{-1} C_{j_1 mn j_2}(g_m, h_n, \vartheta f). \]

Here, \( \partial_n \) denotes the partial derivative with respect to the \( n \)th variable and the functions \( p_\ell \) with \( p \in \{g, h\} \) and \( \ell \in \{m, n\} \), are defined by \( \hat{p}_\ell = \hat{p} \chi_{\mathbb{R}_0^\ell} \) if \( \ell = 1 \) and \( \hat{p}_\ell = \hat{p} \chi_{\mathbb{R}_0^\ell} \) if \( \ell = -1 \). By using (86)–(91) and the Taylor expansion of \( \sigma \) as function of \( m \) around \( m = k \), the symbols \( \hat{q}_{j_1 j_2}^{1,3}, \hat{q}_{j_1 j_2}^{2,5} : \mathbb{R}^3 \to i\mathbb{R} \) can be computed explicitly. But for simplicity we only present those properties of these symbols that we need for the proof of the error estimates.

For later purposes we set

\[ \hat{q}_{j_1 j_2}(k, k - m, m) := \sum_{\mu=1}^{2|j_1|+1} \hat{q}_{j_1 j_2}^{[j_1,\mu]}(k, k - m, m). \]
The symbols $\hat{q}^{2,\mu}_{j_1,j_2}$ with $\mu \in \{1, 2, 3, 4\}$ have the following symmetry properties, which will be essential for the proof of our error estimates. It holds that

$$\hat{q}^{2,\mu}_{j_1,j_2}(-k, k-m, -m) = -\hat{q}^{2,\mu}_{j_1,j_2}(k, k-m, m),$$

$$\hat{q}^{2,\mu}_{j,j_j}(k, k-m, m) = -\hat{q}^{2,\mu}_{j,j_j}(k, k-m, m)$$

for all $j_1, j_2 \in \{\pm 2\}$, $\mu \in \{1, 2, 3, 4\}$, $k \in \mathbb{R}$ and $m \in \mathbb{R} \setminus \{0\}$.

The functions $W_j(\Psi, R)$ are of the form

$$W_1(\Psi, R) = \left( \mathcal{M}_{-1}(\Psi, R) \right),$$

and

$$W_2(\Psi, R) = \left( \begin{array}{cc} \mathcal{M}_{-2}(\Psi, R) & W_{22}(\Psi, R) \\ W_{22}(\Psi, R) & R_2 \end{array} \right) + \left( \begin{array}{c} \mathcal{M}_{-2}(\Psi, R) \\ \mathcal{M}_{2}(\Psi, R) \end{array} \right)$$

where

$$W_{j_1,j_2}(\Psi, R)_{j_2} = \sum_{\mu=1}^{4} W^{\mu}_{j_1,j_2}(\Psi, R)_{j_2},$$

$$W_{1}^{1}_{j_1,j_2}(\Psi, R)_{j_2} = Q_{j_1,j_2}^{1}(\hat{\alpha}_{-2}^2 g_+ (\Psi_2^h, R_2)) R_{j_2},$$

$$W_{2}^{2}_{j_1,j_2}(\Psi, R)_{j_2} = Q_{j_1,j_2}^{2}(\hat{\alpha}_{-2}^2 g_- (\Psi_2^h, R_2) + (k_0 \sigma^{-1}\hat{\alpha})^{-1}c(\Psi, R)) R_{j_2},$$

$$W_{3}^{3}_{j_1,j_2}(\Psi, R)_{j_2} = Q_{j_1,j_2}^{3}(\hat{\alpha}_{-2}^2 g_- (\Psi_2^h, R_2)) R_{j_2},$$

$$W_{4}^{4}_{j_1,j_2}(\Psi, R)_{j_2} = Q_{j_1,j_2}^{4}(\hat{\alpha}_{-2}^2 g_- (\Psi_2^h, R_2)) R_{j_2}.$$
By the mean value theorem we have
\[ \hat{r}_{j_1,j_2}(k, k - m, m, b) \]
\[ = i\left(j_1 \partial_{k} \omega(\theta_0(k, b)k, b) + \partial_{k} \omega(k - m - \theta_1(k, m, b)k, b)\right)k - i(1 + j_2)\omega(m, b), \]  
(178)
with \( \theta_0(k, b), \theta_1(k, m, b) \in [0, 1], \) for all \( k, m \in \mathbb{R}. \) Since \( k \mapsto \partial_k \omega(k, b) \) is a continuous even function which satisfies (21) for all \( b \in \mathbb{R}_0^+ \setminus \{b_*\}, \) where \( b_* \in (0, 1/3), \) there exist a function \( \delta_0 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, (0, k_0/20)) \) such that \( (k, m) \mapsto \hat{r}_{j_1,j_2}(k, k - m, m, b) \) has zeros satisfying \(|k| \leq \delta_0(b)\) and \(|k - m + k_0| \leq \delta_0(b)\) if and only if \( j_2 = -1 \) and then the zeros are \((0, \mp k_0)\). Moreover, there exist a function \( \gamma_0 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, \mathbb{R}^+) \) such that
\[ |\hat{r}_{\pm 1-1}(k, k - m, m, b)| \geq \gamma_0(b)|k| \]  
(179)
for all \((k, m) \in \mathbb{R}^2\) with \(|k| \leq \tilde{\delta}_0(b)\) and \(|k - m + k_0| \leq \tilde{\delta}_0(b)\).

Next, we analyze the zeros of \( k \mapsto \hat{r}_{j_1,j_2}(k, \pm k_0, k \mp k_0, b) \). We have
\[ \hat{r}_{\pm 1-1}(0, \pm k_0, \mp k_0, b) = 0 \]  
(180)
for all \( b \geq 0. \) By the mean value theorem we obtain
\[ \hat{r}_{\pm 1-1}(k, \pm k_0, k \mp k_0, b) = i\left(j_1 \partial_{k} \omega(\theta_0(k, b)k, b) + \partial_{k} \omega(\pm k_0 - \theta_1(k, b)k, b)\right)k \]  
(181)
with \( \theta_0(k, b), \theta_1(k, b) \in [0, 1], \) for all \( k \in \mathbb{R}. \) Hence, there exist functions \( \gamma_1 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, \mathbb{R}^+) \) and \( \delta_1 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, (0, k_0/20)) \) such that
\[ |\hat{r}_{\pm 1-1}(k, \pm k_0, k \mp k_0, b)| \geq \gamma_1(b)|k| \]  
(182)
for \(|k| \leq \tilde{\delta}_1(b)\). Moreover, we have
\[ \hat{r}_{-1+1}(\pm k_0, \pm k_0, 0, b) = 0 \]  
(183)
for all \( b \geq 0. \) Using the mean value theorem again we obtain
\[ \hat{r}_{-1+1}(k, \pm k_0, k \mp k_0, b) \]
\[ = -i\left(\partial_{k} \omega(\pm k_0 + \theta_0(k, b)(k \mp k_0), b) + j_2 \partial_{k} \omega(\theta_1(k, b)(k \mp k_0), b)\right)(k \mp k_0) \]  
(184)
with \( \theta_0(k, b), \theta_1(k, b) \in [0, 1], \) for all \( k \in \mathbb{R}. \) Hence, there exist functions \( \gamma_2 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, \mathbb{R}^+) \) and \( \delta_2 \in C^0(\mathbb{R}_0^+ \setminus \{b_*\}, (0, k_0/20)) \) such that
\[ |\hat{r}_{-1+1}(k, \pm k_0, k \mp k_0, b)| \geq \gamma_2(b)|k \mp k_0| \]  
(185)
for \(|k \mp k_0| \leq \tilde{\delta}_2(b)\).

Since \( k \mapsto \omega(k) \) is strictly monotonically increasing, \( k \mapsto \hat{r}_{j_1,j_2}(k, \pm k_0, k \mp k_0, b) \) has no other zeros if \( j_1 = 1 \) or \( j_2 = 1. \) As discussed in the introduction the remaining zeros of \( k \mapsto \hat{r}_{-1-1}(k, \pm k_0, k \mp k_0, b) \) can be determined by analyzing the zeros of the function \( \hat{r}: \mathbb{R} \times \mathbb{R}_0^+ \to \mathbb{R} \) with
\[ \hat{r}(k, b) = \omega(k, b) - \omega(k - k_0, b) - \omega(k_0, b). \]  
(186)
Because \( k \mapsto \omega(k, b) \) is odd, it holds that
\[
\hat{r}(k_0/2 + k, b) = \hat{r}(k_0/2 - k, b)
\] (187)
for all \( k \in \mathbb{R} \) and all \( b \in \mathbb{R}^+_0 \) such that it is sufficient to analyze \( \hat{r} \) for \( k \in [k_0/2, \infty) \).

In the following, we present a quantitative description of the behavior of \( \hat{r} \) that is illustrated in Fig. 2.

We have
\[
\partial_k \hat{r}(k, b) = \partial_k \omega(k, b) - \partial_k \omega(k - k_0, b) = \partial_k \omega(k, b) - \partial_k \omega(k_0 - k, b)
\] (188)
for all \( k \in [k_0/2, \infty) \) and all \( b \in \mathbb{R}^+_0 \).

Using
\[
\tanh(k) = \text{sgn}(k) + O(e^{-2|k|}),
\] (189)
\[
\frac{d}{dk} \tanh(k) = O(e^{-2|k|})
\] (190)
for \( |k| \to \infty \), we deduce
\[
\omega(k, b) = \text{sgn}(k) |k|^{1/2}(1 + b k^2)^{1/2} + O(e^{-|k|}),
\] (191)
\[
\partial_k \omega(k, b) = \frac{1 + 3bk^2}{2(1 + bk^2)} |k|^{-1/2}(1 + b k^2)^{1/2} + O(e^{-|k|}),
\] (192)
\[
\partial_k^2 \omega(k, b) = O(|k|^{-3/2}(1 + b k^2)^{1/2})
\] (193)
for \( |k| \to \infty \) uniformly with respect to \( b \lesssim 1 \). By Taylor’s theorem we have
\[
\omega(k, b) - \omega(k - k_0, b) = \partial_k \omega(k, b) k_0 - \frac{1}{2} \partial_k^2 \omega(k + \theta(k)(k - k_0), b) k_0^2
\] (194)
with \( \theta(k) \in [0, 1] \). Due to (191)–(194), we obtain
\[
\hat{r}(k, 0) \to -\omega(k_0, 0) \quad \text{for} \quad k \to \infty,
\] (195)
and for all \( b > 0 \),
\[
\hat{r}(k, b) \to \infty \quad \text{for} \quad k \to \infty.
\] (196)
If \( b \geq 1/3 \), then there holds \( \partial_k^2 \omega(k, b) > 0 \) for all \( k > 0 \). Because of \( \omega(0, b) = 0 \) this implies \( \omega(k_0/2, b) < \omega(k_0, b)/2 \) for all \( b \geq 1/3 \) and therefore \( \hat{r}(k_0/2, b) < 0 \) for all \( b \geq 1/3 \). Due to (188), the positivity of \( \partial_k^2 \omega \) also yields \( \partial_k \hat{r}(k, b) > 0 \) for all \( k > k_0/2 \) and all \( b \geq 1/3 \). Moreover, since \( \partial_k^2 \hat{r} \) is continuous with respect to \( k \) and \( b \) for any \( n \in \mathbb{N}_0 \), there exist a constant \( b_1 \in (0, 1/3) \) and functions \( \gamma_j \in C^0((b_1, \infty), \mathbb{R}^+) \), \( j \in \{3, 4, 5\} \), and \( \tilde{\delta}_3 \in C^0((b_1, \infty), (0, k_0/20)) \) such that we have
\[
\hat{r}(k, b) \leq -\gamma_3(b)
\] (197)
for all \( b \in (b_1, \infty) \) and all \( k \in [k_0/2, k_0 - \tilde{\delta}_3(b)] \),
\[
\partial_k \hat{r}(k, b) \geq \gamma_4(b)
\] (198)
for all $b \in (b_1, \infty)$ and all $k \in [k_0 - \tilde{\delta}_3(b), k_0 + \tilde{\delta}_3(b)]$, as well as
\[ \hat{r}(k, b) \geq \gamma_5(b) \] (199)
for all $b \in (b_1, \infty)$ and all $k \in [k_0 + \tilde{\delta}_3(b), \infty)$; compare Fig. 2, Panel (i)–(iii).

If $b = 0$, then there holds $\partial_k^2 \omega(k, b) < 0$ for all $k > 0$, which implies $\omega(k_0/2, 0) > \omega(k_0, 0)/2$ and therefore $\hat{r}(k_0/2, 0) > 0$. Due to (188), the negativity of $\partial_k^2 \omega$ also yields $\partial_k \hat{r}(k, 0) < 0$ for all $k > k_0/2$. Hence, there exist $b_2 \in (0, b_1)$, $\gamma_j \in C^0([0, b_2], \mathbb{R}^+)$ for $j \in \{6, 7\}$, $\gamma_8 \in \mathbb{R}^+$ and $\tilde{\delta}_4 \in C^0([0, b_2], (0, k_0/20))$ such that there holds
\[ \hat{r}(k, b) \geq \gamma_6(b) \] (200)
for all $b \in [0, b_2)$ and all $k \in [k_0/2, k_0 - \tilde{\delta}_4(b)]$,
\[ \partial_k \hat{r}(k, b) \leq -\gamma_7(b) \] (201)
for all $b \in [0, b_2)$ and all $k \in [k_0 - \tilde{\delta}_4(b), k_0 + \tilde{\delta}_4(b)]$, as well as
\[ \hat{r}(k, 0) \leq -\gamma_8 \] (202)
for all $k \in [k_0 + \tilde{\delta}_4(b), \infty)$; compare Fig. 2, Panel (vi)–(ix).

Because of $\hat{r}(k_0, b) = 0$, (196), (201) and the intermediate value theorem there exist functions $k_1, k_2 \in C^0((0, b_2), (k_0, \infty))$ with
\[ \hat{r}(k_1(b), b) = 0, \] (203)
\[ \partial_k \hat{r}(k_2(b), b) = 0 \] (204)
for all $b \in (0, b_2)$.

Since there exists a strictly monotonically decreasing function $k_3 \in C^0((0, 1/3), (0, \infty))$ with $k_3(b) \to 0$ for $b \to 1/3$ and $k_3(b) \to \infty$ for $b \to 0$ such that for all $b \in (0, 1/3)$ there holds
\[ \partial_k^3 \omega(k, b) < 0 \quad \text{if} \quad 0 < k < k_3(b), \] (205)
\[ \partial_k^3 \omega(k, b) = 0 \quad \text{if} \quad k = k_3(b), \] (206)
\[ \partial_k^3 \omega(k, b) > 0 \quad \text{if} \quad k > k_3(b), \] (207)
the mean value theorem yields $k_2(b) > k_3(b)$ for all $b \in (0, b_2)$. Moreover, because there exists a unique function $k_4 \in C^0((0, 1/3))$ with $k_4(b) > k_3(b)$ and
\[ \partial_k^3 \omega(k_4(b), b) = 0 \] (208)
for all $b \in (0, 1/3)$, and since
\[ \partial_k^2 \hat{r}(k, b) = \partial_k^2 \omega(k, b) - \partial_k^2 \omega(k - k_0, b) \] (209)
for all $k \in [k_0/2, \infty)$ and all $b \in \mathbb{R}^+_0$, the function $k \mapsto \partial_k^2 \hat{r}(k, b)$ can have at most one zero on $(k_2(b), \infty)$ for all $b \in (0, b_2)$. Because of $\hat{r}(k_0, b) = 0$, (196), (201) and the mean value theorem it follows that the function $k_2$ and therefore also the function $k_1$ is unique.
Let \( \tilde{k}(c, b) = 4c^2 \tanh(k_0)/9k_0b \) for \( c > 0. \) Using (192)–(194) yields

\[
\lim_{b \to 0} \partial_b \omega(\tilde{k}(c, b), b) k_0 = c \left( k_0 \tanh(k_0) \right)^{1/2} = c \omega(k_0, 0),
\]

which implies

\[
k_1(b) = \frac{4 \tanh(k_0)}{9k_0b} (1 + o(1)) \quad \text{for } b \to 0.
\]

Moreover, due to (155)–(163), there exist \( b_0 \in (0, b_2] \) and functions \( C_{TWI} \in C^0([0, b_0), (1, \infty)), \delta_1 \in C^0([0, b_0), (0, 1)) \) with \( \delta_1(b) < 1 - (20(k_1(b) - k_0))^{-1}k_0 \) for all \( b \in (0, b_0), \)

\[
\lim_{b \to 0} \delta_1(b) \geq \frac{1}{2},
\]

\[
\lim_{b \to 0} C_{TWI}(b) = 1
\]

such that

\[
\frac{1}{C_{TWI}(b)} \leq \frac{-\hat{q}_{j_1j_1}(\pm k, \pm k_0, \pm(k - k_0))}{\hat{q}_{j_1j_1}(\mp(k - k_0), \pm k_0, \mp k)} \leq C_{TWI}(b)
\]

for all \( j_1 \in \{-1, -2\}, b \in (0, b_0) \) and \( k \in [k_1(b) - \delta_1(b)(k_1(b) - k_0), k_1(b) + \delta_1(b)(k_1(b) - k_0)] \). Furthermore, there exist functions \( \gamma_j \in C^0([0, b_0), \mathbb{R}^+), j \in \{9, 10\}, \) such that

\[
\hat{r}(k, b) \leq -\gamma_9(b).
\]

for all \( b \in (0, b_0) \) and all \( k \in [k_0 + \tilde{\delta}_4(b), k_1(b) - \delta_1(b)(k_1(b) - k_0)/2], \) as well as

\[
\hat{r}(k, b) \geq \gamma_{10}(b).
\]

for all \( b \in (0, b_0) \) and all \( k \in [k_1(b) + \delta_1(b)(k_1(b) - k_0)/2, \infty); \) compare Fig. 2, Panel (vii)–(viii).

By using the method from [37] to compute the values of the coefficients in the TWI systems belonging to the two-dimensional water wave problem with finite depth in the Eulerian formulation and in the arc length formulation, one can additionally show that one can choose \( b_0 = b_2. \) We remark that this property of \( b_0 \) is not needed for the proof of our error estimates.

Finally, we define

\[
\mathcal{B} := \{b \in [0, b_0) \cup (b_1, \infty) : k \mapsto \omega(k, b) \text{ satisfies (21), (23) and (24) with } M = 6.\}
\]

and \( \delta_0 = \delta_0(b) \) by

\[
\delta_0(b) = \begin{cases} 
\min\{\tilde{\delta}_j(b) : j \in \{0, 1, 2, 3\}\} & \text{if } b \in (b_1, \infty), \\
\min\{\tilde{\delta}_j(b) : j \in \{0, 1, 2, 4\}\} & \text{if } b \in [0, b_0).
\end{cases}
\]

Now, we are prepared to prove
Theorem 4.1. Let $k_0 > 0$ and $s_A \geq 10$. Then for all $\tau_0, C_0 > 0$ there exists an $\varepsilon_0 > 0$ and a function $C \in C^0(\mathcal{B}, \mathbb{R}^+)$ such that for all $b \in \mathcal{B}$ and all solutions $\tilde{A}_{-11}^0 \in C^0([0, \tau_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ of the NLS equation (107) with

$$\sup_{\tau \in [0, \tau_0]} \|\tilde{A}_{-11}^0(\cdot, \tau)\|_{H^{s_A}(\mathbb{R}, \mathbb{C})} \leq C_0$$

the following holds: for all $\varepsilon \in (0, \varepsilon_0)$ there exists a solution

$${\mathcal U} \in C^0([0, \tau_0 e^{-2}], (L^2(\mathbb{R}, \mathbb{R}))^2 \times (H^s(\mathbb{R}, \mathbb{R}))^2)$$

of (86)–(91), where $s = s_A - 2$, which satisfies

$$\sup_{\tau \in [0, \tau_0 e^{-2}]} \left\|{\mathcal U}(\cdot, \tau) - \varepsilon \left(\frac{\tilde{\Psi}_0^0}{\partial_\alpha^2 \tilde{\Psi}_0^1}(\cdot, \tau)\right)\right\|_{(L^2)^2 \times (H^s)^2} \leq C(b)\varepsilon^{3/2},$$

where

$$\tilde{\Psi}_1^0(\alpha, t) = \tilde{A}_{-11}^0(\varepsilon(\alpha - c_s t), \varepsilon^2 t) E_1^0\left(\begin{array}{c} 1 \\ 0 \end{array}\right) + \text{c.c.}$$

4.2. The Construction of the Energy

For the proof of Theorem 4.1 we introduce the energy

$$\mathcal{E}_\tilde{s} = \mathcal{E}_{1,0} + \mathcal{E}_{2,\tilde{s}}$$

(217)

with $0 \leq \tilde{s} \leq s := s_A - 2$,

$$\mathcal{E}_{2,\tilde{s}} = \sum_{l=0}^{\tilde{s}} \mathcal{E}_{2,l},$$

(218)

$$\mathcal{E}_{j,0} = \sum_{j_1 \in \{\pm j\}} \frac{1}{2} \int_{\mathbb{R}} \tilde{R}_{j_1} \rho_{j_1}^0 \tilde{R}_{j_1} \, d\alpha,$$

(219)

$$\tilde{R}_{j_1} = R_{j_1} + \varepsilon \sum_{j_2 \in \{\pm j_1\}} N_{j_1,j_2}(\psi_c, R_{j_2}) + \varepsilon^2 \sum_{j_2 \in \{\pm j_1\}} T_{j_1,j_2}(\psi_c, \psi_c, R_{j_2})$$

(220)

and

$$\mathcal{E}_{2,l} = \sum_{j_1 \in \{\pm 2\}} \left(\frac{1}{2} \int_{\mathbb{R}} \partial_a R_{j_1} \rho_{j_1}^l \partial_a R_{j_1} \, d\alpha + \varepsilon \sum_{j_2 \in \{\pm j_1\}} \int_{\mathbb{R}} \partial_a R_{j_1} \rho_{j_1}^l \partial_a N_{j_1,j_2}(\psi_c, R_{j_2}) \, d\alpha \right)$$

(221)

for $1 \leq l \leq s$.

Here, $\rho_{j_1}^l$ is defined by its symbol

$$\tilde{\rho}_{j_1}^l(k) = \begin{cases} 1 + \sum_{\ell \in \{\pm 1\}} \tilde{\rho}_{j_1 \ell}^l(k) & \text{if } b \in (0, b_0) \text{ and } \text{sgn}(j_1) = -1, \\ 1 & \text{otherwise}, \end{cases}$$

(222)
where
\[ \hat{\rho}^l_{j_1 \ell}(k) = \left( -\hat{q}_{j_1 j}(k, \ell k_0, k - \ell k_0) \left( \frac{-k + \ell k_0}{k} \right)^{2l} - 1 \right) \hat{\xi}_1 \left( \frac{k + \ell(k_1 - k_0)}{k_1 - k_0} \right), \]
(223)

\[ k_1 = k_1(b) \] is as above and \( \hat{\xi}_i \in C_c^\infty(\mathbb{R}, \mathbb{R}), i \in \{0, 1\}, \) satisfies
\[ \hat{\xi}_i(k) = \begin{cases} 1 & \text{if } |k| \leq \delta_i/2, \\ 0 & \text{if } |k| \geq \delta_i, \\ \hat{\xi}_i(|k|) & \text{otherwise} \end{cases} \]
(244)

with \( \delta_i = \delta_i(b) \) as above. Because of (214) there exist a constant \( C_\rho \geq 1 \) such that there holds
\[ C_\rho^{-1} \leq \hat{\rho}^l_{j_1} \leq C_\rho \]
(225)

for all \( j_1 \in \{\pm 1, \pm 2\} \) and all \( 0 \leq l \leq s \) uniformly on compact subsets of \( B \).

The functions \( N_{j_1 j_2}, \tilde{N}_{j_1 j_2} \) and \( \tilde{T}_{j_1 j_2} \) are defined as follows. We set
\[ N_{j_1 j_2}(\varphi, f) = \begin{cases} N^1_{j_1 j_2}(\varphi, f) & \text{if } |j_1| = 1, \\ N^2_{j_1 j_2}(\varphi, f) & \text{if } |j_1| = 2, \end{cases} \]
(226)

where
\[ N^j_{j_1 j_2}(\varphi, g) = \sum_{\ell \in \{\pm 1\}} N^j_{j_1 j_2 \ell}(\varphi, g) \]
(227)

with
\[ \hat{N}_{j_1 j_2 \ell}(\varphi, g, k) = \int_{\mathbb{R}} \hat{n}_{j_1 j_2 \ell}(k, \varphi \omega(k - m)) g(m) dm, \]
(228)

\[ \hat{n}_{j_1 j_2 \ell}(k) = \frac{\tilde{q}_{j_1 j}(k, \ell k_0, k - \ell k_0)}{\tilde{r}_{j_1 j}(k, \ell k_0, k - \ell k_0)} \hat{\xi}_{j_1 j_2 \ell}(k) \left( \hat{\vartheta} - \varepsilon \hat{\xi}_0(k) k - \ell k_0 \right), \]
(229)

\[ \tilde{q}_{j_1 j_2}(k, k - m, m) = \begin{cases} \tilde{q}_{j_1 j_2}(k, k - m, m) & \text{if } |j_1| = 1, \\ \tilde{q}_{j_1 j_2}(k, k - m, m) - \tilde{q}_{j_1 j_2}^2(k, k - m, m) & \text{if } |j_1| = 2, \end{cases} \]
(230)

\[ \tilde{r}_{j_1 j_2}(k, k - m, m) = i m \tilde{q}_{j_1 j_2}^2(k, k - m, m), \]
(231)

and
\[ \hat{\xi}_{j_1 j_2 \ell}(k) = \begin{cases} 1 - \hat{\xi}_1 \left( \frac{k - \ell k_1}{k_1 - k_0} \right) - \hat{\xi}_1 \left( \frac{k + \ell(k_1 - k_0)}{k_1 - k_0} \right) & \text{if } b \in (0, b_0) \text{ and } \text{sgn}(j_1) = \text{sgn}(j_2) = -1, \\ 1 & \text{otherwise}, \end{cases} \]
(233)

as well as
\[ N_{j_1 j_2}(\varphi, f) = \begin{cases} N^1_{j_1 j_2}(\varphi, f) & \text{if } |j_1| = 1, \\ N^2_{j_1 j_2}(\varphi, f) & \text{if } |j_1| = 2. \end{cases} \]
(234)
where

\[ \mathcal{N}_{j_1j_2}^i (\varphi, g) = \sum_{j=0}^{1} \mathcal{N}_{j_1j_2}^{i,j} (\varphi, g) \]  \hspace{1cm} (235)

with

\[ \hat{\mathcal{N}}_{j_1j_2}^{i,0} (\varphi, g)(k) = \int_{\mathbb{R}} \hat{\mathcal{N}}_{j_1j_2}^{i,0} (k, k - m, m) \hat{\varphi}(k - m) \hat{g}(m) \, dm, \]  \hspace{1cm} (236)

\[ \hat{n}_{j_1j_2}^{i,0} (k, k - m, m) = \hat{\mathcal{P}}_{0,\delta_0} (k) \frac{\hat{a}_{j_1j_2}^i (k, k - m, m)}{\hat{r}_{j_1j_2} (k, k - m, m)} \hat{g}(m) \]  \hspace{1cm} (237)

and

\[ \hat{\mathcal{N}}_{j_1j_2}^{i,1} (\varphi, g)(k) = \hat{\mathcal{P}}_{\delta,\infty} (k) \hat{\mathcal{N}}_{j_1j_2}^{i} (\varphi, g)(k). \]  \hspace{1cm} (238)

Moreover, we set

\[ \mathcal{T}_{j_1j_2} (g, h, f) = \sum_{j=|j_1|}^{2} \mathcal{T}_{j_1j_2}^j (g, h, f), \]  \hspace{1cm} (239)

where

\[ \mathcal{T}_{j_1j_2}^j (g, h, f) = \sum_{\ell \in \{\pm 1\}} \mathcal{T}_{j_1j_2}^j (g_\ell, h_\ell, f). \]  \hspace{1cm} (240)

\[ \tilde{\mathcal{T}}_{j_1j_2}^j (g_\ell, h_\ell, f)(k) = \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{\mathcal{T}}_{j_1j_2}^j (k, \ell k_0, \ell k_0, k - 2\ell k_0) \hat{g}(k - m) \hat{h}(m - n) \hat{f}(n) \, dm \, dn, \]  \hspace{1cm} (241)

\[ \tilde{\mathcal{T}}_{j_1j_2}^j (k) = \hat{\mathcal{P}}_{0,\delta_0} (k) \frac{\hat{c}_{j_1j_2 \ell j_2} (k, \ell k_0, \ell k_0, k - 2\ell k_0)}{\hat{c}_{j_1j_2 \ell j_2} (k, \ell k_0, \ell k_0, k - 2\ell k_0) \hat{g}(k - 2\ell k_0) \hat{h}(k - 2\ell k_0) \hat{f}(n)} \]  \hspace{1cm} (242)

\[ \tilde{\mathcal{T}}_{j_1j_2}^2 (k) = \sum_{j_2 \in \{\pm 1\}} \hat{\mathcal{P}}_{0,\delta_0} (k) \frac{\hat{a}_{j_1j_2j_2} (k, \ell k_0, \ell k_0, k - 2\ell k_0)}{\hat{r}_{j_1j_2j_2} (k, \ell k_0, \ell k_0, k - 2\ell k_0) \hat{g}(k) \hat{h}(k) \hat{f}(n)} \]  \hspace{1cm} (243)

\[ \tilde{\mathcal{V}}_{j_1j_2} (k, k - m, m - n, n) = i \left( \text{sgn}(j_1) \omega(k) + \omega(k - m) + \omega(m - n) - \text{sgn}(j_2) \omega(n) \right), \]  \hspace{1cm} (244)

\[ \tilde{\mathcal{V}}_{j_1j_2} (k, k - m, m - n, n) = \tilde{\mathcal{V}}_{j_1j_2} (k, k - m, m - n, n), \]  \hspace{1cm} (245)

\[ \tilde{\mathcal{A}}_{j_1j_2j_2} (k, k - m, m - n, n) = \tilde{\mathcal{A}}_{j_1j_2j_2} (k, k - m, m - n, n) \]  \hspace{1cm} (246)

and \( \tilde{c}_{j_1j_2} \) is given by

\[ \int_{\mathbb{R}} \int_{\mathbb{R}} \tilde{c}_{j_1j_2} (k, k - m, m - n, n) \hat{g}(k - m) \hat{h}(m - n) \hat{f}(n) \, dm \, dn = \tilde{c}_{j_1j_2} (g_\ell, h_\ell, f)(k). \]  \hspace{1cm} (247)

Finally, the functions \( p_\ell \) with \( p \in \{\varphi, g, h\} \) are defined by \( \hat{p}_\ell = \hat{\rho} \chi_{\mathbb{R}^+}^\ell \) if \( \ell = 1 \) and \( \hat{p}_\ell = \hat{\rho} \chi_{\mathbb{R}^-}^\ell \) if \( \ell = -1 \).

As explained in the introduction, the construction of the energy \( \mathcal{E}_s \) is inspired by the method of normal-form transforms, where the normal-form transform is incorporated directly in \( \mathcal{E}_s \). The normal-form transform we use is an extension of a normal-form transform of the form (35) and (40) in order to handle the non-trivial resonances being present in the two-dimensional water wave problem with finite depth.
The weight function \( \hat{\vartheta} \) and the correction function \( \hat{\varepsilon} \hat{\xi_0} \) are included to handle the non-trivial resonances at \( \pm k_0 \). The trilinear mappings \( T_{j_1 j_2} \) are constructed in such a way that they generate terms in the evolution equation of \( \mathcal{E}_s \) which cancel all the terms of order \( O(\varepsilon) \) in the evolution equation which are caused by the fact that \( \hat{\vartheta}^{-1} \) is of order \( O(\varepsilon^{-1}) \) for \( |k| \leq \delta_0 \). The weight functions \( \hat{\rho}_{j_1 \ell} \) and the correction functions \( \hat{\xi}_{j_1 j_2 \ell} \) are included to control the additional non-trivial resonances. Their form is motivated by the conserved quantity (45). The factor \( 1/(k_1 - k_0) \) in the definition of \( \hat{\xi}_{j_1 j_2 \ell} \) is chosen in such a way that we obtain error estimates which are uniform with respect to \( b \) as \( b \) and \( \varepsilon \) go to 0.

The following lemma will allow us to show that it is sufficient for our goals that \( \hat{\rho}_{j_1 \ell} \), \( \hat{n}_{j_1 j_2 \ell} \) and \( \hat{\tau}_{j_1 j_2 \ell} \) in \( \mathcal{E}_s \) depend only on \( k \) and not on \( k \) and \( m \) like the kernel (39).

**Lemma 4.2.** Let \( p \in \mathbb{R} \), \( g \in C^2(\mathbb{R}, \mathbb{C}) \) have a compactly supported Fourier transform, \( \hat{R} : \mathbb{R}^3 \to \mathbb{C} \cup \{\infty\} \) be uniformly bounded for all \((k,l,m) \in \mathbb{R}^3 \) with \( l - p, k - m - p \in \text{supp} \, \hat{g} \) and \( f \in H^s(\mathbb{R}, \mathbb{C}) \) for \( s \geq 0 \).

1. If \( l \mapsto \hat{R}(k,l,m) \) is Lipschitz continuous in some neighborhood of \( p \) with a Lipschitz constant \( L \) being independent of \( k \) and \( m \) if \( k - m - p \in \text{supp} \, \hat{g} \), then there exist constants \( \varepsilon_0 > 0 \) and \( C_2 > 0 \) with \( C_2 \lesssim L \| \hat{R} \|_{L^1(s+1)} \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) it holds that

\[
\left\| \int \left( \hat{R}(\cdot, -m, m) - \hat{R}(\cdot, p, m) \right) \varepsilon^{-1} \hat{g}(\cdot - \frac{m - p}{\varepsilon}) \hat{f}(m) \, dm \right\|_{L^2(s)} \leq C_2 \varepsilon \| f \|_{H^s}.
\]

(248)

2. If \( m \mapsto \hat{R}(k,m) \) is Lipschitz continuous for all \( m \in \mathbb{R} \) for which \( k - m - p \in \text{supp} \, \hat{g} \) with a Lipschitz constant \( L \) being independent of \( k \) and \( m \), then there exist a constant \( C_3 > 0 \) with \( C_3 \lesssim L \| \hat{R} \|_{L^1(s+1)} \) such that for all \( \varepsilon > 0 \) it holds that

\[
\left\| \int \left( \hat{R}(\cdot, -m, m) - \hat{R}(\cdot, -m, \cdot - p) \right) \varepsilon^{-1} \hat{g}(\cdot - \frac{m - p}{\varepsilon}) \hat{f}(m) \, dm \right\|_{L^2(s)} \leq C_3 \varepsilon \| f \|_{H^s}.
\]

(249)

**Proof.** The lemma is proven analogously as Lemma 3.5 in [14]. \( \square \)

In order to verify Theorem 4.1 we would like to prove that the energy \( \mathcal{E}_s \) controls the error \( \mathcal{R} \) in \( (L^2(\mathbb{R}, \mathbb{R}))^2 \times (H^s(\mathbb{R}, \mathbb{R}))^2 \) and remains bounded for all \( t \in [0, \tau_0 \varepsilon^{-2}] \). However, we will see that the right-hand side of the evolution equation for \( \mathcal{E}_s \) contains terms which cannot be estimated by a multiple of \( \varepsilon^2 (\mathcal{E}_s + 1) \) such that Gronwall’s inequality cannot be applied to obtain the desired bound. Nevertheless, these terms can be identified as time derivatives of time dependent integrals. Hence, by adding these integrals to \( \mathcal{E}_s \) we will obtain a new energy \( \tilde{\mathcal{E}}_s \) at the end of Section 4.4 and this energy will have the desired properties. We will prove

**Lemma 4.3.** For sufficiently small \( \varepsilon > 0 \), \( \tilde{\mathcal{E}}_s \) satisfies

\[
\frac{d}{dt} \tilde{\mathcal{E}}_s \lesssim \varepsilon^2 (\tilde{\mathcal{E}}_s + 1),
\]

(250)
\[ \tilde{E}_s \lesssim \| \mathcal{R}_1 \|_{(L^2)^2}^2 + \| \mathcal{R}_2 \|_{(H^s)^2}^2 + \varepsilon^5, \]  
\[ \| \mathcal{R}_1 \|_{(L^2)^2}^2 + \| \mathcal{R}_2 \|_{(H^s)^2}^2 \lesssim \tilde{E}_s + \varepsilon^5, \]

as long as \( \varepsilon^{5/2} \| \mathcal{R}_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| \mathcal{R}_2 \|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( B \).

With the help of the energy \( \tilde{E}_s \) it is also possible to give an alternative proof of local existence and uniqueness of solutions \( \mathcal{R} \) to (133)–(136) in Sobolev spaces without using the local existence and uniqueness results for the water wave problem (52)–(58). In this proof, the energy \( \tilde{E}_s \) takes the role of the energy used in the proof of the local existence and uniqueness result for the water wave problem in the arc length formulation in [3].

Because \( \tilde{E}_s \) remains \( O(1) \)-bounded for a timespan of order \( O(\varepsilon^{-2}) \), the solutions \( \mathcal{R} \) to (133)–(136) even exist for this long timespan and belong to \( (L^2(\mathbb{R}, \mathbb{R}))^2 \times (H^s(\mathbb{R}, \mathbb{R}))^2 \). The Fourier truncation procedure from Section 3 allows us to choose \( s = s_A - 2 \) such that the error \( \mathcal{R} \) has the same Sobolev regularity as the solution of the NLS equation.

Since the energy estimates (250)–(252) are only valid as long as \( \varepsilon^{5/2} \| \mathcal{R}_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| \mathcal{R}_2 \|_{(H^s)^2} \ll 1 \), they are not sufficient to guarantee global existence of solutions to (133)–(136) and global existence of small oscillating wave packet-like solutions to the two-dimensional water wave problem with finite depth by iterating the arguments of the proof of local existence in time.

For the three-dimensional water wave problem with infinite depth, for the three-dimensional water wave problem with finite depth and either gravity or surface tension as well as for the two-dimensional water wave problem with infinite depth and either gravity or surface tension it is possible to combine energy estimates of this type (so-called high-order energy estimates) with dispersive decay estimates to establish global existence of solutions for small data. A detailed explanation of this procedure is given, for example, in [30] and global existence results for the water wave problem obtained with the help of this procedure can be found, for example, in [2, 7, 16, 22, 23, 25, 26, 46–48].

However, for the two-dimensional water wave problem with infinite depth, gravity and surface tension the optimal dispersive decay rate is not strong enough to ensure global existence of solutions and for the two-dimensional water wave problem with finite depth as well as for the three-dimensional water wave problem with finite depth, gravity and surface tension such dispersive decay estimates cannot be expected because of the existence of solitary waves solutions, which are localized traveling waves of permanent form. In all these cases, the global existence of solutions is still an open problem. For a more detailed discussion of the existence issue for solutions to the water wave problem we refer, for example, to the review article [10] and the references therein.

### 4.3. Fundamental Properties of the Energy

Now, we show several fundamental properties of the operators \( N_{j_1,j_2}, \mathcal{N}_{j_1,j_2} \) and \( T_{j_1,j_2} \), which will be mandatory for the proof of our energy estimates.
Lemma 4.4. The operators $N^i_{j_1, j_2}$ have the following properties:

a) Fix $\varphi \in L^2(\mathbb{R}, \mathbb{R})$ with $\text{supp } \hat{\varphi} = \text{supp } \hat{\varphi}_c$. Then $f \mapsto N^i_{j_1}(\varphi, f)$ defines a continuous linear map from $H^1(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$ and $f \mapsto N^i_{j_1}(\varphi, f)$ a continuous linear map from $H^{(1-(|j_1|-1))/2}(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$. Furthermore, there exists a constant $C > 0$ with $C \lesssim \|\varphi\|_{L^1}$ such that for all $f \in H^{(1-(|j_1|-1))/2}(\mathbb{R}, \mathbb{R})$, all $g \in H^1(\mathbb{R}, \mathbb{R})$ and all $h \in H^{1+(1-(|j_1|-1))/2}(\mathbb{R}, \mathbb{R})$ and $p \in H^2(\mathbb{R}, \mathbb{R})$ there holds

\[
\|N^i_{j_1}(\varphi, g)\|_{L^2} \leq C \varepsilon^{-1} \|g\|_{H^1},
\]
\[
\|N^i_{j_1-j}(\varphi, f)\|_{L^2} \leq C \varepsilon^{-1} \|f\|_{H^{(1-(|j_1|-1))/2}},
\]
\[
\|\mathcal{P} N^i_{j_1-j}(\varphi, g)\|_{L^2} \leq C \|g\|_{H^1},
\]
\[
\|\mathcal{P} N^i_{j_1-j}(\varphi, f)\|_{L^2} \leq C \|f\|_{H^{(1-(|j_1|-1))/2}},
\]
\[
\|\partial_\alpha N^i_{j_1-j}(\varphi, p)\|_{L^2} \leq C \|p\|_{H^2},
\]
\[
\|\partial_\alpha N^i_{j_1-j}(\varphi, h)\|_{L^2} \leq C \|h\|_{H^{1+(1-(|j_1|-1))/2}}
\]

uniformly on compact subsets of $\mathcal{B}$, where $\mathcal{P} = P_{\delta_0, \infty}$ or $\mathcal{P} = \emptyset$.

b) Let $\varphi$ be as in a). Then for all $f \in L^2(\mathbb{R}, \mathbb{R})$ it holds that

\[
P_{0, \delta_0} N^i_{j_1, j_2} (\varphi, P_{0, \delta_0} f) = 0.
\]

Proof. a) The key step of the proof is to discuss systematically the kernels $\hat{\mathcal{N}}^i_{j_1, j_2 \ell}$ for all $i \in \{1, 2\}$. We start by analyzing the behavior of $\hat{\mathcal{N}}^i_{j_1, j_2 \ell}$ in a neighborhood of the zeros of the factor $\hat{r}_{j_1, j_2}$ in the denominator. As shown above, we have zeros at

- $k = 0$ if $\text{sgn}(j_2) = -1$.
- $k = 0$ if $\text{sgn}(j_2) = -1$.

Because of (155)–(163) we have

\[
|\hat{P}_{0, \delta_0}(k) \hat{\mathcal{N}}^i_{j_1, j_2}(k, \ell k_0, k - \ell k_0)| \lesssim |k|
\]

uniformly with respect to $b \lesssim 1$. Hence, due to (182), the singularity of $\hat{\mathcal{N}}^i_{j_1, j_2 \ell}$ at $k = 0$ can be removed and then $\hat{P}_{0, \delta_0} \hat{\mathcal{N}}^i_{j_1, j_2 \ell}$ is bounded uniformly on compact subsets of $\mathcal{B}$. However, because of

\[
\hat{P}_{0, \delta_0}(k) \hat{\vartheta}^{-1}(k) = O(\varepsilon^{-1})
\]

we have

\[
\hat{P}_{0, \delta_0}(k) \hat{\mathcal{N}}^i_{j_1, j_2 \ell}(k) = O(\varepsilon^{-1})
\]

uniformly on compact subsets of $\mathcal{B}$.

- $k = \ell k_0$ if $\text{sgn}(j_2) = -1$.

By construction of $(\hat{\vartheta} - \varepsilon \hat{\vartheta}_0)(k - \ell k_0)$ we have

\[
|\hat{P}_{0, \delta_0}(k - \ell k_0) (\hat{\vartheta} - \varepsilon \hat{\vartheta}_0)(k - \ell k_0)| \leq (1 + \varepsilon) |k - \ell k_0| \delta_0.
\]
Hence, because of (185) and since \( \sigma \) is differentiable with respect to \( k \), where \( \partial_k \sigma \) depends continuously on \( b \), the singularity of \( \hat{w}_{j_1j_2}^i \) at \( k = \ell k_0 \) can be removed such that we obtain
\[
\hat{P}_{0,0}(k - \ell k_0) \hat{w}_{j_1j_2}^i(k) = \mathcal{O}(1)
\] (264)
uniformly on compact subsets of \( B \).

- \( k = \ell k_1 \) and \( k = -\ell(k_1 - k_0) \) if \( b \in (0, b_2) \) and \( \text{sgn}(j_1) = \text{sgn}(j_2) = -1 \).

The function \( \xi_{j_1j_2} \) is constructed in such a way that the singularities of \( \hat{w}_{j_1j_2}^i \) at \( k = \ell k_1 \) and \( k = -\ell(k_1 - k_0) \) can be removed and then, due to (215)–(216) and the fact that \( k \rightarrow \omega(k, b) \) is odd, we obtain
\[
\hat{P}_{0,0}(k - \ell k_0) \hat{w}_{j_1j_2}^i(k) = \mathcal{O}(\hat{w}_{j_1j_2}^i(k, \ell k_0, k - \ell k_0)),
\] (265)
\[
\hat{P}_{0,0}(k + \ell(k_1 - k_0)) \hat{w}_{j_1j_2}^i(k) = \mathcal{O}(\hat{w}_{j_1j_2}^i(k, \ell k_0, k - \ell k_0))
\] (266)
uniformly on compact subsets of \( B \).

Next, we analyze the asymptotic behavior of the kernels \( \hat{w}_{j_1j_2}^i \) for \( |k| \rightarrow \infty \). Because of (189)–(190) we have
\[
\sigma(k, b) = |k|^{1/2}(1 + bk^2)^{1/2} + \mathcal{O}(e^{-|k|}),
\] (267)
\[
\partial_k \sigma(k, b) = \text{sgn}(k) \frac{1 + 3bk^2}{2(1 + bk^2)} |k|^{-1/2}(1 + bk^2)^{1/2} + \mathcal{O}(e^{-|k|})
\] (268)
for \( |k| \rightarrow \infty \) uniformly with respect to \( b \lesssim 1 \). Inserting (267) in (159)–(160) yields
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2,2} = \mathcal{O}(|k|^{-1/2}(1 + bk^2)^{-1/2}),
\] (269)
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2,3} = \mathcal{O}(b^{1/2}|k|^{-1/2})
\] (270)
for \( |k| \rightarrow \infty \) uniformly with respect to \( b \lesssim 1 \). Furthermore, with the help of (102) and the mean value theorem we derive
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2,4}(k, \ell k_0, k - \ell k_0)
\]
\[
= \text{sgn}(j_1) \sigma^{-1}(\ell k_0)(\ell k_0)^2 \frac{\sigma(k - \theta(k)(k - \ell k_0)) \partial_k \sigma(k - \theta(k)(k - \ell k_0))}{\sigma(k + \sigma(\ell k_0))(k - \ell k_0)}
\] (271)
with \( \theta(k) \in [0, 1] \) for all \( k \in \mathbb{R} \setminus \{\ell k_0\} \). Because of (267)–(268) we obtain
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2,4}(k, \ell k_0, k - \ell k_0) = \mathcal{O}(|k|^{-3/2}(1 + bk^2)^{1/2})
\] (272)
for \( |k| \rightarrow \infty \) uniformly with respect to \( b \lesssim 1 \).

Due to (155)–(158), (162), (269)–(270) and (272), we conclude
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{1}(k, \ell k_0, k - \ell k_0) = -1 + \mathcal{O}(|k|^{-1/2}),
\] (273)
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2}(k, \ell k_0, k - \ell k_0) = \mathcal{O}(|k|^{-(|j|-1)/2}),
\] (274)
\[
(i k)^{-1} \hat{q}_{j_1j_2}^{2}(k, \ell k_0, k - \ell k_0) = \mathcal{O}(|k|^{-1/2}(1 + bk^2)^{1/2})
\] (275)
for $|k| \to \infty$ uniformly with respect to $b \lesssim 1$.

Moreover, by the mean value theorem we have

$$\hat{r}_{jj}(k, \ell k_0, k - \ell k_0) = i(\omega(\ell k_0) + \sgn(j)\omega'(k - \theta(k)\ell k_0)\ell k_0), \quad (276)$$

$$\hat{r}_{j-j}(k, \ell k_0, k - \ell k_0) = i(\omega(\ell k_0) + 2\sgn(j)\omega(k) - \sgn(j)\omega'(k - \theta(k)\ell k_0)\ell k_0) \quad (277)$$

with $\theta(k) \in [0, 1]$ for all $k \in \mathbb{R}$. Due to (191)–(192) and (215)–(216), this implies

$$\hat{\zeta}_{jj}(k) (\hat{r}_{jj}(k, \ell k_0, k - \ell k_0))^{-1} = O((1 + |k|^{-1/2}(1 + bk^2)^{1/2})^{-1}), \quad (278)$$

$$\hat{r}_{j-j}(k, \ell k_0, k - \ell k_0))^{-1} = O(|k|^{-1/2}(1 + bk^2)^{-1/2}) \quad (279)$$

for $|k| \to \infty$ uniformly on compact subsets of $\mathcal{B}$.

Now, using (132), (273)–(274), (275) and (278)–(279), we obtain

$$\hat{\mathcal{H}}_{j1j2}^{-1}(k) = O((1 + |k|^{-1/2}(1 + bk^2)^{1/2})^{-1}), \quad (280)$$

$$\hat{\mathcal{H}}_{j1j2} = O(|k|^{(1-(|j|^{-1})/2}(1 + bk^2)^{-1/2}), \quad (281)$$

$$\hat{\mathcal{H}}_{j1j2}^{-1}(k) = O(1), \quad (282)$$

$$\hat{\mathcal{H}}_{j1j2}(k) = O(1) \quad (283)$$

for $|k| \to \infty$ uniformly on compact subsets of $\mathcal{B}$. Hence, combining (129)–(131), (262), (264), (265)–(266), (280)–(283) and Young’s inequality for convolutions, we arrive at (253)–(258). Since

$$\hat{\mathcal{H}}_{j1j2}^{-1}(-k) = \hat{\mathcal{H}}_{j1j2}(k)$$

and $\varphi$ is real-valued, $f \mapsto N^i_{j}(\varphi, f)$ is a continuous linear map from $H^1(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$ and $f \mapsto N^i_{j-j}(\varphi, f)$ a continuous linear map from $H^{(1-(|j|^{-1})/2}(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$, such that we have proven all assertions of a).

b) is a direct consequence of

$$\hat{\mathcal{P}}_{0, \delta_0}(k) \hat{\mathcal{P}}_{0, \delta_0}(m) \chi_c(k - m) = 0.$$

\end{proof}

**Lemma 4.5.** Let $R_{-1}, R_1 \in L^2(\mathbb{R}, \mathbb{R}), R_{-2}, R_2 \in H^s(\mathbb{R}, \mathbb{R})$ and $1 \leq l \leq s - 1$. Then there holds

$$\int_{\mathbb{R}} \partial_\alpha^l R_{j1} \rho_{j1}^l \partial_\alpha^l (\sgn(j_1) i \omega_{j1j2}(\psi_c, R_{j2}) + N_{j1j2}(i \omega \psi_c, R_{j2})) d\alpha$$

$$- \sgn(j_2) N_{j1j2}(\psi_c, i \omega R_{j2})) d\alpha$$

$$= \int_{\mathbb{R}} \partial_\alpha^l R_{j1} \rho_{j1}^l \partial_\alpha^l Q_{j1j2}(\psi_c) R_{j2} d\alpha + \epsilon \int_{\mathbb{R}} \partial_\alpha^l R_{j1} \rho_{j1}^l \partial_\alpha^l Y_{j1j2}(\psi_c, R_{j2}) d\alpha \quad (284)$$
with
\[
\left| \sum_{j_1 \in \{\pm 1\}, j_2 \in \{\pm j_1\}} \int_\mathbb{R} \alpha(a) R_{j_1} \rho_{j_1} \partial_a Y_{j_1 j_2}(\psi_c, R_{j_2}) \, da \right| \lesssim \|R_1\|_{(L^2)^2}^2 + \|R_2\|_{(H^{\max(2,j)}_1)^2}^2 + \varepsilon^4,
\]
(285)
as long as \(\varepsilon^{5/2} \|R_1\|_{(L^2)^2}, \varepsilon^{5/2} \|R_2\|_{(H^{\max(2,j)}_1)^2} \ll 1\), and
\[
\left| \sum_{j_1 \in \{\pm 1\}, j_2 \in \{\pm j_1\}} \int_\mathbb{R} \alpha(a) R_{j_1} \rho_{j_1} \partial_a Y_{j_1 j_2}(\psi_c, R_{j_2}) \, da \right| \lesssim \|R_2\|_{(H^1)^2} \left( \|R_1\|_{(L^2)^2} + \|R_2\|_{(H^{1+2})^2} + \varepsilon^2 \right),
\]
(286)
as long as \(\varepsilon^{5/2} \|R_1\|_{(L^2)^2}, \varepsilon^{5/2} \|R_2\|_{(H^{1+2})^2} \ll 1\), uniformly on compact subsets of \(\mathcal{B}\).

**Proof.** Let
\[
\varepsilon Y_{j_1 j_2}(\psi_c, R_{j_2}) := \text{sgn}(j_1)i \omega N_{j_1 j_2}(\psi_c, R_{j_2}) + N_{j_1 j_2}(i \omega \psi_c, R_{j_2}) - \text{sgn}(j_2) N_{j_1 j_2}(\psi_c, i \omega R_{j_2}) - Q_{j_1 j_2}(\psi_c) R_{j_2}.
\]
Then we have
\[
\hat{Y}_{j_1 j_2}(\psi_c, R_{j_2})(k) = \varepsilon^{-1} \sum_{\mu=1}^{|j_1|} \sum_{\ell \in \{\pm 1\}} \int_\mathbb{R} \hat{\gamma}^{-1}(k) K_{j_1 j_2 \ell}(k, k - m, m) \hat{\psi}(k - m)
\]
\[
\times (im)^{-(\mu-1)} \hat{R}_{j_2}(m) \, dm \tag{287}
\]
with
\[
K_{j_1 j_2 \ell}(k, k - m, m)
\]
\[
= \frac{\hat{r}_{j_1 j_2}(k, k - m, m) \hat{\gamma}_{j_1 j_2 \ell}(k) \hat{\alpha}_{j_1 j_2 \ell}(k, \ell k_0, k - \ell k_0)}{\hat{r}_{j_1 j_2}(k, \ell k_0, k - \ell k_0) (\hat{\theta} - \hat{\xi}_0)(k - \ell k_0)} - \hat{q}_{j_1 j_2}(k, k - m, m) \hat{\gamma}(m). \tag{288}
\]
We split \(K_{j_1 j_2 \ell}^\mu\) into
\[
K_{j_1 j_2 \ell}^\mu(k, k - m, m) = K_{j_1 j_2 \ell}^\mu(k, k - m, m) - K_{j_1 j_2 \ell}^\mu(k, \ell k_0, m)
\]
\[
+ K_{j_1 j_2 \ell}^\mu(k, \ell k_0, m) - K_{j_1 j_2 \ell}^\mu(k, \ell k_0, k - \ell k_0)
\]
\[
+ K_{j_1 j_2 \ell}^\mu(k, \ell k_0, k - \ell k_0). \tag{289}
\]
We have
\[
K_{j_1 j_2 \ell}^\mu(k, \ell k_0, k - \ell k_0) = -(1 - \hat{\gamma}_{j_1 j_2 \ell}(k) + \varepsilon \hat{\xi}_0(k - \ell k_0)) \hat{q}_{j_1 j_2 \ell}(k, \ell k_0, k - \ell k_0), \tag{290}
\]
where \(\hat{\gamma}_{j_1 j_2 \ell}(k) = 1\) if \(b \in \mathcal{B} \setminus (0, b_0)\) or \(\text{sgn}(j_1) = 1\) or \(\text{sgn}(j_2) = 1\).
Let \( b \in (0, b_0) \), \( j \in \{-1, -2\} \),
\[
\widehat{S}_{j \ell}^{l \mu}(k, m) = -\hat{\varrho}^{-1}(k) \hat{\rho}_j^l(k)^{2l} (1 - \hat{\zeta}_{j \ell}(k)) \widehat{S}_{j j}^{l \mu}(k, \ell k_0, k - \ell k_0)(im)^{-\mu-1}
\]
for all \( k, m \in \mathbb{R} \) and the function \( g_j^{l \mu} \) be defined by its Fourier transform
\[
\widehat{g}_j^{l \mu}(k) = \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} (-ik)^{-l} \widehat{S}_{j \ell}^{l \mu}(k, m) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dm
\]
for all \( k \in \mathbb{R} \). Then, due to (148)–(147) and Young’s inequality for convolutions, we have \( g_j^{l \mu} \in L^2(\mathbb{R}, \mathbb{R}) \) and with the help of Fubini’s theorem we deduce
\[
\sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \widehat{S}_{j \ell}^{l \mu}(k, m) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk
\]
\[
= \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \widehat{S}_{j \ell}^{l \mu}(k, m) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk
\]
\[
+ \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \widehat{S}_{j \ell}^{l \mu}(k, m) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk
\]
\[
= \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \widehat{S}_{j \ell}^{l \mu}(m, k) \hat{\psi}_\ell(m - k) \hat{R}_j(m) \, dmdk
\]
\[
+ \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \widehat{S}_{j \ell}^{l \mu}(-m, -k) \hat{\psi}_{-\ell}(k - m) \hat{R}_j(m) \, dmdk
\]
\[
= \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \left( \widehat{S}_{j \ell}^{l \mu}(k, k - \ell k_0) + \widehat{S}_{j \ell}^{l \mu}(-k + \ell k_0, -k) \right) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk
\]
\[
+ \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \left( \widehat{S}_{j \ell}^{l \mu}(k, m) - \widehat{S}_{j \ell}^{l \mu}(k, -\ell k_0) \right) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk
\]
\[
+ \frac{1}{2} \sum_{\ell \in \{\pm \}} \int_{\mathbb{R}} \hat{R}_j(k) \left( \widehat{S}_{j \ell}^{l \mu}(-m, -k) - \widehat{S}_{j \ell}^{l \mu}(-k + \ell k_0, -k) \right) \hat{\psi}_\ell(k - m) \hat{R}_j(m) \, dmdk.
\]
Since \( \hat{\xi}_1(-k) = \hat{\xi}_1(k) \) for all \( k \in \mathbb{R} \), we have by construction of \( \hat{\rho}_j^l \) and \( \widehat{S}_{j \ell}^{l \mu} \):
\[
- \sum_{\mu=1}^{\lfloor j \rfloor} \widehat{S}_{j \ell}^{l \mu}(k, k - \ell k_0) + \widehat{S}_{j \ell}^{l \mu}(-k + \ell k_0, -k)
\]
Because of (112), (118), (192), (273)–(275) and (280)–(283) the functions 

\[ \text{Fix} \ a) \quad (\text{continuous linear map from } H) \]

and all \( g \in H \) exist a constant \( C > 0 \) which yields

\[ K \text{ have the properties of (123), (146)–(148), (248)–(249), (289)–(291) and Young’s inequality for convolution.} \]

\( \hat{\varphi} \mid \mu \) \( \hat{\mu} \)

\( \sum_{|j|} \sum_{\mu=1}^{|j|} \int_{\mathbb{R}} \int_{\mathbb{R}} R_j(k) \frac{\xi_j^{|l|/\ell}}{\xi_1^{|\ell|}} (k, m) \tilde{\psi}_\ell(k-m) R_j(m) \, dmdk \)

\[ = \frac{1}{2} \sum_{\mu=1}^{(||j||)} \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} R_j(k) \left( \xi_j^{|l|/\ell}(k, m) - \xi_j^{|l|/\ell}(k, k - \ell k_0) \right) \tilde{\psi}_\ell(k-m) R_j(m) \, dmdk \]

\[ + \frac{1}{2} \sum_{\mu=1}^{(||j||)} \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} R_j(k) \left( \xi_j^{|l|/\ell}(-m, -k) - \xi_j^{|l|/\ell}(-k + \ell k_0, -k) \right) \tilde{\psi}_\ell(k-m) R_j(m) \, dmdk. \]

Because of (112), (118), (192), (273)–(275) and (280)–(283) the functions

\[ (k, k - m, m) \mapsto (1 + k^{2j})^{-1} \xi_j^{|l|/\ell}(k, m), \]

\[ (k, k - m, m) \mapsto (1 + k^{2j})^{-1} \xi_j^{|l|/\ell}(-m, -k), \]

\[ (k, k - m, m) \mapsto (1 + |k|)^{-(\mu+1)/2} \tilde{\varphi}^{-1}(k) K_{jj\ell}^{|l|/\ell}(k, k - m, m) \]

have the properties of \( \mathcal{A} \) in Lemma 4.2 a) and b), respectively, such that we can use (123), (146)–(148), (248)–(249), (289)–(291) and Young’s inequality for convolutions to obtain (284)–(285) uniformly on compact subsets of \( \mathcal{B} \).

**Lemma 4.6.** The operators \( \mathcal{N}_{jj\ell}^{|l|} \) have the following properties:

**a)** Fix \( \varphi \in L^2(\mathbb{R}, \mathbb{R}) \) with \( \text{supp } \varphi = \text{supp } \tilde{\psi}_c \). Then \( f \mapsto \mathcal{N}_{jj}^{|l|}(\varphi, f) \) defines a continuous linear map from \( H^1(\mathbb{R}, \mathbb{R}) \) into \( L^2(\mathbb{R}, \mathbb{R}) \) and \( f \mapsto \mathcal{N}_{jj}^{|l|}(\varphi, f) \) a continuous linear map from \( H^{1-(|j|+1)/2}(\mathbb{R}, \mathbb{R}) \) into \( L^2(\mathbb{R}, \mathbb{R}) \). Furthermore, there exists a constant \( C > 0 \) with \( C \lesssim \| \varphi \|_{L^1} \) such that for all \( f \in H^{1-(|j|+1)/2}(\mathbb{R}, \mathbb{R}) \) and all \( g \in H^1(\mathbb{R}, \mathbb{R}) \) it holds that

\[ \| \mathcal{N}_{jj}^{|l|}(\varphi, g) \|_{L^2} \leq C \varepsilon^{-1} \| g \|_{H^1}, \quad (292) \]

\[ \| \mathcal{N}_{j-j}^{|l|}(\varphi, f) \|_{L^2} \leq C \varepsilon^{-1} \| f \|_{H^{1-(|j|+1)/2}}, \quad (293) \]

\[ \| P_{\delta_0, \infty} \mathcal{N}_{jj}^{|l|}(\varphi, g) \|_{L^2} \leq C \| g \|_{H^1}, \quad (294) \]
\[ \| P_{\delta_0, \infty} \mathcal{N}_j^i (\varphi, f) \|_{L^2} \leq C \| f \|_{H^{1-(|j|-1)/2}}, \]  

(295)

uniformly on compact subsets of \( \mathcal{B} \).

b) Let \( \varphi \) be as in a). Then for all \( f \in L^2(\mathbb{R}, \mathbb{R}) \) it holds that

\[ P_{0, \delta_0} \mathcal{N}_j^i (\varphi, P_{0, \delta_0} f) = 0. \]  

(296)

Proof. a) The first step of the proof is to analyze the behavior of \( \hat{\pi}_{j_1, j_2}^i \) for all \( i \in \{1, 2\} \) in a neighborhood of the zeros of the factor \( \hat{\pi}_{j_1, j_2}^i \) in the denominator. Due to the localization of the supports of \( \hat{P}_{0, \delta_0} \) and \( \hat{\varphi} \), it is sufficient to consider only the zeros satisfying \( |k| \leq \delta_0 \) and \( |k - m \mp k_0| \leq \delta_0 \). As shown above, the only zeros satisfying \( |k| \leq \delta_0 \) and \( |k - m \mp k_0| \leq \delta_0 \) are \((0, \pm k_0, \mp k_0)\), which appear if \( \text{sgn}(j_2) = -1 \). By the same arguments as those in the proof of Lemma 4.4 it follows that the singularities of \( \hat{\pi}_{j_1, j_2}^i \) at \((0, \pm k_0, \mp k_0)\) can be removed and then it holds that

\[ \hat{P}_{0, \delta_0}(k) \hat{P}_{0, \delta_0}(k - m \mp k_0) \hat{\pi}_{j_1, j_2}^{i, 0}(k, k - m, m) = O(e^{-1}) \]  

(297)

uniformly on compact subsets of \( \mathcal{B} \). Hence, because of (255)–(256) all assertions of a) are valid.

b) is again a direct consequence of

\[ \hat{P}_{0, \delta_0}(k) \hat{P}_{0, \delta_0}(m) \chi_{\varepsilon}(k - m) = 0. \]

□

Lemma 4.7. Let \( R_- \), \( R_1 \in L^2(\mathbb{R}, \mathbb{R}) \) and \( R_- \), \( R_2 \in H^2(\mathbb{R}, \mathbb{R}) \). Then it holds that

\[
\int_{\mathbb{R}} R_j \rho_{j_1}^0 (\text{sgn}(j_1) i \omega \mathcal{N}_{j_1, j_2}(\psi_c, R_{j_2}) + \mathcal{N}_{j_1, j_2}(i \omega \psi_c, R_{j_2})) \\
- \text{sgn}(j_2) \mathcal{N}_{j_1, j_2}(\psi_c, i \omega R_{j_2})) \, d\alpha \\
= \int_{\mathbb{R}} R_j \rho_{j_1}^0 Q_{j_1, j_2}(\psi_c) R_{j_2} \, d\alpha + \varepsilon \int_{\mathbb{R}} R_j \rho_{j_1}^0 \mathcal{Y}_{j_1, j_2}(\psi_c, R_{j_2}) \, d\alpha 
\]  

(298)

with

\[
\left| \sum_{j_1 \in \{\pm 1, \pm 2\}, j_2 \in \{\pm 1\}} \int_{\mathbb{R}} R_j \rho_{j_1}^0 \mathcal{Y}_{j_1, j_2}(\psi_c, R_{j_2}) \, d\alpha \right| \lesssim \| R_1 \|_{L^2}^2 + \| R_2 \|_{H^2}^2 + \varepsilon^4,
\]  

(299)

as long as \( \varepsilon^{5/2} \| R_1 \|_{L^2}, \varepsilon^{5/2} \| R_2 \|_{H^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Proof. Let

\[ \varepsilon \mathcal{Y}_{j_1, j_2}(\psi_c, R_{j_2}) := \text{sgn}(j_1) i \omega \mathcal{N}_{j_1, j_2}(\psi_c, R_{j_2}) + \mathcal{N}_{j_1, j_2}(i \omega \psi_c, R_{j_2}) \\
- \text{sgn}(j_2) \mathcal{N}_{j_1, j_2}(\psi_c, i \omega R_{j_2}) \\
- Q_{j_1, j_2}(\psi_c) R_{j_2}. \]
Then we have
\[
\hat{\gamma}_{j_1 j_2}(\psi_e, R_{j_2})(k) = \varepsilon^{-1} \sum_{j_1} \sum_{\mu=1}^{j_1} \int_{\mathbb{R}} \hat{P}_{0, \infty}(k) K_{j_1 j_2}^\mu(k, k - m, m) \hat{\psi}_\ell(k - m) \times (im)^{-\mu-1} \hat{R}_{j_2}(m) \, dm,
\]
where \( K_{j_1 j_2}^\mu \) is defined by (288). Hence, the assertion of the lemma follows by the same arguments as those from the proof of Lemma 4.5.

Let \( i \in \{1, 2\}, l \geq 1, \varphi \in L^2(\mathbb{R}, \mathbb{R}) \) with \( \text{supp} \hat{\varphi} = \text{supp} \hat{\psi}_e, f := (f_{-2}, f_2)^T \in (H^1(\mathbb{R}, \mathbb{R}))^2 \),
\[
(\rho^i N^i)(\varphi) f = \left( \begin{array}{c}
(\rho^i N^i)_{-2}(-2)(\varphi) (\rho^i N^i)_{-22}(\varphi) \\
(\rho^i N^i)_{2}(-2)(\varphi) (\rho^i N^i)_{22}(\varphi)
\end{array} \right) \left( \begin{array}{c}
 f_{-2} \\
 f_2
\end{array} \right)
\]
with
\[
(\rho^i N^i)_{j_1 j_2}(\varphi) f_{j_2} = \rho^l_{j_1} \vartheta N^i_{j_1 j_2}(\varphi) f_{j_2} = \rho^l_{j_1} \vartheta N^i_{j_1 j_2}(\varphi, f_{j_2})
\]
for \( j_1, j_2 \in \{\pm 2\} \). Then it holds that
\[
((\hat{\rho^i N^i})_{j_1 j_2}(\varphi) f_{j_2})(k) = \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} (\hat{\rho^i n^i})_{j_1 j_2 \ell}(k) \hat{\varphi}_\ell(k - m) \hat{f}_{j_2}(m) \, dm
\]
with
\[
(\hat{\rho^i n^i})_{j_1 j_2 \ell}(k) = \hat{\rho}^l_{j_1}(k) \vartheta(k) \hat{N}^i_{j_1 j_2 \ell}(k).
\]
Moreover, let
\[
(\rho^i N^i)^*(\varphi) = \left( \begin{array}{c}
(\rho^i N^i)^*_{-2}(-2)(\varphi) (\rho^i N^i)^*_{-22}(\varphi) \\
(\rho^i N^i)^*_{2}(-2)(\varphi) (\rho^i N^i)^*_{22}(\varphi)
\end{array} \right)
\]
be the adjoint operator of \((\rho^i N^i)(\varphi)\). That means, we have
\[
\langle (\rho^i N^i)(\varphi) f, g \rangle_{(L^2)^2} = \langle f, (\rho^i N^i)^*(\varphi) g \rangle_{(L^2)^2}
\]
for all \( f, g \in (H^1(\mathbb{R}, \mathbb{R}))^2 \), as well as
\[
(\rho^i N^i)^*(\varphi) = \left( \begin{array}{c}
(\rho^l_{-2} \vartheta N^i_{-2})(\varphi) (\rho^l_{-2} \vartheta N^i_{-2})(\varphi) \\
(\rho^l_{2} \vartheta N^i_{2})(\varphi) (\rho^l_{2} \vartheta N^i_{2})(\varphi)
\end{array} \right)^T = \left( \begin{array}{c}
(\rho^l_{-2} \vartheta N^i_{-2})(\varphi) (\rho^l_{-2} \vartheta N^i_{-2})(\varphi) \\
(\rho^l_{2} \vartheta N^i_{2})(\varphi) (\rho^l_{2} \vartheta N^i_{2})(\varphi)
\end{array} \right),
\]
where
\[
\langle \rho^l_{j_1} \vartheta N^i_{j_1 j_2}(\varphi) h, p \rangle_{L^2} = \langle h, (\rho^l_{j_1} \vartheta N^i_{j_1 j_2})^*(\varphi) p \rangle_{L^2}
\]
for all \( h, p \in H^1(\mathbb{R}, \mathbb{R}) \). Then
\[
(\rho^i N^i)^*(\varphi) = \left( \begin{array}{c}
(\rho^i N^i)^*_{-2}(-2)(\varphi) (\rho^i N^i)^*_{-22}(\varphi) \\
(\rho^i N^i)^*_{2}(-2)(\varphi) (\rho^i N^i)^*_{22}(\varphi)
\end{array} \right) := \frac{1}{2}((\rho^i N^i)(\varphi) + (\rho^i N^i)^*(\varphi))
\]
denotes the symmetric part and

\[(\rho^l \mathcal{M}^a)(\varphi) = \left(\begin{array}{cc}
(\rho^l \mathcal{M}^a)_{2-2}(\varphi) & (\rho^l \mathcal{M}^a)_{2-2}(\varphi) \\
(\rho^l \mathcal{M}^a)_{2-2}(\varphi) & (\rho^l \mathcal{M}^a)_{2-2}(\varphi)
\end{array}\right)
= \frac{1}{2}((\rho^l \mathcal{M}^a)(\varphi) - (\rho^l \mathcal{M}^a)^*(\varphi))
\]

the antisymmetric part of \((\rho^l \mathcal{M}^a)(\varphi)\). It holds that

\[((\rho^l \mathcal{M}^a)^*_{j_1 j_2}(\varphi) f_{j_2})(k) = \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} (\rho^l \mathcal{n}^s_{j_1 j_2 \ell})(m) \hat{\varphi}_\ell(k-m) \hat{f}_{j_2}(m) \, dm\]

with

\[(\rho^l \mathcal{n}^s_{j_1 j_2 \ell}(m) = (\rho^l \mathcal{n}^s)_{j_2 j_1 \ell}(-m) = \hat{\rho}_{j_2}(-m) \hat{\varphi}_{j_2 j_1 \ell}(-m)\]

and therefore

\[((\rho^l \mathcal{M}^a)^*_{j_1 j_2}(\varphi) f_{j_2})(k) = \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} (\rho^l \mathcal{n}^s_{j_1 j_2 \ell})(k, m) \hat{\varphi}_\ell(k-m) \hat{f}_{j_2}(m) \, dm\]

with

\[(\rho^l \mathcal{n}^s_{j_1 j_2 \ell}(k, m) = \frac{1}{2}((\rho^l \mathcal{n}^s_{j_1 j_2 \ell}(k) + (\rho^l \mathcal{n}^s)_{j_2 j_1 \ell}(-m))\]

as well as

\[((\rho^l \mathcal{M}^a)^{1a}_{j_1 j_2}(\varphi) f_{j_2})(k) = \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} (\rho^l \mathcal{n}^{1a}_{j_1 j_2 \ell})(k, m) \hat{\varphi}_\ell(k-m) \hat{f}_{j_2}(m) \, dm\]

with

\[(\rho^l \mathcal{n}^{1a}_{j_1 j_2 \ell}(k, m) = \frac{1}{2}((\rho^l \mathcal{n}^{1a}_{j_1 j_2 \ell}(k) - (\rho^l \mathcal{n}^s)_{j_2 j_1 \ell}(-m))\]

More generally, for any densely defined linear operator \(L\) on a Hilbert space we denote its adjoint operator by \(L^*\), its symmetric part by \(L^s\) and its antisymmetric part by \(L^a\).

**Lemma 4.8.** The operators \((\rho^l \mathcal{M}^1)^{s}_{j_1 j_2}\) with \(j_1, j_2 \in \{\pm 2\}\) and \(l \geq 1\) have the following properties:

\(a)\) Fix \(\varphi \in L^2(\mathbb{R}, \mathbb{R})\) with \(\text{supp} \hat{\varphi} = \text{supp} \hat{\psi}_c\). Then \(f \mapsto (\rho^l \mathcal{M}^1)^{s}_{j_1 j_2}(\varphi) f\) can be extended to a continuous linear map from \(L^2(\mathbb{R}, \mathbb{R})\) into \(L^2(\mathbb{R}, \mathbb{R})\) and there exists a constant \(C_1 > 0\) with \(C_1 \lesssim \|\hat{\varphi}\|_{L^1}\) such that for all \(f \in L^2(\mathbb{R}, \mathbb{R})\) it holds that

\[\|(\rho^l \mathcal{M}^1)^{s}_{j_1 j_2}(\varphi) f\|_{L^2} \leq C_1 \|f\|_{L^2}\]  

(301)

uniformly on compact subsets of \(\mathcal{B}\). Moreover, for all \(\psi \in H^{2+p}(\mathbb{R}, \mathbb{R})\) with \(p > 1/2\) there exists a constant \(C_2 > 0\) with \(C_2 \lesssim \|\hat{\varphi}\|_{L^1} \|\partial_\alpha \psi\|_{H^{1+p}}\) such that for all \(f \in L^2(\mathbb{R}, \mathbb{R})\) it holds that

\[\|\psi, (\rho^l \mathcal{M}^1)^{s}_{j_1 j_2}(\varphi) f\|_{H^1} \leq C_2 \|f\|_{L^2}\]  

(302)
uniformly on compact subsets of $B$.

b) For all $\varphi \in L^2(\mathbb{R}, \mathbb{R})$ with $\text{supp} \tilde{\varphi} = \text{supp} \hat{\psi}_c$ there exists a symmetric linear operator $G_j^1(\varphi) : L^2(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R})$ such that

$$ (\rho^j \Upsilon^1)_{j-j}(\varphi) f = G_j^1(\varphi) f + M_j^1(\varphi) f \quad (303) $$

and there exists a constant $C_3 > 0$ with $C_3 \lesssim \|\varphi\|_{L^1}$ such that for all $f \in L^2(\mathbb{R}, \mathbb{R})$ it holds that

$$ \|M_j^1(\varphi) f\|_{H^{1/2}} \leq C_3 \|f\|_{L^2} \quad (304) $$

uniformly on compact subsets of $B$. Moreover, for all $\psi \in H^{2+p}(\mathbb{R}, \mathbb{R})$ with $p > 1/2$ there exists a constant $C_4 > 0$ with $C_4 \lesssim \|\varphi\|_{L^1} \|\partial_\alpha \psi\|_{H^{1+p}}$ such that for all $f \in L^2(\mathbb{R}, \mathbb{R})$ it holds that

$$ \|[\psi, G_j^1(\varphi)] f\|_{H^1} \leq C_4 \|f\|_{L^2} \quad (305) $$

uniformly on compact subsets of $B$.

Proof. a) Let $j \in \{\pm 2\}$. We have

$$ \partial_1 \sum_{\mu=1}^3 \hat{q}_{j \mu}^2(k, \ell k_0, k - \ell k_0) = -i + O(|k|^{-1/2}), \quad (306) $$

$$ \partial_1 \sum_{\mu=1}^3 \hat{q}_{j-j \mu}^2(k, \ell k_0, k - \ell k_0) = O(|k|^{-1/2}), \quad (307) $$

$$ \partial_1^2 \sum_{\mu=1}^3 \hat{q}_{j1j2}^2(k, \ell k_0, k - \ell k_0) = 0, \quad (308) $$

$$ \partial_3 \sum_{\mu=1}^3 \hat{q}_{j1j2}^2(k, \ell k_0, k - \ell k_0) = O(|k|^{-1/2}), \quad (309) $$

$$ \partial_3^2 \sum_{\mu=1}^3 \hat{q}_{j1j2}^2(k, \ell k_0, k - \ell k_0) = O(|k|^{-3/2}), \quad (310) $$

$$ \partial_1 \partial_3 \sum_{\mu=1}^3 \hat{q}_{j1j2}^2(k, \ell k_0, k - \ell k_0) = \partial_3 \partial_1 \sum_{\mu=1}^3 \hat{q}_{j1j2}^2(k, \ell k_0, k - \ell k_0) = O(|k|^{-3/2}) \quad (311) $$

for $|k| \to \infty$ uniformly with respect to $b \ll 1$.

Because of (162)–(163), (166)–(167), (211), (273)–(274) and (306)–(311), we obtain by construction of $\hat{\gamma}_j^1$ and $\hat{\xi}_{j1j2}^\ell$ for all $j \in \{\pm 2\}$, $\beta \in \{1, 2\}$, $\gamma \in \{0, 1\}$ and $n \in \{1, 3\}$ that

$$ \left( \frac{d^\beta}{dk^\beta} (\rho_j^\beta \hat{\xi}_{j1j2}^\ell) \right)(-k) \partial_\gamma^\beta \hat{\gamma}_j^1(-k, \ell k_0, -k + \ell k_0) = O(|k|^{-(\beta + \gamma - 1)}), \quad (312) $$
\[
\left( \frac{d^\beta}{dk^\beta} (\hat{\rho}_{j-i} \hat{\zeta}_{j-i}) \right) (-k) \hat{a}_n^i \hat{q}^{-1}_{j-i} (-k, \ell k_0, -k + \ell k_0) = \mathcal{O}(|k|^{-(\beta+\gamma-1/2)})
\]  
(313) 

for \(|k| \to \infty\) uniformly with respect to \(b \lesssim 1\).

Moreover, since \(\omega\) is odd, we have
\[
\hat{r}_{j_2j_1} (-k, \ell k_0, -k - \ell k_0) = \hat{r}_{j_1j_2} (k + \ell k_0, \ell k_0, k).
\]  
(314) 

Let \(\hat{r}_{j_1j_2\ell} : \mathbb{R} \to \mathbb{C}\) defined by \(\hat{r}_{j_1j_2\ell}(k) = \hat{r}_{j_1j_2}(k, \ell k_0, k - \ell k_0)\). Then it holds for all \(\beta \in \{1, 2\}\) that
\[
\frac{d^\beta}{dk^\beta} \hat{r}_{j_1j_2\ell}(k) = \mathcal{O}(|k|^{-(\beta+1/2)}(1 + bk^2)^{1/2}),
\]  
(315) 
\[
\frac{d^\beta}{dk^\beta} \hat{r}_{j-j\ell}(k) = \mathcal{O}(|k|^{-(\beta-1/2)}(1 + bk^2)^{1/2})
\]  
(316) 

for \(|k| \to \infty\) uniformly with respect to \(b\) in compact subsets of \(\mathcal{B}\).

With the help of Taylor’s theorem as well as (162)–(163), (166)–(167), (273)–(274), (278)–(279) and (306)–(316) we derive
\[
(\rho^{n1})^{s}_{j_1j_2\ell}(k, m) \chi_{\ell}(k - m)
\]
\[
= (\rho^{n1})^{s,1}_{j_1j_2\ell}(k) + (\rho^{n1})^{s,2}_{j_1j_2\ell}(k - m) + (\rho^{n1})^{s,3}_{j_1j_2\ell}(k - m)
\]
\[
+ (\rho^{n1})^{s,4}_{j_1j_2\ell}(k) (k - m - 2\ell k_0) + (\rho^{n1})^{s,5}_{j_1j_2\ell}(k) (k - m - \ell k_0)
\]
\[
+ \delta_{j_1j_2} \mathcal{O}(|k|^{-1}(1 + |k|^{-1/2})(1 + bk^2)^{1/2})^{-1})
\]
\[
+ \delta_{j_1j_2} \mathcal{O}(|k|^{-3/2}(1 + bk^2)^{-1/2})
\]  
(317) 

for \(|k| \to \infty\) uniformly with respect to \(m \in \mathbb{R}\) and \(b\) in compact subsets of \(\mathcal{B}\), where
\[
(\rho^{n1})^{s,1}_{j_1j_2\ell}(k) = \frac{\rho^{l}_{j_1}(k) \hat{\zeta}_{j_1j_2\ell}(k) \hat{q}^{-1}_{j_1j_2\ell}(k, \ell k_0, k - \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)}
\]
\[
+ \frac{\rho^{l}_{j_2}(-k) \hat{\zeta}_{j_2j_1\ell}(-k) \hat{q}^{-1}_{j_2j_1\ell}(-k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)},
\]  
(318) 
\[
(\rho^{n1})^{s,2}_{j_1j_2\ell}(k) = -\frac{\rho^{l}_{j_2}(-k) \hat{\zeta}_{j_2j_1\ell}(-k) \hat{q}^{-1}_{j_2j_1\ell}(-k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)},
\]  
(319) 
\[
(\rho^{n1})^{s,3}_{j_1j_2\ell}(k) = -\frac{\rho^{l}_{j_2}(-k) \hat{\zeta}_{j_2j_1\ell}(-k) \hat{q}^{-1}_{j_2j_1\ell}(-k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)},
\]  
(320) 
\[
(\rho^{n1})^{s,4}_{j_1j_2\ell}(k) = -\frac{\rho^{l}_{j_2}(-k) \hat{\zeta}_{j_2j_1\ell}(-k) \hat{q}^{-1}_{j_2j_1\ell}(-k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)} \frac{d}{dk} \hat{r}_{j_1j_2\ell}(k),
\]  
(321) 
\[
(\rho^{n1})^{s,5}_{j_1j_2\ell}(k) = \frac{\rho^{l}_{j_2}(-k) \hat{\zeta}_{j_2j_1\ell}(-k) \hat{q}^{-1}_{j_2j_1\ell}(-k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{j_1j_2\ell}(k)} \frac{d}{dk} \hat{r}_{j_1j_2\ell}(k),
\]  
(322)
and consequently
\[ (\hat{\rho}^1\mathcal{N})^{s}_{jj\ell}(k, m) \chi_\ell(k - m) = \mathcal{O}((1 + |k|^{-1/2})^{-1}), \quad (323) \]
\[ (\hat{\rho}^1\mathcal{N})^{s}_{j-j\ell}(k, m) \chi_\ell(k - m) = \mathcal{O}((1 + bk^2)^{-1/2}) \]
for \(|k| \to \infty\) uniformly with respect to \(m \in \mathbb{R}\) and \(b\) in compact subsets of \(\mathcal{B}\), which implies (301).

To prove the second assertion of a) we consider
\[ \mathcal{F}(\{\psi, (\rho^1\mathcal{N})^{s}_{j1, j2}(\varphi)\}f)(k) \]
\[ \quad = \sum_{\ell \in \{\pm 1\}} \left( \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{\psi}(k - m) (\hat{\rho}^1\mathcal{N})^{s}_{jj\ell}(m, n) \hat{\varphi}(m - n) \hat{f}(n) \, dn \, dm \right. \]
\[ \quad \quad \quad - \int_{\mathbb{R}} \int_{\mathbb{R}} (\hat{\rho}^1\mathcal{N})^{s}_{j, j\ell}(k, m) \hat{\varphi}(k - m) \hat{\psi}(m - n) \hat{f}(n) \, dn \, dm \right). \quad (325) \]

Expanding the kernels \((\hat{\rho}^1\mathcal{N})^{s}_{j1, j2}(m, n)\) and \((\hat{\rho}^1\mathcal{N})^{s}_{j, j\ell}(k, m)\) by (317)–(322), rewriting the factors \((\hat{\rho}^1\mathcal{N})^{s, i}_{j1, j2}(m)\) for \(i = 1, \ldots, 5\) as
\[ (\hat{\rho}^1\mathcal{N})^{s, i}_{j1, j2}(m) = (\hat{\rho}^1\mathcal{N})^{s, i}_{j1, j2}(k) + ((\hat{\rho}^1\mathcal{N})^{s, i}_{j1, j2}(m) - (\hat{\rho}^1\mathcal{N})^{s, i}_{j1, j2}(k)) \]
and using (162)–(163), (166)–(167), (273)–(274), (278)–(279), (306)–(316), the mean value theorem and the fact that the convolution is a commutative operation yields (302).

b) Let \(G^l_j(\varphi)\) defined by
\[ G^l_j(\varphi) f = \frac{1}{2} ((\rho^1\mathcal{N})^{s}_{j-j}(\varphi) f + (\rho^1\mathcal{N})^{s}_{-jj}(\varphi) f) \]
for all \(f \in L^2(\mathbb{R}, \mathbb{R})\) and \(M^l_j(\varphi) := (\rho^1\mathcal{N})^{s}_{j-j}(\varphi) - G^l_j(\varphi)\). Then, \(G^l_j(\varphi)\) is symmetric. Moreover, we have
\[ M^l_j(\varphi) f = \frac{1}{2} ((\rho^1\mathcal{N})^{s}_{j-j}(\varphi) f - (\rho^1\mathcal{N})^{s}_{-jj}(\varphi) f) \]
for all \(f \in L^2(\mathbb{R}, \mathbb{R})\). There holds
\[ (\hat{\rho}^1\mathcal{N})^{s}_{j-j\ell}(k, m) - (\hat{\rho}^1\mathcal{N})^{s}_{-jj\ell}(k, m) \]
\[ = \frac{1}{2} (\hat{\rho}^l_j(k) \hat{\vartheta}(k) \hat{n}^1_{j-j\ell}(k) - \hat{\rho}^l_{-j}(k) \hat{\vartheta}(k) \hat{n}^1_{-jj\ell}(k)) \]
\[ + \frac{1}{2} (\hat{\rho}^l_j(-m) \hat{\vartheta}(-m) \hat{n}^1_{j-j\ell}(-m) - \hat{\rho}^l_{-j}(-m) \hat{\vartheta}(-m) \hat{n}^1_{-jj\ell}(-m)) \]
\[ = \left( \hat{\rho}^l_j(k) \tilde{q}^1_{j-j\ell}(k, \ell k_0, k - \ell k_0) - \hat{\rho}^l_{-j}(k) \tilde{q}^1_{-jj\ell}(k, \ell k_0, k - \ell k_0) \right) \]
\[ = \left( \hat{\rho}^l_{-j}(k) \tilde{q}^1_{-jj\ell}(k, \ell k_0, k - \ell k_0) - \hat{\rho}^l_j(k) \tilde{q}^1_{j-j\ell}(k, \ell k_0, k - \ell k_0) \right). \quad (328) \]
\[
\hat{\mathbf{m}}_{ij}^{(1)}(k) = \hat{\rho}_j'(k) \frac{\hat{q}_{j-j}(k, \ell k_0, k - \ell k_0) + \hat{q}_{-j-j}(k, \ell k_0, k - \ell k_0)}{\hat{\gamma}_{j-j}(k, \ell k_0, k - \ell k_0)},
\]

\[
\hat{\mathbf{m}}_{ij}^{(2)}(k) = -\hat{\rho}_j'(k) \frac{\hat{\gamma}_{j-j}(k, \ell k_0, k - \ell k_0) + \hat{\gamma}_{-j-j}(k, \ell k_0, k - \ell k_0) - \hat{q}_{-j-j}(k, \ell k_0, k - \ell k_0)}{\hat{\gamma}_{j-j}(k, \ell k_0, k - \ell k_0) - \hat{q}_{-j-j}(k, \ell k_0, k - \ell k_0)},
\]

\[
\hat{\mathbf{m}}_{ij}^{(3)}(k) = (\hat{\rho}_j'(k) - \hat{\rho}_{-j}'(k)) \frac{\hat{q}_{-j-j}(k, \ell k_0, k - \ell k_0)}{\hat{\gamma}_{-j-j}(k, \ell k_0, k - \ell k_0)}.
\]

Using (162), (167) and (279) we obtain

\[
\hat{\mathbf{m}}_{ij}^{(1)}(k) = \mathcal{O}(|k|^{-1/2}(1 + b k^2)^{-1/2})
\]

for \(|k| \to \infty\) uniformly on compact subsets of \(B\). Because of

\[
\hat{\gamma}_{j-j}(k, \ell k_0, k - \ell k_0) + \hat{\gamma}_{-j-j}(k, \ell k_0, k - \ell k_0) = 2i \omega(\ell k_0),
\]

(274) and (279) we conclude

\[
\hat{\mathbf{m}}_{ij}^{(2)}(k) = \mathcal{O}(|k|^{-1/2}(1 + b k^2)^{-1/2})
\]

for \(|k| \to \infty\) uniformly on compact subsets of \(B\). Furthermore, we have

\[
\hat{\rho}_j'(k) - \hat{\rho}_{-j}'(k) = \begin{cases} 
-\text{sgn}(j) \sum_{v \in \{\pm 1\}} \hat{\rho}_v^{(j)}(k) & \text{if } b \in (0, b_0), \\
0 & \text{otherwise},
\end{cases}
\]

and

\[
\hat{\rho}_{-j}'(v)(k) = \left( -\frac{\hat{q}_{-j-|j|}(k, v k_0, k - v k_0)}{\hat{q}_{-j-|j|}(k, v k_0, k - v k_0)} - \frac{\hat{q}_{-j-|j|}(k, v k_0, k - v k_0)}{\hat{q}_{-j-|j|}(k, v k_0, k - v k_0)} \left( 1 - \frac{v k_0}{k} \right)^{2l} - 1 \right)
\]

\[
\times \hat{\xi}_1 \left( k + \nu(k_1 - k_0) \right) \frac{k_1 - k_0}{k_1 - k_0} = \mathcal{O}(|k|^{-1})
\]

for \(|k| \to \infty\) uniformly on compact subsets of \(B \cap (0, b_0)\), where the last equality holds because of (162), (166), (211), (214), (272), (273), (306), (309),

\[
\partial_n \hat{q}_{j_1,j_2}^{(2)}(k, \ell k_0, k - \ell k_0) = \mathcal{O}(|k|^{-3/2}(1 + b k^2)^{1/2})
\]
for $n \in \{1, 3\}$ and $|k| \to \infty$ uniformly with respect to $b \lesssim 1$ and the mean value theorem. Hence, due to (274) and (279), we obtain

$$\tilde{m}_{j\ell}^{l,3}(k) = O(|k|^{-1}(1 + bk^2)^{-1/2})$$

(339) for $|k| \to \infty$ uniformly on compact subsets of $B$.

Now, combining (327)–(332), (333), (335) and (339), we arrive at (304).

Finally, (305) is proven in an analogous manner as (302).

Lemma 4.9. The operators $(\rho^l \mathcal{N}^1)^{a}_{j_1j_2}$ with $i \in \{1, 2\}$, $j_1, j_2 \in \{\pm 2\}$ and $l \geq 1$ have the following properties:

a) Fix $\varphi \in L^2(\mathbb{R}, \mathbb{R})$ with supp $\hat{\varphi} = \text{supp } \hat{\psi}_c$. Then $f \mapsto (\rho^l \mathcal{N}^1)^{a}_{j_1j_2}(\varphi) f$ defines a continuous linear map from $H^1(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$ and $f \mapsto (\rho^l \mathcal{N}^1)^{a}_{j_1-j}(\varphi) f$ can be extended to a continuous linear map from $L^2(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$. Furthermore, there exists a constant $C_1 > 0$ with $C_1 \lesssim \|\hat{\varphi}\|_L^1$ such that for all $f \in L^2(\mathbb{R}, \mathbb{R})$ and all $g \in H^1(\mathbb{R}, \mathbb{R})$ it holds that

$$\|(\rho^l \mathcal{N}^1)^{a}_{j_1j_2}(\varphi) g\|_{L^2} \leq C_1 \|g\|_{H^1},$$

(340)

$$\|(\rho^l \mathcal{N}^1)^{a}_{j_1-j}(\varphi) f\|_{L^2} \leq C_1 \|f\|_{L^2}$$

(341)

uniformly on compact subsets of $B$.

b) Let $\varphi$ be as in a). Then there exists a constant $C_2 > 0$ with $C_2 \lesssim \|\hat{\varphi}\|_L^1$ such that for all $g \in H^1(\mathbb{R}, \mathbb{R})$ it holds that

$$(\rho^l \mathcal{N}^1)^{a}_{j_1j_2}(\varphi) g = (\partial_{\alpha}^{-1} \rho^l \mathcal{N}^1)^{s}_{j_1j_2}(\varphi) \partial_{\alpha} g + \tilde{M}^l_j(\varphi) g$$

(342)

with

$$\|\tilde{M}^l_j(\varphi) g\|_{L^2} \leq C_2 \|g\|_{L^2}$$

(343)

uniformly on compact subsets of $B$. Moreover, for all $\psi \in H^{2+p}(\mathbb{R}, \mathbb{R})$ with $p > 1/2$ there exists a constant $C_3 > 0$ with $C_3 \lesssim \|\hat{\psi}\|_L^1 \|\hat{\varphi}\|_L^1 \|\partial_{\alpha} \psi\|_{H^1+p}$ such that for all $f \in L^2(\mathbb{R}, \mathbb{R})$ it holds that

$$\|f, (\partial_{\alpha}^{-1} \rho^l \mathcal{N}^1)^{s}_{j_1j_2}(\varphi) \|_{H^1} \leq C_3 \|f\|_{L^2}$$

(344)

uniformly on compact subsets of $B$.

c) Let $\varphi$ be as in a). Then $f \mapsto (\rho^l \mathcal{N}^2)^{a}_{j_1j_2}(\varphi) f$ defines a continuous linear map from $L^2(\mathbb{R}, \mathbb{R})$ into $L^2(\mathbb{R}, \mathbb{R})$ and there exists a constant $C_4 > 0$ with $C_4 \lesssim \|\hat{\varphi}\|_L^1$ such that for all $f \in L^2(\mathbb{R}, \mathbb{R})$ it holds that

$$\|(\rho^l \mathcal{N}^2)^{a}_{j_1j_2}(\varphi) f\|_{L^2} \leq C_4 \|f\|_{L^2}$$

(345)

uniformly on compact subsets of $B$.

Proof. a) follows directly from Lemmas 4.4 a) and 4.8 a).

b) It holds that

$$(\rho^l \mathcal{N}^1)^{a}_{j_1j_2}(k, m)$$
\[
\begin{align*}
&= \frac{1}{2}((ik)^{-1} \hat{\rho}'_j(k) \hat{\vartheta}(k) \hat{n}^1_{ij\ell}(k) + (i(-m))^{-1} \hat{\rho}'_j(-m) \hat{\vartheta}(-m) \hat{n}^1_{ij\ell}(-m)) i m \\
&+ \frac{1}{2} \hat{\rho}'_j(k) \hat{\vartheta}(k) (ik)^{-1} \hat{n}^1_{ij\ell}(k) i (k - m),
\end{align*}
\]

which, due to (280), implies (342)–(343).

(344) is proven in an analogous manner as (302).

c) By construction of \( \hat{\rho}'_j \xi_{jj\ell} \) and by using Taylor’s theorem as well as (166), (211), (275), (278), (314)–(315), (338) and

\[
\partial_n \partial_{n_2} \hat{q}^{2,4}_{j_1j_2}(k, \ell k_0, k - \ell k_0) = \mathcal{O}(|k|^{-5/2}(1 + bk^2)^{1/2}),
\]

for \( n_1, n_2 \in \{1, 3\} \) and \(|k| \to \infty\) uniformly with respect to \( b \lesssim 1 \) we obtain

\[
\frac{1}{2} \partial^2 \hat{q}^{2,4}_{j_1j_2}(k, m) \chi_{\ell}(k - m)
\]

\[
= \frac{\hat{\rho}'_j(-k) \hat{\xi}_{jj\ell}(-k) \partial_1 \hat{q}^2_{jj}(k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{jj\ell}(k)} (k - m) \\
+ \frac{\left( \frac{d}{dk} (\hat{\rho}'_j \xi_{jj\ell}) \right)(-k) \hat{q}^2_{jj}(k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{jj\ell}(k)} (k - m) \\
+ \frac{\hat{\rho}'_j(-k) \hat{\xi}_{jj\ell}(-k) \partial_3 \hat{q}^2_{jj}(k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{jj\ell}(k)} (k - m - 2\ell k_0) \\
- \frac{\hat{\rho}'_j(-k) \xi_{jj\ell}(-k) \hat{q}^2_{jj}(k, \ell k_0, -k + \ell k_0)}{2\hat{r}_{jj\ell}(k)} \frac{d}{dk} \hat{r}_{jj\ell}(k) - (k - m - \ell k_0) \\
+ \mathcal{O}(|k|^{-1}) \\
= \mathcal{O}(1)
\]

for \(|k| \to \infty\) uniformly with respect to \( m \in \mathbb{R} \) and \( b \) in compact subsets of \( \mathcal{B} \), which implies (345).

**Lemma 4.10.** The operators \( T_{j_1j_2} \) have the following properties:

a) Fix functions \( g, h \) with \( \hat{g}, \hat{h} \in L^1(\mathbb{R}, \mathbb{C}) \) and \( \supp \hat{g} = \supp \hat{h} = \supp \hat{\psi}_c \). Then \( f \mapsto T_{j_1j_2}(g, h, f) \) defines a continuous linear map from \( L^2(\mathbb{R}, \mathbb{C}) \) into \( L^2(\mathbb{R}, \mathbb{C}) \), and for all \( f \in L^2(\mathbb{R}, \mathbb{C}) \) we have

\[
\| T_{j_1j_2}(g, h, f) \|_{L^2} \lesssim \varepsilon^{-1} \| \hat{g} \|_{L^1} \| \hat{h} \|_{L^1} \| f \|_{L^2},
\]

uniformly on compact subsets of \( \mathcal{B} \). If \( f, g \) and \( h \) are real-valued, then \( T_{j_1j_2}(g, h, f) \) is also real-valued.

b) For sufficiently small \( \varepsilon > 0 \) we have

\[
\begin{align*}
\text{sgn}(j_1) \iota \omega T^1_{j_1j_2}(\psi_c, \psi_c, R_{j_2}) + T^1_{j_1j_2}(i \omega \psi_c, \psi_c, R_{j_2}) \\
+ T^1_{j_1j_2}(\psi_c, i \omega \psi_c, R_{j_2}) \text{sgn} (j_2) T^1_{j_1j_2}(\psi_c, \psi_c, i \omega R_{j_2}) \\
= \varepsilon^{-1} C_{j_1j_2}(\psi_c, \psi_c) R_{j_2} + Y^1_{j_1j_2}(\psi_c, \psi_c, R_{j_2})
\end{align*}
\]

with

\[
\| Y^1_{j_1j_2}(\psi_c, \psi_c, R_{j_2}) \|_{L^2} = \mathcal{O}(\| R_{j_2} \|_{L^2} + \| R_{j_2} \|_{H^{3/2}}).
\]
uniformly on compact subsets of $\mathcal{B}$ if $j_1$, $j_2 \in \{ \pm 1 \}$, and

\[
\text{sgn}(j_1)i\omega T_{j_1j_2}^2(\psi_c, \psi_c, R_{j_2}) + T_{j_1j_2}^2(i\omega \psi_c, \psi_c, R_{j_2})
\]

\[
+ T_{j_1j_2}^2(\psi_c, i\omega \psi_c, R_{j_2}) - \text{sgn}(j_2)T_{j_1j_2}^2(\psi_c, \psi_c, i\omega R_{j_2})
\]

\[
= \sum_{j_3 \in \{ \pm |j_1| \}} N_{j, j_3}^{j_1}(\psi_c, Q_{j_3j_2}(\psi_c)R_{j_2}) + Y_{j_1j_2}^1(\psi_c, \psi_c, R_{j_2})
\]  \(352\)

with

\[
\|Y_{j_1j_2}^1(\psi_c, \psi_c, R_{j_2})\|_{L^2} = \mathcal{O}(\|R_{j_2}\|_{H^2})
\]  \(353\)

uniformly on compact subsets of $\mathcal{B}$.

c) For all $f \in \mathcal{L}^2(\mathbb{R}, \mathbb{C})$ we have

\[
P_{\delta_0, \infty} T_{j_1j_2}(\psi_c, \psi_c, f) = 0.
\]  \(354\)

Proof. a) Because of (24), (129), (242)–(247) and (262) we have

\[
\|\hat{T}_{j_1j_2}\|_{L^\infty} = \mathcal{O}(\varepsilon^{-1})
\]

uniformly on compact subsets of $\mathcal{B}$ for $j_1 \in \{ \pm 1, \pm 2 \}$, $j_2 \in \{ \pm j_1 \}$, $|j_1| \leq j \leq 2$ and $\ell \in \{ \pm 1 \}$. With the help of Young’s inequality for convolutions we obtain

\[
\|T_{j_1j_2}(g, h, f)\|_{L^2} \leq \sum_{|j_1| \leq j \leq 2, \ell \in \{ \pm 1 \}} \|\hat{T}_{j_1j_2}\|_{L^\infty} \|\hat{g}\|_{L^1} \|\hat{h}\|_{L^1} \|f\|_{L^2} \lesssim \varepsilon^{-1} \|\hat{g}\|_{L^1} \|\hat{h}\|_{L^1} \|f\|_{L^2}
\]

uniformly on compact subsets of $\mathcal{B}$. Furthermore, since

\[
\hat{T}_{j_1j_2, -\ell}(k) = \overline{\hat{T}_{j_1j_2, \ell}(k)},
\]

we conclude that $T_{j_1j_2}(g, h, f)$ is real-valued if $f$, $g$ and $h$ are real-valued.

b) To prove the first assertion of b), we first show that it holds that

\[
\varepsilon^{-1} C_{j_1j_2}(\psi_c, \psi_c)R_{j_2}
\]

\[
= \sum_{\ell \in \{ \pm 1 \}} \varepsilon^{-1} P_{\delta_0, \infty} C_{j_1j_2}(\psi_\ell, \psi_\ell, R_{j_2}) + \mathcal{O}(\|R_{j_2}\|_{L^2} + \|R_{j_2}\|_{H^{3/2}})
\]  \(355\)

uniformly on compact subsets of $\mathcal{B}$ such that it is sufficient to prove that the $L^2$-norm of

\[
\tilde{Y}_{j_1j_2}^1 := \sum_{\ell \in \{ \pm 1 \}} \left(\text{sgn}(j_1)i\omega T_{j_1j_2, \ell}(\psi_\ell, \psi_\ell, R_{j_2}) + T_{j_1j_2, \ell}(i\omega \psi_\ell, \psi_\ell, R_{j_2})
\]

\[
+ T_{j_1j_2, \ell}(\psi_\ell, i\omega \psi_\ell, R_{j_2}) - \text{sgn}(j_2)T_{j_1j_2, \ell}(\psi_\ell, i\omega R_{j_2})
\]

\[
- \varepsilon^{-1} P_{\delta_0, \infty} C_{j_1j_2}(\psi_\ell, \psi_\ell)R_{j_2}\right)
\]

\[
\|\tilde{Y}_{j_1j_2}^1\|_{L^2} = \mathcal{O}(\|R_{j_2}\|_{H^3/2})
\]
is of order $O(||R_{j_2}||_{L^2})$ uniformly on compact subsets of $B$, which we will obtain by construction of $T_{j_1j_2}^1$ and because of Lemma 4.2. To verify (355), we split $\varepsilon^{-1}C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2}$ into

$$\varepsilon^{-1}C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2} = \sum_{\ell \in \{\pm 1\}} \varepsilon^{-1}P_{0, \delta_0}C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2}$$

$$+ \sum_{\ell \in \{\pm 1\}} \varepsilon^{-1}P_{0, \delta_0}C_{j_1j_2}(\psi_\epsilon, \psi_{-\ell})R_{j_2}$$

$$+ \varepsilon^{-1}P_{0, \infty}C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2}.$$  

Due to (123), (130), (140) and (164) we have

$$\|\varepsilon^{-1}P_{0, \infty}C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2}\|_{L^2} = O(||R_{j_2}\|_{L^2} + ||R_{2j_2}\|_{H^{3/2}})$$

uniformly on compact subsets of $B$.

It follows from (52)–(58) that each summand of $C_{j_1j_2}(\psi_\epsilon, \psi_\epsilon)R_{j_2}$ contains at least one $\alpha$-derivative. Using this fact as well as $\psi_\epsilon, \partial_\alpha \psi_\epsilon = (\partial_\alpha (\psi_\epsilon^2))/2$ and the inequality $|n| \leq |k| + |k - n|$, we obtain

$$\varepsilon^{-1}F(P_{0, \delta_0}C_{j_1j_2}(\psi_\epsilon, \psi_{-\ell})R_{j_2})(k)$$

$$= \widehat{P}_{0, \delta_0}(k) \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{\phi}^{-1}(k) \hat{c}_{j_1j_2}(k, k - m, m - n, n) \hat{\psi}_\epsilon(k - m)$$

$$\hat{\psi}_{-\ell}(m - n) \widehat{R}_{j_2}(n) dndm$$

with

$$|\hat{c}_{j_1j_2}(k, k - m, m - n, n)| \lesssim |k| + |k - n|$$

(356) uniformly on compact subsets of $B$. (129), (131) and (356), as well as Fubini’s theorem, yield

$$|\widehat{P}_{0, \delta_0}(k) \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{\phi}^{-1}(k) \hat{c}_{j_1j_2}(k, k - m, m - n, n)$$

$$\hat{\psi}_\epsilon(k - m) \hat{\psi}_{-\ell}(m - n) \widehat{R}_{j_2}(n) dndm|$$

$$\lesssim \int_{\mathbb{R}} \int_{\mathbb{R}} \widehat{P}_{0, \delta_0}(k) (1 + \frac{|k - n|}{\varepsilon}) |\hat{\psi}_\epsilon(k - m)| |\hat{\psi}_{-\ell}(m - n)||\widehat{R}_{j_2}(n)| dndm$$

$$= \int_{\mathbb{R}} \widehat{P}_{0, \delta_0}(k) (1 + \frac{|k - n|}{\varepsilon}) (|\hat{\psi}_\epsilon| * |\hat{\psi}_{-\ell}|)(k - n) \widehat{R}_{j_2}(n) dn$$

uniformly on compact subsets of $B$. Because of (112) and (118), the function $|\hat{\psi}_\epsilon| * |\hat{\psi}_{-\ell}|$ is strongly concentrated near 0, more precisely, it has a compact support being independent of $\varepsilon$ and is of the form

$$(|\hat{\psi}_\epsilon| * |\hat{\psi}_{-\ell}|)(k) = \varepsilon^{-1} \hat{g}(\varepsilon^{-1} k)$$

for all $k \in \mathbb{R}$, with $\hat{g} \in L^1(1)$. Hence, by using (123) and Young’s inequality for convolutions, we conclude that

$$\|\varepsilon^{-1}P_{0, \delta_0}C_{j_1j_2}(\psi_\epsilon, \psi_{-\ell})R_{j_2}\|_{L^2} = O(||R_{j_2}\|_{L^2})$$
uniformly on compact subsets of $\mathcal{B}$. Therefore, we have verified (355).

To estimate $\|\tilde{Y}^1_{j_1 j_2}\|_{L^2}$, we use

$$\tilde{Y}^1_{j_1 j_2}(k) = \sum_{\ell \in \{-1, 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} K^1_{j_1 j_2}(k, k - m, m - n, n, \ell k_0, k - 2\ell k_0) \psi\ell(k - m) \psi\ell(m - n) \tilde{R}_{j_2}(n) \, d\nu \, d\mu,$$

with

$$K^1_{j_1 j_2}(k, k - m, m - n, n) = \tilde{P}_{0, \delta_0}^{\ell}(k) \tilde{v}_{j_1 j_2}(k, k - m, m - n, n) \tilde{c}_{j_1 \ell j_2}(k, \ell k_0, \ell k_0, k - 2\ell k_0) \hat{\vartheta}(n)$$

We split $K^1_{j_1 j_2}$ into

$$K^1_{j_1 j_2}(k, k - m, m - n, n) = K^1_{j_1 j_2}(k, k - m, m - n, n) - K^1_{j_1 j_2}(k, \ell k_0, m - n, n) + K^1_{j_1 j_2}(k, \ell k_0, m - n, n) - K^1_{j_1 j_2}(k, \ell k_0, \ell k_0, k - 2\ell k_0) + K^1_{j_1 j_2}(k, \ell k_0, \ell k_0, k - 2\ell k_0).$$

Since

$$K^1_{j_1 j_2}(k, \ell k_0, \ell k_0, k - 2\ell k_0) = 0,$$

we deduce by applying (123) and Lemma 4.2 that it holds that

$$\|\tilde{Y}^1_{j_1 j_2}\|_{L^2} = O(\|\tilde{R}_{j_2}\|_{L^2})$$

uniformly on compact subsets of $\mathcal{B}$. Hence, we have proven the first assertion of b).

To prove the second assertion of b), we first show that it holds that

$$\mathcal{N}_{j_1 j_3}(\psi\ell, Q_{j_3 j_2}(\psi\ell) R_{j_2}) = \sum_{\ell \in \{-1, 1\}} P_{0, \delta_0}^{j_1 j_3}(\psi\ell, Q_{j_3 j_2}(\psi\ell) R_{j_2}) + O(\|\tilde{R}_{j_2}\|_{H^2})$$

(357)

uniformly on compact subsets of $\mathcal{B}$ such that it is sufficient to prove that the $L^2$-norm of

$$\tilde{Y}^2_{j_1 j_2} := \sum_{\ell \in \{-1, 1\}} (\text{sgn}(j_1) i \omega T^2_{j_1 j_2}(\psi\ell, \psi\ell, R_{j_2}) + T^2_{j_1 j_2}(i \omega \psi\ell, \psi\ell, R_{j_2}) + T^2_{j_1 j_2}(\psi\ell, i \omega \psi\ell, R_{j_2}) - \text{sgn}(j_2) T^2_{j_1 j_2}(\psi\ell, \psi\ell, i \omega R_{j_2})$$
\[
- \sum_{j_3 \in \{\pm |j_1|\}} P_{0, \delta_0} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2})
\]

is of order \(O(\|R_{j_2}\|_{L^2})\) uniformly on compact subsets of \(B\), which we will obtain by construction of \(T_{j_1 j_2}^z\), and because of Lemma 4.2.

To verify (357), we split \(N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2})\) into

\[
N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2}) = \sum_{\ell \in \{\pm 1\}} P_{0, \delta_0} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2})
+ \sum_{\ell \in \{\pm 1\}} P_{0, \delta_0} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell - \ell) R_{j_2})
+ P_{\delta_0, \infty} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2}).
\]

Because of (123), (153)–(163), (255)–(256) and Young’s inequality for convolutions we conclude that the \(L^2\)-norm of the last summand is of order \(O(\|R_{j_2}\|_{H^2})\) uniformly on compact subsets of \(B\). Moreover, due to Lemma 4.2 and (296), it holds that

\[
\mathcal{F} P_{0, \delta_0} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_\ell) R_{j_2}) (k)
= \int_{\mathbb{R}} \int_{\mathbb{R}} K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, j k_0, k)
- (\ell + j) k_0) \tilde{\psi}_\ell (k - m) \tilde{\psi}_j (m - n) \tilde{R}_{j_2} (n) \, dn \, dm
+ O(\|R_{j_2}\|_{L^2})
\]

uniformly on compact subsets of \(B\), where

\[
K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, j k_0, k - (\ell + j) k_0)
= \tilde{P}_{0, \delta_0} (k) \frac{\tilde{q}_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, j k_0, k - (\ell + j) k_0) \tilde{\vartheta} (k - (\ell + j) k_0)}{\tilde{r}_{j_1 j_3} (k, \ell k_0, k - \ell k_0) \tilde{\vartheta} (k)}.
\]

If \(j = -\ell\), then we have

\[
K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, j k_0, k - (\ell + j) k_0)
= K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, -\ell k_0, k)
= \tilde{P}_{0, \delta_0} (k) \frac{\tilde{q}_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, -\ell k_0, k) \tilde{\vartheta} (k)}{\tilde{r}_{j_1 j_3} (k, \ell k_0, k - \ell k_0) \tilde{\vartheta} (k)}
\]

and the factor \(\tilde{\vartheta} (k)\) in the denominator is canceled by the same factor in the numerator. \(K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, -\ell k_0, k)\) contains no factors which are of order \(O(\varepsilon^{-1})\) such that there holds

\[
\sup_{k \in \mathbb{R}} |K_{j_1 j_2 j_3} (k, \ell k_0, k - \ell k_0, -\ell k_0, k)| = O(1)
\]

uniformly on compact subsets of \(B\). Hence, we obtain

\[
\|\mathcal{F} N_{j_1 j_3} (\psi_\ell, Q_{j_3 j_2} (\psi_{-\ell}) R_{j_2})\|_{L^2} = O(\|R_{j_2}\|_{L^2})
\]
uniformly on compact subsets of \( \mathcal{B} \) such that we have verified (357).

We have

\[
\hat{Y}^2_{j_1j_2}(k) = \sum_{\ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} K^2_{j_1j_2}(k, k-m, m-n, n, \ell k_0, k-\ell k_0, k-2\ell k_0) \hat{\psi}_\ell(k-m) \times \\
\times \hat{\psi}_\ell(m-n) \tilde{R}_{j_2}(n) dndm \\
+ O(\|R_{j_2}\|_{L^2})
\]

uniformly on compact subsets of \( \mathcal{B} \), where

\[
K^2_{j_1j_2}(k, k-m, m-n, n, \ell k_0, k-\ell k_0, k-2\ell k_0)
= \sum_{j \in \{\pm 1\}} \left( P_{0, \delta_0}(k) \hat{q}_{j_1j_2j_3}(k, \ell k_0, k-\ell k_0, \ell k_0, k-2\ell k_0) \tilde{\varphi}(k-2\ell k_0) \times \\
\times (\hat{\psi}_{j_1j_2}(k, k-m, m-n, n) - \hat{\psi}_{j_1j_2}(k, \ell k_0, \ell k_0, k-2\ell k_0)) \right).
\]

Now, we can apply again Fubini’s theorem, Young’s inequality for convolutions and Lemma 4.2 to obtain

\[
\|\hat{Y}^2_{j_1j_2}\|_{L^2} = O(\|R_{j_2}\|_{L^2})
\]

uniformly on compact subsets of \( \mathcal{B} \). Hence, we have proven the second assertion of b).

c) follows directly by the definition of \( T_{j_1j_2} \). \( \square \)

Now, we are able to compare our energy with Sobolev norms of the error. We obtain

**Lemma 4.11.** For sufficiently small \( \varepsilon > 0 \), we have

\[
\sum_{j \in \{\pm 1, \pm 2\}} \|\tilde{R}_j\|_{L^2}^2 \lesssim \sum_{j \in \{\pm 1, \pm 2\}} \|R_j\|_{L^2}^2 + \|\partial_\alpha \mathcal{R}_2\|_{(H^1)^2}^2 + \varepsilon^4, \quad (358)
\]

\[
\sum_{j \in \{\pm 1, \pm 2\}} \|R_j\|_{L^2}^2 \lesssim \sum_{j \in \{\pm 1, \pm 2\}} \|\tilde{R}_j\|_{L^2}^2 + \varepsilon \|\partial_\alpha \mathcal{R}_2\|_{(H^1)^2}^2 + \varepsilon^5, \quad (359)
\]

as long as \( \varepsilon^{5/2}\|\mathcal{R}_1\|_{(L^2)^2}, \varepsilon^{5/2}\|\mathcal{R}_2\|_{(H^2)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

**Proof.** Estimate (358) follows from the estimates (146)–(148), (292)–(293) and (349).

To prove (359) we introduce \( R^0_{j_1} := P_{0, \delta_0} R_{j_1}, \hat{R}^0_{j_1} := P_{0, \delta_0} \hat{R}_{j_1}, R^1_{j_1} := P_{\delta_0, \infty} R_{j_1}, \hat{R}^1_{j_1} := P_{\delta_0, \infty} \hat{R}_{j_1} \) and split \( R_{j_1}, \hat{R}_{j_1} \) into \( R_{j_1} = R^0_{j_1} + R^1_{j_1} \) and \( \hat{R}_{j_1} = \hat{R}^0_{j_1} + \hat{R}^1_{j_1} \). Because of (296) and (354), \( R^0_{j_1} \) satisfies
\[ R^0_{j_1} + \varepsilon^2 \sum_{j_2 \in \{\pm j_1\}} T_{j_1j_2}(\psi_c, \psi_c, R^0_{j_2}) \]
\[ = \tilde{R}^0_{j_1} - \varepsilon P_{0, \delta_0} \sum_{j_2 \in \{\pm j_1\}} N_{j_1j_2}(\psi_c, R^0_{j_2}) - \varepsilon^2 \sum_{j_2 \in \{\pm j_1\}} T_{j_1j_2}(\psi_c, \psi_c, R^1_{j_2}). \] (360)

Multiplying this equation with \( R^0_{j_1} \), integrating, summing over \( j_1 \in \{\pm 1, \pm 2\} \) and using (349) yields
\[ \sum_{j_1 \in \{\pm 1, \pm 2\}} \| R^0_{j_1} \|_{L^2}^2 \lesssim \sum_{j_1 \in \{\pm 1, \pm 2\}} \| \tilde{R}^0_{j_1} \|_{L^2} + \| R^1_{j_1} \|_{L^2} \] (361)

uniformly on compact subsets of \( \mathcal{B} \) for sufficiently small \( \varepsilon > 0 \). Moreover, \( R^1_{j_1} \) satisfies
\[ R^1_{j_1} + \varepsilon P_{\delta_0, \infty} \sum_{j_2 \in \{\pm j_1\}} N_{j_1j_2}(\psi_c, R^1_{j_2}) = \tilde{R}^1_{j_1} - \varepsilon P_{\delta_0, \infty} \sum_{j_2 \in \{\pm j_1\}} N_{j_1j_2}(\psi_c, R^0_{j_2}). \] (362)

Multiplying this equation with \( R^1_{j_1} \), integrating, summing over \( j_1 \in \{\pm 1, \pm 2\} \) and using (294)–(295) yields
\[ \sum_{j_1 \in \{\pm 1, \pm 2\}} \| R^1_{j_1} \|_{L^2}^2 \lesssim \sum_{j_1 \in \{\pm 1, \pm 2\}} \| \tilde{R}^1_{j_1} \|_{L^2} \| R^1_{j_1} \|_{L^2} \]
\[ + \varepsilon \sum_{j_1 \in \{\pm 1, \pm 2\}, j_2 \in \{\pm j_1\}} (\| R^0_{j_2} \|_{L^2} + \| R^1_{j_2} \|_{H^1}) \| R^1_{j_1} \|_{L^2} \]
\[ + \varepsilon \sum_{j_1 \in \{\pm 2\}} (\| \partial_\alpha^{-1} R_{1 - 2} \|_{L^2} + \| \partial_\alpha^{-1} R_2 \|_{L^2}) \| R^1_{j_1} \|_{L^2} \]

uniformly on compact subsets of \( \mathcal{B} \). With the help of (146)–(148), (361) we deduce
\[ \sum_{j_1 \in \{\pm 1, \pm 2\}} \| R^1_{j_1} \|_{L^2}^2 \lesssim \sum_{j_1 \in \{\pm 1, \pm 2\}} \| \tilde{R}^1_{j_1} \|_{L^2}^2 + \varepsilon \| R^0_{j_1} \|_{L^2}^2 + \varepsilon \| R^1_{j_1} \|_{H^1}^2 \]
\[ + \varepsilon \| \partial_\alpha^{-1} R_{1 - 2} \|_{L^2}^2 + \varepsilon \| \partial_\alpha^{-1} R_2 \|_{L^2}^2 \]
\[ \lesssim \sum_{j_1 \in \{\pm 1, \pm 2\}} \| \tilde{R}^1_{j_1} \|_{L^2}^2 + \varepsilon \| \partial_\alpha R_2 \|_{(H^1)^2}^2 + \varepsilon^5 \] (363)

and
\[ \sum_{j_1 \in \{\pm 1, \pm 2\}} \| R^0_{j_1} \|_{L^2}^2 \lesssim \sum_{j_1 \in \{\pm 1, \pm 2\}} \| \tilde{R}^1_{j_1} \|_{L^2}^2 + \varepsilon \| \partial_\alpha R_2 \|_{(H^1)^2}^2 + \varepsilon^5 \] (364)

uniformly on compact subsets of \( \mathcal{B} \) for sufficiently small \( \varepsilon > 0 \). Combining (363) and (364) yields (359).

The analysis of \( \mathcal{E}_{2,s} \) will be simplified by
Lemma 4.12. Let \( f \in H^l(\mathbb{R}, \mathbb{R}) \) and \( g \in H^m(\mathbb{R}, \mathbb{R}) \) with \( l, m \geq 0 \). Then we have

\[
\int_{\mathbb{R}} \partial_{\alpha}^l f \partial_{\alpha}^m g \, d\alpha = \int_{\mathbb{R}} \partial_{\alpha}^l f \partial_{\alpha}^m g \, d\alpha + \mathcal{O}(\|f\|_{L^2} \|g\|_{L^2}).
\]

(365)

\[
\int_{\mathbb{R}} \partial_{\alpha}^l f \partial_{\alpha}^m g^{-1} \, d\alpha = \int_{\mathbb{R}} \partial_{\alpha}^l f \partial_{\alpha}^m g^{-1} \, d\alpha + \mathcal{O}(\|f\|_{L^2} \|g\|_{L^2}).
\]

(366)

Proof. The proof is analogous to the proof of Lemma 4.4 in [12].  

We obtain

Lemma 4.13. For sufficiently small \( \varepsilon > 0 \) and \( 2 \leq \tilde{s} \leq s \), we have

\[
\mathcal{E}_{\tilde{s}} \lesssim \|R_1\|_{(L^2)^2}^2 + \|R_2\|_{(H^{\tilde{s}})^2}^2 + \varepsilon^5,
\]

(367)

\[
\|R_1\|_{(L^2)^2}^2 + \|R_2\|_{(H^{\tilde{s}})^2}^2 \lesssim \mathcal{E}_{\tilde{s}} + \varepsilon^5,
\]

(368)

as long as \( \varepsilon^{5/2} \|R_1\|_{(L^2)^2} \), \( \varepsilon^{5/2} \|R_2\|_{(H^{\tilde{s}})^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Proof. Because of (225), (148)–(147), Leibniz’s rule, Lemmas 4.4 and 4.12 there holds

\[
\mathcal{E}_{2,l} = \frac{1}{2} \sum_{j_1 \in \{\pm 1\}} \int_{\mathbb{R}} \partial_{\alpha}^l R_{j_1} \partial_{\alpha}^l \partial_{\alpha}^l R_{j_1} \, d\alpha + \varepsilon \mathcal{O}(\|R_1\|_{(L^2)^2}^2 + \|R_2\|_{(H^{\max(2,l)})^2}^2 + \varepsilon^4)
\]

(369)

uniformly on compact subsets of \( \mathcal{B} \) for all \( 1 \leq l \leq s \). Moreover, we have

\[
\langle \partial_{\alpha}^l R_2, (\rho^l \eta^1)(\psi) \partial_{\alpha}^l R_2 \rangle_{(L^2)^2} = \langle \partial_{\alpha}^l R_2, (\rho^l \eta^1) s(\psi) \partial_{\alpha}^l R_2 \rangle_{(L^2)^2}.
\]

(370)

Hence, due to Lemmas 4.8, 4.11, 4.12, (146) and (225), we obtain (367)–(368) uniformly on compact subsets of \( \mathcal{B} \).  

4.4. The Energy Estimates

Now, we are prepared to estimate \( \frac{d}{dt} \mathcal{E}_{\tilde{s}} \). First, we show

Lemma 4.14. For sufficiently small \( \varepsilon > 0 \), we have

\[
\frac{d}{dt} \mathcal{E}_0 \lesssim \varepsilon^2 (\mathcal{E}_2 + 1),
\]

(371)

as long as \( \varepsilon^{5/2} \|R_1\|_{(L^2)^2} \), \( \varepsilon^{5/2} \|R_2\|_{(H^{\tilde{s}})^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).
Proof. Because of (149) and (220) and since $\rho_{j_1}^0$ is symmetric, we have

$$\frac{d}{dt} \mathcal{E}_0 = \sum_{j_1 \in \{\pm 1, \pm 2\}} \int_{\mathbb{R}} \tilde{R}_{j_1} \rho_{j_1}^0 \partial_t \tilde{R}_{j_1} \, d\alpha$$

with

$$\partial_t \tilde{R}_{j_1} = \text{sgn}(j_1) i \omega \tilde{R}_{j_1} + \epsilon^2 \sum_{k=1}^{12} F_{j_1}^k (\Psi, \mathcal{R}),$$

where

$$F_{j_1}^1 (\Psi, \mathcal{R}) = \epsilon^{-1} \sum_{j_2 \in \{\pm j_1\}} \left( Q_{j_1,j_2}(\psi_c) R_{j_2} - \text{sgn}(j_1) i \omega N_{j_1,j_2}(\psi_c, R_{j_2}) 
- N_{j_1,j_2}(i \omega \psi_c, R_{j_2}) + \text{sgn}(j_2) N_{j_1,j_2}(\psi_c, i \omega R_{j_2}) \right),$$

$$F_{j_1}^2 (\Psi, \mathcal{R}) = \epsilon^{-9/2} \partial^{-1} \text{res}_{\text{sgn}(j_1)} (\epsilon \Psi),$$

$$F_{j_1}^3 (\Psi, \mathcal{R}) = \epsilon^{-9/2} \sum_{j_2 \in \{\pm j_1\}} N_{j_1,j_2}(\partial_t \psi_c + i \omega \psi_c, R_{j_2}),$$

$$F_{j_1}^4 (\Psi, \mathcal{R}) = \epsilon^{-7/2} \sum_{j_2 \in \{\pm j_1\}} N_{j_1,j_2}(\psi_c, \partial^{-1} \text{res}_{\text{sgn}(j_2)} (\epsilon \Psi)),$$

$$F_{j_1}^5 (\Psi, \mathcal{R}) = \sum_{j_2 \in \{\pm j_1\}} \left( \epsilon^{-1} C_{j_1,j_2}(\psi_c, \psi_c) R_{j_2} + \sum_{j_3 \in \{\pm j_1\}} N_{j_1,j_3}(\psi_c, Q_{j_3,j_2}(\psi_c) R_{j_2}) 
- \text{sgn}(j_1) i \omega T_{j_1,j_2}(\psi_c, \psi_c, R_{j_2}) - T_{j_1,j_2}(i \omega \psi_c, \psi_c, R_{j_2}) 
- T_{j_1,j_2}(\psi_c, i \omega \psi_c, R_{j_2}) + \text{sgn}(j_2) T_{j_1,j_2}(\psi_c, \psi_c, i \omega R_{j_2}) \right),$$

$$F_{j_1}^6 (\Psi, \mathcal{R}) = \sum_{j_2 \in \{\pm j_1\}} T_{j_1,j_2}(\partial_t \psi_c + i \omega \psi_c, \psi_c, R_{j_2}) + T_{j_1,j_2}(\psi_c, \partial_t \psi_c + i \omega \psi_c, R_{j_2}),$$

$$F_{j_1}^7 (\Psi, \mathcal{R}) = \mathcal{M}_{j_1}(\Psi, \mathcal{R}),$$

$$F_{j_1}^8 (\Psi, \mathcal{R}) = \sum_{j_2, j_3 \in \{\pm j_1\}} N_{j_1,j_2}(\psi_c, C_{j_3,j_2}(\psi_c, \psi_c) R_{j_2}),$$

$$F_{j_1}^9 (\Psi, \mathcal{R}) = \epsilon \sum_{j_2 \in \{\pm j_1\}} N_{j_1,j_2}(\psi_c, \mathcal{M}_{j_2}(\Psi, \mathcal{R})),$$

$$F_{j_1}^{10} (\Psi, \mathcal{R}) = \epsilon \sum_{j_2, j_3 \in \{\pm j_1\}} T_{j_1,j_2}(\psi_c, \psi_c, Q_{j_3,j_2}(\psi_c) R_{j_2} + C_{j_3,j_2}(\psi_c, \psi_c) R_{j_2}),$$

$$F_{j_1}^{11} (\Psi, \mathcal{R}) = \epsilon^2 \sum_{j_2 \in \{\pm j_1\}} T_{j_1,j_2}(\psi_c, \psi_c, \mathcal{M}_{j_2}(\Psi, \mathcal{R})),$$

$$F_{j_1}^{12} (\Psi, \mathcal{R}) = \epsilon^{-5/2} \sum_{j_2 \in \{\pm j_1\}} T_{j_1,j_2}(\psi_c, \psi_c, \partial^{-1} \text{res}_{\text{sgn}(j_2)} (\epsilon \Psi)).$$

Due to the skew symmetry of $i \omega$ we obtain

$$\frac{d}{dt} \mathcal{E}_0 = \epsilon^2 \sum_{k=1}^{12} \sum_{j_1 \in \{\pm 1, \pm 2\}} \int_{\mathbb{R}} \tilde{R}_{j_1} \rho_{j_1}^0 F_{j_1}^k (\Psi, \mathcal{R}) \, d\alpha.$$
Because of (146)–(148), (225), (292)–(293), (298)–(299) and (349) we deduce
\[ \left| \sum_{j_i \in \{\pm 1, \pm 2\}} \int_{\mathbb{R}} \tilde{R}_i \rho_{j_i}^0 F^1_{j_i} (\Psi, \mathcal{R}) \, d\alpha \right| \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4 \]
and with the help of the Cauchy-Schwarz inequality and (225) we conclude
\[ \frac{d}{dt} \mathcal{E}_0 \lesssim \varepsilon^2 \left( \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4 + \mathcal{E}_0 + \sum_{k=2}^{12} \sum_{j_i \in \{\pm 1, \pm 2\}} \| F^k_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 \right). \]
Because of (121), (129)–(130), (292)–(295), (296) and (349) we have
\[ \| F^2_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2} \lesssim 1, \]
\[ \| F^3_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2} \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4. \]
Due to (146)–(148) and (350)–(353) we deduce
\[ \| F^5_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2} \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4. \]
Furthermore, (127), (146)–(148), (292)–(293) and (349) yield
\[ \| F^3_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2} + \| F^6_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2} \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4. \]
Using (141), (146)–(148), (292)–(293) and (349), we obtain
\[ \| F^7_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 + \| F^9_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 + \| F^{11}_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4. \]
Finally, (129)–(130), (140), (146)–(148), (153)–(157), (164), (292)–(295), (296) and (349) imply
\[ \| F^8_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 + \| F^{10}_{j_i} (\Psi, \mathcal{R}) \|_{L_2^2}^2 \lesssim \| \mathcal{R}_1 \|_{(L_2^2)^2}^2 + \| \mathcal{R}_2 \|_{(H_2^2)^2}^2 + \varepsilon^4. \]
All bounds are uniform on compact subsets of $\mathcal{B}$.

Hence, because of (368) we arrive at
\[ \frac{d}{dt} \mathcal{E}_0 \lesssim \varepsilon^2 (\mathcal{E}_2 + 1) \]
as long as $\varepsilon^{5/2} \| \mathcal{R}_1 \|_{(L_2^2)^2}, \varepsilon^{5/2} \| \mathcal{R}_2 \|_{(H_2^2)^2} \ll 1$, uniformly on compact subsets of $\mathcal{B}$. \hfill \Box

For the estimates of $\frac{d}{dt} E_{2,l}$ with $1 \leq l \leq s$ we use

**Lemma 4.15.** Let $s \geq 0$ and $A : H^s(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R})$ be an antisymmetric linear operator.

a) Let $S_j : L^2(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R}), \ j \in \{1, 2\}$, be symmetric linear operators with $S_j H^s(\mathbb{R}, \mathbb{R}) \subseteq H^s(\mathbb{R}, \mathbb{R})$. Then for all $f \in H^s(\mathbb{R}, \mathbb{R})$ it holds that
\[ \int_{\mathbb{R}} S_1 S_2 f A f \, d\alpha = -\frac{1}{2} \int_{\mathbb{R}} f [A, S_1 S_2] f \, d\alpha - \frac{1}{2} \int_{\mathbb{R}} f A [S_1, S_2] f \, d\alpha. \]
b) Let $S, S_j : L^2(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R})$, $j \in \{\pm 1\}$, be symmetric linear operators with $SH^s(\mathbb{R}, \mathbb{R}), S_jH^s(\mathbb{R}, \mathbb{R}) \subseteq H^s(\mathbb{R}, \mathbb{R})$. Then for all $f \in H^s(\mathbb{R}, \mathbb{R})$ it holds that

$$
\sum_{j \in \{\pm 1\}} \int SS_j f_j Af_{-j} \, d\alpha
= \frac{1}{2} \int (f_{-1} - f_1) S(S_{-1} - S_1) A(f_{-1} + f_1) \, d\alpha
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int f_j [A, S_j S] f_{-j} \, d\alpha + \int f_{-j} A[S_j, S_j] f_j \, d\alpha \right)
- \sum_{j \in \{\pm 1\}} \frac{j}{4} \left( \int f_j [A, (S_{-1} - S_1)S] f_{j} \, d\alpha + \int f_{-j} A[S_{-1} - S_1, S] f_j \, d\alpha \right).
$$

(373)

c) Let $L_j : L^2(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R})$, $j \in \{\pm 1\}$, be linear operators with $L_jH^s(\mathbb{R}, \mathbb{R}) \subseteq H^s(\mathbb{R}, \mathbb{R})$ and $L_1^+ = L_{-1}$. Moreover, let $S : L^2(\mathbb{R}, \mathbb{R}) \to L^2(\mathbb{R}, \mathbb{R})$ be a symmetric linear operator with $SH^s(\mathbb{R}, \mathbb{R}) \subseteq H^s(\mathbb{R}, \mathbb{R})$. Then for all $f \in H^s(\mathbb{R}, \mathbb{R})$ it holds that

$$
\sum_{j \in \{\pm 1\}} \int SL_j f_j Af_{-j} \, d\alpha = - \int f_1 [A, SL_{-1}] f_{-1} \, d\alpha - \int f_{-1} A[S, L_1] f_1 \, d\alpha.
$$

(374)

Proof. We have

$$
\int S_1 S_2 f Af \, d\alpha = \int S_2 S_1 f Af \, d\alpha + \int [S_1, S_2] f Af \, d\alpha
= \int f S_1 S_2 Af \, d\alpha + \int [S_1, S_2] f Af \, d\alpha
= \int f A S_1 S_2 f \, d\alpha - \int f [A, S_1 S_2] f \, d\alpha + \int [S_1, S_2] f Af \, d\alpha
= - \int S_1 S_2 f Af \, d\alpha - \int f [A, S_1 S_2] f \, d\alpha - \int f A[S_1, S_2] f \, d\alpha,
$$

which implies (372), and

$$
\sum_{j \in \{\pm 1\}} \int SS_j f_j Af_{-j} \, d\alpha
= \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int S_j S f_j Af_{-j} \, d\alpha + \int SS_j f_j Af_{-j} \, d\alpha + \int [S, S_j] f_j Af_{-j} \, d\alpha \right)
= \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int S_j S f_j Af_{-j} \, d\alpha + \int f_j A S_j S f_{-j} \, d\alpha \right)
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int f_j [A, S_j S] f_{-j} \, d\alpha + \int f_{-j} A[S, S_j] f_j \, d\alpha \right)
= \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int S_j S f_j Af_{-j} \, d\alpha - \int S_j S f_{-j} Af_j \, d\alpha \right).
$$
Moreover, we deduce
\[
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int_R f_j [A, S_j] f_{-j} \, d\alpha + \int_R f_{-j} A[S, S_j] f_j \, d\alpha \right)
= \frac{1}{2} \left( \int_R (S_{-1} - S_1) S f_{-1} A f_1 \, d\alpha - \int_R (S_{-1} - S_1) S f_1 A f_{-1} \, d\alpha \right)
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int_R f_j [A, S_j] f_{-j} \, d\alpha + \int_R f_{-j} A[S, S_j] f_j \, d\alpha \right)
= \frac{1}{2} \left( \int_R (S_{-1} - S_1) S f_{-1} A(f_{-1} + f_1) \, d\alpha - \int_R (S_{-1} - S_1) S f_1 A(f_{-1} + f_1) \, d\alpha \right)
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int_R f_j [A, S_j] f_{-j} \, d\alpha + \int_R f_{-j} A[S, S_j] f_j \, d\alpha \right)
= \frac{1}{2} \int_R (f_{-1} - f_1) S(S_{-1} - S_1) A(f_{-1} + f_1) \, d\alpha
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int_R f_j [A, S_j] f_{-j} \, d\alpha + \int_R f_{-j} A[S, S_j] f_j \, d\alpha \right)
- \frac{1}{2} \sum_{j \in \{\pm 1\}} \left( \int_R f_j [A, S_j] f_{-j} \, d\alpha + \int_R f_{-j} A[S, S_j] f_j \, d\alpha \right)
- \sum_{j \in \{\pm 1\}} \frac{i}{4} \left( \int_R f_j [A, (S_{-1} - S_1)] f_{-j} \, d\alpha + \int_R f_{-j} A[S_{-1} - S_1, S] f_j \, d\alpha \right)
\]

Moreover, we deduce
\[
\sum_{j \in \{\pm 1\}} \int_R S L_j f_j A f_{-j} \, d\alpha = \int_R S L_{-1} f_{-1} A f_1 \, d\alpha + \int_R S L_1 f_1 A f_{-1} \, d\alpha
+ \int_R [S, L_1] f_1 A f_{-1} \, d\alpha
= \int_R S L_{-1} f_{-1} A f_1 \, d\alpha + \int_R f_1 A S L_{-1} f_{-1} \, d\alpha
- \int_R f_1 [A, SL_{-1}] f_{-1} \, d\alpha - \int_R f_{-1} A[S, L_1] f_1 \, d\alpha
= - \int_R f_1 [A, SL_{-1}] f_{-1} \, d\alpha - \int_R f_{-1} A[S, L_1] f_1 \, d\alpha.
\]

From now on, let $1 \leq l \leq s$. We compute
\[
\frac{d}{dt} E_{2,l} = \sum_{j_1 \in \{\pm 2\}} \left( \int_R \partial_{\rho_{j_1}} R_{j_1} \rho_{j_1} \partial_{\tau_{j_1}} \partial_{\tau_{j_1}} R_{j_1} \, d\alpha + \epsilon \sum_{j_2 \in \{\pm 2\}} \left( \int_R \partial_{\rho_{j_1}} \partial_{\rho_{j_1}} R_{j_1} \rho_{j_1} \partial_{\tau_{j_1}} N_{j_1, j_2} (\psi_{\epsilon}) R_{j_2} \, d\alpha \right.ight.
+ \int_R \partial_{\rho_{j_1}} R_{j_1} \rho_{j_1} \partial_{\tau_{j_1}} N_{j_1, j_2} (\psi_{\epsilon}) \partial_{\tau_{j_2}} \, d\alpha + \int_R \partial_{\rho_{j_1}} R_{j_1} \rho_{j_1} \partial_{\tau_{j_1}} N_{j_1, j_2} (\partial_{\tau_{j_2}} R_{j_2} \, d\alpha)) \right).
\]
Using (149) we obtain

\[
\frac{d}{dt} E_{2,l} = \sum_{j_1 \in \{\pm 2\}} \left( \text{sgn}(j_1) \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} i \omega \partial^l_{\alpha} R_{j_1} \, d\alpha \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \varepsilon^{-5/2} \partial^{-1}_{\alpha} \vartheta^{-1} \text{res}_{j_1}(\varepsilon \Psi) \, d\alpha \right) \\
+ \varepsilon \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \right) \\
+ \int_{\mathbb{R}} \text{sgn}(j_1) i \omega \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \\
+ \int_{\mathbb{R}} \text{sgn}(j_2) \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) i \omega R_{j_2} \, d\alpha \\
- \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2} (i \omega \psi_c) R_{j_2} \, d\alpha \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2} (\partial \psi_c + i \omega \psi_c) R_{j_2} \, d\alpha \\
- \int_{\mathbb{R}} \varepsilon^{-5/2} \partial^{l+1}_{\alpha} \vartheta^{-1} \text{res}_{j_1}(\varepsilon \Psi) \rho^l_{j_1} \partial^{-1}_{\alpha} N_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) \varepsilon^{-5/2} \vartheta^{-1} \text{res}_{j_2}(\varepsilon \Psi) \, d\alpha \right) \\
+ \varepsilon^2 \sum_{j_1, j_2, j_3 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} Q_{j_1 j_3}(\psi_c) R_{j_3} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \right) \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) Q_{j_2 j_3}(\psi_c) R_{j_3} \, d\alpha \\
+ \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} W_{j_1 j_2}(\Psi, \mathcal{R}) R_{j_2} \, d\alpha \right) \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} M_{j_1}(\Psi, \mathcal{R}) \, d\alpha \\
+ \varepsilon^3 \sum_{j_1, j_2, j_3 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} W_{j_1 j_3}(\Psi, \mathcal{R}) R_{j_3} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \right) \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) W_{j_2 j_3}(\Psi, \mathcal{R}) R_{j_3} \, d\alpha \\
+ \varepsilon^3 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} M_{j_1}(\Psi, \mathcal{R}) \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha \right) \\
+ \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho^l_{j_1} \partial^l_{\alpha} N_{j_1 j_2}(\psi_c) M_{j_2}(\Psi, \mathcal{R}) \, d\alpha \right) \cdot
\]
be bounded by $C \varepsilon^3 (E_x + 1)$ for a constant $C > 0$, as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^s)^2} \ll 1$, uniformly on compact subsets of $\mathcal{B}$. Hence, using (225), (284)–(285), (365)–(366) and (368), we obtain

$$\frac{d}{dt} E_{2,l} = \sum_{j=1}^{7} I_j + \varepsilon^2 \mathcal{O}(E_x + 1),$$

as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^s)^2} \ll 1$, uniformly on compact subsets of $\mathcal{B}$, where

\begin{align*}
I_1 &= -\varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} Y_{j_1 j_2} (\psi_c, R_{j_2}) \, d\alpha, \\
I_2 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} Q_{j_1 j_2} (\psi_c) R_{j_2} \rho_{j_1} \partial^l_{\alpha} N_{j_1 j_1} (\psi_c) R_{j_1} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} Q_{j_1 j_2} (\psi_c) R_{j_2} \, d\alpha \right), \\
I_3 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} Q_{j_1 j_2} (\psi_c) R_{j_2} \rho_{j_1} \partial^l_{\alpha} N_{j_1 j_1} (\psi_c) R_{j_1} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} Q_{j_1 j_2} (\psi_c) R_{j_2} \, d\alpha \right), \\
I_4 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \, d\alpha \right), \\
I_5 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \rho_{j_1} \partial^l_{\alpha} N_{j_1 j_1} (\psi_c) R_{j_1} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \, d\alpha \right), \\
I_6 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \rho_{j_1} \partial^l_{\alpha} N_{j_1 j_1} (\psi_c) R_{j_1} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} W_{j_1 j_2} (\Psi, \mathcal{R}) R_{j_2} \, d\alpha \right), \\
I_7 &= \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^l_{\alpha} \mathcal{M}_{j_1} (\Psi, \mathcal{R}) \rho_{j_1} \partial^l_{\alpha} N_{j_1 j_2} (\psi_c) R_{j_2} \, d\alpha \right. \\
&\quad + \left. \int_{\mathbb{R}} \partial^l_{\alpha} R_{j_1} \rho_{j_1} \partial^l_{\alpha} \psi_{j_1} \mathcal{M}_{j_2} (\Psi, \mathcal{R}) \, d\alpha \right).
\end{align*}

First, we analyze $I_2$. To extract all terms with more than $l$ spatial derivatives falling on $R_2$ or $R_{-2}$ we use Leibniz’s rule, integration by parts, (158)–(163), (255), (301),...
(340), (342)-(343) and (345) to obtain

\[
I_2 = \varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^j_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} \rho^j_{\alpha} \partial N_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha \right)
+ I \int_{\mathbb{R}} \partial^j_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} \rho^j_{\alpha} \partial N_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha
+ \int_{\mathbb{R}} \partial^j_{\alpha} R_{j_1} \rho^j_{\alpha} \partial N_{j_1, j_1}(\psi_c) \partial^j_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} \, d\alpha
+ I \int_{\mathbb{R}} \partial^j_{\alpha} R_{j_1} \rho^j_{\alpha} \partial N_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha
+ \varepsilon^2 O(E_s + 1)
+ 2\varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^j_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} (\rho^j_{\alpha} \mathfrak{N}^1_j)^r_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha \right)
- \int_{\mathbb{R}} \partial^{j-1}_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} (\rho^j_{\alpha} \mathfrak{N}^2_j)^a_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha
+ I \int_{\mathbb{R}} \partial^j_{\alpha} Q_{j_1 j_2}(\psi_c) R_{j_2} (\rho^j_{\alpha} \mathfrak{N}^1_j)^a_{j_1, j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha
+ \varepsilon^2 O(E_s + 1)
= \sum_{i=1}^4 2\varepsilon^2 \sum_{j_1, j_2 \in \{\pm 2\}} \int_{\mathbb{R}} Q_{j_1 j_2}(\psi_c) \partial^j_{\alpha} R_{j_2} S^j_{j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 O(E_s + 1)
=: \sum_{i=1}^4 I_2^i + \varepsilon^2 O(E_s + 1),
\]

as long as \( \varepsilon^{5/2}\|\mathcal{R}_1\|_{L^2}, \varepsilon^{5/2}\|\mathcal{R}_2\|_{H^s} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \), where

\[
S^j_{j_1}(\psi_c) = (\rho^j_{\alpha} \mathfrak{N}^1_j)^r_{j_1, j_1}(\psi_c) + I(\partial^{j-1}_{\alpha} \rho^j_{\alpha} \mathfrak{N}^1_j)^r_{j_1, j_1}(\partial^j_{\alpha} \psi_c).
\]

Due to (365)-(366), we have

\[
I_2^1 = -2\varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} \partial^j_{\alpha} (\psi_c \partial^j_{\alpha} R_{j_1}) S^j_{j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 O(E_s + 1)
= -2\varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} \partial^{j+1}_{\alpha} R_{j_1} \psi_c S^j_{j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 O(E_s + 1).
\]

Since \( \psi_c \) is real-valued and \( S^j_{j_1}(\psi_c) \) symmetric, using (302), (344) and (372) yields

\[
I_2^1 = \varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} \partial^j_{\alpha} R_{j_1} [\partial^j_{\alpha}, \psi_c S^j_{j_1}(\psi_c)] \partial^j_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 O(E_s + 1)
= \varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial^j_{\alpha} R_{j_1} \partial_s \psi_c S^j_{j_1}(\psi_c) \partial^j_{\alpha} R_{j_1} \, d\alpha \right)
\]
with a function \( f \) satisfying
\[
\|f(\Psi, \mathcal{R}_2)\|_{H^\epsilon} = \mathcal{O}((\mathcal{E}_s + 1)^{1/2}),
\]  
(377)
as long as \( \varepsilon^{5/2}\|\mathcal{R}_1\|_{(L^2)^2}, \varepsilon^{5/2}\|\mathcal{R}_2\|_{(H^\epsilon)^2} \ll 1 \). Hence, with the help of (101), (103), (148), (302), (323), (344), (365)–(366) and (372) we obtain
\[
I_2^2 = \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \partial_\alpha^j R_{j_1} \psi_c S_{j_1}^j (\partial_\alpha \partial_\alpha^j R_{j_1}) \, \mathrm{d}\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1) \]

\[
= \frac{\varepsilon^2}{4} \frac{d}{dt} \int_{\mathbb{R}} \partial_\alpha^j R_{j_2} (K_0 \sigma^{-1} \partial_\alpha \psi_c) (S_{j_2}^l + S_{l_2}^j) (\psi_c) (1 - b \partial_\alpha^2) \partial_\alpha^{j_1} (R_{j_2} - R_2) \, \mathrm{d}\alpha
\]

\[
\quad + \frac{\varepsilon^3}{2} \int_{\mathbb{R}} \partial_\alpha^j (R_{j_1} - R_2) (K_0 \sigma^{-1} \partial_\alpha \psi_c) (S_{j_1}^l + S_{l_2}^j) (\psi_c) (1 - b \partial_\alpha^2) \partial_\alpha^{l_1} (R_{j_1} - R_2) \, \mathrm{d}\alpha
\]

\[
\quad + \frac{\varepsilon^3}{4} \int_{\mathbb{R}} \partial_\alpha^j (R_{j_1} - R_2) (K_0 \sigma^{-1} \partial_\alpha \psi_c) (S_{j_1}^l + S_{l_2}^j) (\psi_c) (1 - b \partial_\alpha^2) \partial_\alpha^{l_1} (R_{j_1} - R_2) \, \mathrm{d}\alpha
\]

\[
\quad + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1) \]

(375)
as long as \( \varepsilon^{5/2}\|\mathcal{R}_1\|_{(L^2)^2}, \varepsilon^{5/2}\|\mathcal{R}_2\|_{(H^\epsilon)^2} \ll 1 \). Hence, with the help of (101), (103), (148), (302), (323), (344), (365)–(366) and (372) we obtain

\[
I_2^2 = \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \partial_\alpha^j R_{j_1} \psi_c S_{j_1}^j (\partial_\alpha \partial_\alpha^j R_{j_1}) \, \mathrm{d}\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\]

(375)
as long as \( \varepsilon^{5/2}\|\mathcal{R}_1\|_{(L^2)^2}, \varepsilon^{5/2}\|\mathcal{R}_2\|_{(H^\epsilon)^2} \ll 1 \). Hence, with the help of (101), (103), (148), (302), (323), (344), (365)–(366) and (372) we obtain

\[
I_2^2 = \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \partial_\alpha^j R_{j_1} \psi_c S_{j_1}^j (\partial_\alpha \partial_\alpha^j R_{j_1}) \, \mathrm{d}\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\]

(375)
as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^1)^2} \ll 1$, uniformly on compact subsets of $B$.

For $I_2^3$ and $I_2^4$ we can also use (101), (103), (148), (280), (302), (323), (344), (365)–(366) and (372)–(373) to deduce

$$I_2^3 = -\varepsilon^2 \sum_{j_l \in \{\pm 1\}} \int_{\mathbb{R}} b(\sigma^{-1} \partial^2_{\langle \alpha^{-1} (R_{-2} - R_2) S^l_{j_l}(\psi_c) \rangle} R_{j_l} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

and

$$I_2^4 = \varepsilon^2 \sum_{j_l \in \{\pm 1\}} \int_{\mathbb{R}} \text{sgn}(j_l) ([\sigma, \partial_{\langle \alpha^{-1} (R_{-2} + R_2) \rangle} S^l_{j_l}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_l} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^1)^2} \ll 1$, uniformly on compact subsets of $B$.

Next, we examine $I_3$. Using Leibniz’s rule, integration by parts, (158)–(163), (256), (301), (303)–(304) and (365)–(366) we conclude that

$$I_3 = \varepsilon^2 \sum_{j_l, j_2 \in \{\pm 1\}} \left( \int_{\mathbb{R}} \partial^l_{\langle \alpha^{-1} \rangle} Q_{j_1 j_2}(\psi_c) R_{j_2} \rho_{j_1} \partial N^l_{j_1, j_2} (\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

$$= 4 \varepsilon^2 \sum_{i=1}^{4} \sum_{j_l, j_2 \in \{\pm 1\}} \int_{\mathbb{R}} Q^l_{j_1 j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} (\rho^l \Omega^l)^{j_2}_{j_1, j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

$$= 2 \varepsilon^2 \sum_{j_l, j_2 \in \{\pm 1\}} \int_{\mathbb{R}} Q^l_{j_1 j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} (\rho^l \Omega^l)^{j_2}_{j_1, j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

$$= 4 \varepsilon^2 \sum_{i=1}^{4} \sum_{j_l, j_2 \in \{\pm 1\}} \int_{\mathbb{R}} Q^l_{j_1 j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} (\rho^l \Omega^l)^{j_2}_{j_1, j_2}(\psi_c) \partial^l_{\langle \alpha^{-1} \rangle} R_{j_2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1)$$

$$= 4 \varepsilon^2 \sum_{i=1}^{4} I^3_i + \varepsilon^2 \mathcal{O}(\varepsilon_s + 1),$$
as long as \( \varepsilon^{5/2}\|R_1\|_{(L^2)^2}, \varepsilon^{5/2}\|R_2\|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Because of \((302), (365)–(366)\) and \((374)\) we obtain

\[
I_3^1 = -2\varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} \partial^{j_1+1}_\alpha R_{-j_1} \psi_c (\rho^{j_1} M_{-j_1,j_1}(\psi_c) \partial^j\psi \partial^j_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= 2\varepsilon^2 \int_{\mathbb{R}} \partial^{j_1}_\alpha R_2 \partial^{j_1}_{\alpha} \psi_c (\rho^{j_1} M_{-j_1,j_1}(\psi_c) \partial^j_{\alpha} R_{-2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= 2\varepsilon^2 \int_{\mathbb{R}} \partial^{j_1}_\alpha R_2 \partial^{j_1}_{\alpha} \psi_c (\rho^{j_1} M_{-j_1,j_1}(\psi_c) \partial^j_{\alpha} R_{-2} \, d\alpha
\]

\[
+ 2\varepsilon^2 \int_{\mathbb{R}} \partial^{j_1}_\alpha R_2 \psi_c (\rho^{j_1} M_{-j_1,j_1}(\psi_c) \partial^j_{\alpha} R_{-2} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

as long as \( \varepsilon^{5/2}\|R_1\|_{(L^2)^2}, \varepsilon^{5/2}\|R_2\|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Due to \((305), (365)–(366), (372)–(373)\) and \((375)–(377)\), we have

\[
I_3^2 = \varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} (\left\|K_0\sigma^{-1}_\alpha \partial^{j_1+1}_\alpha \right\|(\left\|R_{-2} - R_2\right\| G_{j_1}(\psi_c) \partial^j_{\alpha} \right\| R_{j_1} \, d\alpha + \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \partial^{j_1}_\alpha (R_{-2} - R_2)(\left\|K_0\sigma^{-1}_\alpha \partial^{j_1+1}_\alpha \right\|(\left\|G_{j_1}^l + G_{j_2}^l\right\| \psi_c \sigma^{-1}_\alpha \partial^{j_1+1}_\alpha (R_{-2} + R_2) \, d\alpha
\]

\[
+ \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= \frac{\varepsilon^2}{4} \frac{d}{dt} \int_{\mathbb{R}} \partial^{j_1}_\alpha (R_{-2} - R_2)(\left\|K_0\sigma^{-1}_\alpha \partial^{j_1+1}_\alpha \right\|(\left\|G_{j_1}^l + G_{j_2}^l\right\| \psi_c (1 - b \partial^{j_1+2}_\alpha \sigma^{-1}_\alpha \partial^{j_1+1}_\alpha (R_{-2} - R_2) \, d\alpha
\]

\[
+ \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

as long as \( \varepsilon^{5/2}\|R_1\|_{(L^2)^2}, \varepsilon^{5/2}\|R_2\|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

For \( I_3^3 \) and \( I_3^4 \) we can use \((101), (103), (148), (191), (274), (305), (324), (365)–(366) \) and \((372)–(373)\) again to deduce

\[
I_3^3 = -\varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} b(\sigma^{-1}_\alpha \partial^{j_1+2}_\alpha \sigma^{-1}_\alpha \partial^{j_1+2}_\alpha (R_{-2} - R_2) \sigma^{-1}_\alpha \partial^{j_1+2}_\alpha \right\| R_{j_1} \, d\alpha
\]

\[
+ \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \partial^{j_1}_\alpha (R_{-2} - R_2)(\sigma^{-1}_\alpha \partial^{j_1+2}_\alpha \psi_c) (G_{j_1}^l + G_{j_2}^l) \psi_c (1 - b \partial^{j_1+2}_\alpha \sigma^{-1}_\alpha \partial^{j_1+2}_\alpha (R_{-2} - R_2) \, d\alpha
\]

\[
+ \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

\[
= \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]

and

\[
I_3^4 = \varepsilon^2 \sum_{j_1 \in \{\pm 2\}} \int_{\mathbb{R}} \operatorname{sgn}(j_1)(\sigma^{-1}_\alpha \partial^{j_1-1}_\alpha (R_{-2} + R_2) \sigma^{-1}_\alpha \partial^{j_1-1}_\alpha \psi_c) G_{j_1}^l \psi_c \partial^j_{\alpha} \right\| R_{j_1} \, d\alpha
\]

\[
+ \varepsilon^2 \mathcal{O}(\varepsilon_x + 1)
\]
as long as \( \varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{L^2}^2 \ll 1 \), uniformly on compact subsets of \( B \).

Now, we investigate \( I_4 \). Because of (142) we have

\[
I_4 = \sum_{j_1, j_2 \in \{-1, 1\}} \varepsilon^2 \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) \sigma^{-1} \partial_{\alpha}^2 \Psi_{\alpha}^l (G_{-2}^l + G_2^l) \sigma_{\alpha} \partial_{\alpha}^{l-1} (R_{-2} + R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

\[
= 4 \sum_{i=1}^{4} I_4^i + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\]

as long as \( \varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{L^2}^2 \ll 1 \), uniformly on compact subsets of \( B \). Analogously to the case of \( I_2 \), we conclude that

\[
I_4^1 = -\varepsilon^2 \sum_{j_1 \in \{-1, 1\}} \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) \sigma^{-1} \partial_{\alpha}^2 g_{\alpha}^l (\Psi_{\alpha}^l, R_2) \sigma_{\alpha} \partial_{\alpha}^{l+1} R_{j_1} \sigma_{\alpha} \partial_{\alpha} R_{j_1} \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

\[
= \frac{\varepsilon^2}{2} \sum_{j_1 \in \{-1, 1\}} \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) \sigma^{-1} \partial_{\alpha}^2 g_{\alpha}^l (\Psi_{\alpha}^l, R_2) (\partial_{\alpha} R_{j_1})^2 \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

\[
= \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

and

\[
I_4^2 = -\varepsilon^2 \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) (K_0 \sigma^{-1} \partial_{\alpha}^2 g_{\alpha}^l (\Psi_{\alpha}^l, R_2) + c(\Psi, \mathcal{R}))
\]

\[
\times \sigma^{-1} \partial_{\alpha}^{l+1} (R_{-2} + R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

\[
= \frac{\varepsilon^2}{4} \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) (K_0 \sigma^{-1} \partial_{\alpha}^2 g_{\alpha}^l (\Psi_{\alpha}^l, R_2) + c(\Psi, \mathcal{R}))
\]

\[
\times (1 - b \partial_{\alpha}^2)^{-1} \partial_{\alpha}^l (R_{-2} - R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\]

as long as \( \varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{L^2}^2 \ll 1 \), uniformly on compact subsets of \( B \). Furthermore, we deduce

\[
I_4^3 = \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) (\sigma^{-1} g_{\alpha}^l (\Psi_{\alpha}^l, R_2)) b K_0 \sigma^{-1} \partial_{\alpha}^{l+2} (R_{-2} + R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]

\[
= -\frac{\varepsilon^2}{4} \int_{\mathbb{R}} \delta^l_{\alpha} (R_{-2} - R_2) (\sigma^{-1} g_{\alpha}^l (\Psi_{\alpha}^l, R_2)) b K_0 \partial_{\alpha}
\]

\[
\times (1 - b \partial_{\alpha}^2)^{-1} \partial_{\alpha}^l (R_{-2} - R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1)
\]
as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^2)^2} \ll 1$, uniformly on compact subsets of $B$. Moreover, due to (101), (103), (148), we have

$$I_4^4 = -\frac{\varepsilon^2}{2} \int_{\mathbb{R}} \sigma \partial_{\alpha}^{-1} (R_{-2} + R_2) (\sigma^{-1} g_+ (\Psi_2^h, R_2)) \partial_{\alpha} (R_{-2} - R_2) \, d\alpha$$

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

and because of

$$\sigma \partial_{\alpha}^{-1} = (-K_0 \partial_{\alpha})^{-1} i \omega$$

and (376)–(377) we obtain

$$I_4^4 = \frac{\varepsilon^2}{4} \frac{d}{dt} \int_{\mathbb{R}} \partial_{\alpha}^j (R_{-2} - R_2) \sigma^{-1} g_+ (\Psi_2^h, R_2) (-K_0 \partial_{\alpha})^{-1} \partial_{\alpha}^j (R_{-2} - R_2) \, d\alpha$$

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

as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^2)^2} \ll 1$, uniformly on compact subsets of $B$.

Next, we consider $I_5$ and $I_6$. Analogously to the cases of $I_2$ and $I_3$ we obtain

$$I_5 = \sum_{j_1, j_2 \in \{\pm 2\}, i \in \{1, \ldots, 4\}} 2\varepsilon^3 \int_{\mathbb{R}} W_{j_1, j_2} (\Psi, R) \partial_{\alpha}^j R_{j_2} S_{j_1} (\psi_c) \partial_{\alpha}^j R_{j_1} \, d\alpha$$

$$+ \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1)$$

$$=: \sum_{i=1}^4 I_5^i + \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1),$$

$$I_6 = \sum_{j_1, j_2 \in \{\pm 2\}, i \in \{1, \ldots, 4\}} 2\varepsilon^3 \int_{\mathbb{R}} W_{j_1, j_2} (\Psi, R) \partial_{\alpha}^j R_{j_2} (\rho \gamma_1')_{j_1 - j_1} (\psi_c) \partial_{\alpha}^j R_{-j_1} \, d\alpha$$

$$+ \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1)$$

$$= \sum_{i=1}^4 I_6^i + \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1)$$

as long as $\varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^2)^2} \ll 1$, uniformly on compact subsets of $B$ and therefore

$$I_5^1 = \varepsilon^3 \sum_{j_1 \in \{\pm 2\}} \left( \int_{\mathbb{R}} \partial_{\alpha}^j R_{j_1} \partial_{\alpha}^{-1} g_+ (\Psi_2^h, R_2) S_{j_1} (\psi_c) \partial_{\alpha}^j R_{j_1} \, d\alpha \right)$$

$$+ \int_{\mathbb{R}} \partial_{\alpha}^j R_{j_1} \partial_{\alpha}^{-2} g_+ (\Psi_2^h, R_2) S_{j_1} (\partial_{\alpha} \psi_c) \partial_{\alpha}^j R_{j_1} \, d\alpha$$

$$+ \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1)$$

$$= \varepsilon^3 \mathcal{O}(\mathcal{E}_s + 1),$$

$$I_6^1 = 2\varepsilon^3 \int_{\mathbb{R}} \partial_{\alpha}^j R_2 \partial_{\alpha}^{-1} g_+ (\Psi_2^h, R_2) (\rho \gamma_1')_{j_2 - j_2} (\psi_c) \partial_{\alpha}^j R_{-j_2} \, d\alpha$$
\[+2e^3 \int R_2 \partial_{\alpha}^{-2} g_{+}(\Psi^h_j, \mathcal{R}_2)(\rho^j \eta_1)_{j_{i=2}}(\partial_{\alpha} \psi c) \partial_{\alpha} R_{-2} \, d\alpha \]
\[= e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_5^2 = \frac{e^3}{4} \frac{d}{d\alpha} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (S_{-2}^2 + S_{-2}^2)(\psi c) \right) \] 
\[\times (1 - b \partial_{\alpha}^{-2})^2 \partial_{\alpha}^2 (R_2 - R_2) \, d\alpha + e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_6^2 = \frac{e^3}{4} \frac{d}{d\alpha} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (G_{-2} + G_{-2}^2)(\psi c) \right) \] 
\[\times (1 - b \partial_{\alpha}^{-2})^2 \partial_{\alpha}^2 (R_2 - R_2) \, d\alpha + e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_5^3 = \frac{e^3}{2} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (S_{-2}^2 + S_{-2}^2)(\psi c) b K_{0} \partial_{\alpha}^{-1} \partial_{\alpha}^2 (R_2 + R_2) \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[= e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_6^3 = \frac{e^3}{2} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (G_{-2} + G_{-2}^2)(\psi c) b K_{0} \partial_{\alpha}^{-1} \partial_{\alpha}^2 (R_2 + R_2) \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_5^4 = \frac{e^3}{2} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (S_{-2}^2 + S_{-2}^2)(\psi c) \sigma \partial_{\alpha}^{-1} (R_2 + R_2) \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[= e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[I_6^4 = - \frac{e^3}{2} \int R_2 \partial_{\alpha}^{-1} g_{+}(\Psi^h_j, \mathcal{R}_2) + c(\Psi, \mathcal{R}) \left( (G_{-2} + G_{-2}^2)(\psi c) \sigma \partial_{\alpha}^{-1} (R_2 + R_2) \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[= e^3 \mathcal{O}(E_\varepsilon + 1),\]

as long as \( e^{5/2} \| \mathcal{R}_1 \| (L^2)^2, \ e^{5/2} \| \mathcal{R}_2 \| (H^2)^2 \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Next, we estimate \( I_7 \). Due to (142), (256) and (301), we deduce
\[I_7 = e^3 \sum_{j_1 \in \{ \pm 2 \}} \left( \int R_j \partial_{\alpha}^j \mathcal{M}_{j_1}(\Psi, \mathcal{R}) \partial_{\alpha}^j \rho^j \partial_{\alpha}^j \partial_{\alpha}^j \partial_{\alpha}^j \psi c \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[\sum_{j_1 \in \{ \pm 2 \}} \int R_j \partial_{\alpha}^j \mathcal{M}_{j_1}(\Psi, \mathcal{R}) \left( (\rho^j \eta_1)_{j_{i=2}}(\psi c) \partial_{\alpha}^j R_j \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[= e^3 \sum_{j_1 \in \{ \pm 2 \}} \int R_j \partial_{\alpha}^j \partial_{\alpha}^j \partial_{\alpha}^j \partial_{\alpha}^j \psi c \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[\sum_{j_1 \in \{ \pm 2 \}} \int R_j \partial_{\alpha}^j \mathcal{M}_{j_1}(\Psi, \mathcal{R}) \left( (\rho^j \eta_1)_{j_{i=2}}(\psi c) \partial_{\alpha}^j R_j \right) \] 
\[+ e^3 \mathcal{O}(E_\varepsilon + 1),\]
\[= e^3 \mathcal{O}(E_\varepsilon + 1),\]
as long as \( e^{5/2} \| \mathcal{R}_1 \| (L^2)^2, \ e^{5/2} \| \mathcal{R}_2 \| (H^2)^2 \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).
Finally, we bound $I_1$ by improving the estimate (286). It follows from the proof of this estimate that

\[ I_1 = \sum_{\nu=1}^{2} I_{1,1}^{\nu} + \sum_{\nu=1}^{2} I_{1,2}^{\nu} + \sum_{\nu=2}^{4} I_{1,3}^{\nu} + I_{1,4} + \varepsilon^2 O(E_s + 1) \]

with

\[ I_{1,1}^{\nu} = -\varepsilon \sum_{j_1, j_2 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) \left( \hat{r}_{j_1, j_2} (k, k - m, m) - \hat{r}_{j_1, j_2} (k, \ell k_0, m) \right) \]

\[ \times (ik)^l (\hat{\rho}^{\nu})_{j_1, j_2} (k) \hat{\psi}_\ell (k - m) (im)^{-\nu - 1} \hat{R}_{j_2} (m) \, dk \, d\nu, \]

\[ I_{1,2}^{\nu} = -\varepsilon \sum_{j_1, j_2 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) \left( \hat{r}_{j_1, j_2} (k, \ell k_0, m) - \hat{r}_{j_1, j_2} (k, \ell k_0, k - k_0) \right) \]

\[ \times (ik)^l (\hat{\rho}^{\nu})_{j_1, j_2} (k) \hat{\psi}_\ell (k - m) (im)^{-\nu - 1} \hat{R}_{j_2} (m) \, dk \, d\nu, \]

\[ I_{1,3}^{\nu} = \varepsilon \sum_{j_1, j_2 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) \left( \hat{q}^{\nu}_{j_1, j_2} (k, k - m, m) - \hat{q}^{\nu}_{j_1, j_2} (k, \ell k_0, m) \right) \]

\[ \times (ik)^l \hat{\psi}_\ell (k - m) \hat{\vartheta} (m) \hat{R}_{j_2} (m) \, dk \, d\nu, \]

\[ I_{1,4} = \varepsilon \sum_{j_1, j_2 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) \left( \hat{q}^{\nu}_{j_1, j_2} (k, \ell k_0, m) - \hat{q}^{\nu}_{j_1, j_2} (k, \ell k_0, k - k_0) \right) \]

\[ \times (ik)^l \hat{\psi}_\ell (k - m) \hat{\vartheta} (m) \hat{R}_{j_2} (m) \, dk \, d\nu. \]

Because of

\[ \hat{r}_{j_1, j_2} (k, k - m, m) - \hat{r}_{j_1, j_2} (k, \ell k_0, m) = i (\omega (k - m) - \omega (\ell k_0)) \]

as well as (280)–(281) and Lemma 4.2 it holds that

\[ I_{1,1}^{\nu} = -\varepsilon \sum_{j_1 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) i (\omega (k - m) - \omega (\ell k_0)) \]

\[ \times (\hat{\rho}^{\nu})_{j_1, j_2} (k) \hat{\psi}_\ell (k - m) (im)^{\nu} \hat{R}_{j_2} (m) \, dk \, d\nu + \varepsilon^2 O(E_s + 1). \]

For symmetry reasons it follows that

\[ I_{1,1}^{\nu} = -\varepsilon \sum_{j_1 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_1} (k) i (\omega (k - m) - \omega (\ell k_0)) \]

\[ \times (\hat{\rho}^{\nu})_{j_1, j_2} (k) \hat{\psi}_\ell (k - m) (im)^{\nu} \hat{R}_{j_2} (m) \, dk \, d\nu + \varepsilon^2 O(E_s + 1). \]
such that, due to (323) and Lemma 4.2, we conclude
\[
I_{1,1}^1 = \varepsilon^2 \mathcal{O}(\mathcal{E}_\varepsilon + 1),
\]
as long as \( \varepsilon^{5/2} \| \mathcal{R}_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| \mathcal{R}_2 \|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Furthermore, by using (148), (282)–(283), (379) and Lemma 4.2 we obtain
\[
I_{1,1}^2 = \varepsilon^2 \mathcal{O}(\mathcal{E}_\varepsilon + 1),
\]
as long as \( \varepsilon^{5/2} \| \mathcal{R}_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| \mathcal{R}_2 \|_{(H^s)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \).

Also, because of (280)–(281) and Lemma 4.2, we deduce that
\[
I_{1,2}^1 = -\varepsilon \sum_{j_1 \in \{\pm 2\}, \ell \in \{\pm 1\}} \int_{\mathbb{R}} \int_{\mathbb{R}} \left( ik \right) \mathcal{R}_{j_1}(k) \left( \hat{\mathcal{R}}_{j_1,j_1}(k, \ell k_0, m) - \hat{\mathcal{R}}_{j_1,j_1}(k, \ell k_0, k - \ell k_0) \right)
\times \left( \rho^1 \mathbf{n}^1 \right)_{j_1,j_1 \ell}(k) \hat{\psi}_{\ell}(k - m) \left( im \right) \mathcal{R}_{j_1}(m) \, dmdk + \varepsilon^2 \mathcal{O}(\mathcal{E}_\varepsilon + 1).
\]

We split the integral kernel into
\[
\left( \hat{\mathcal{R}}_{j_1,j_1}(k, \ell k_0, m) - \hat{\mathcal{R}}_{j_1,j_1}(k, \ell k_0, k - \ell k_0) \right) \left( \rho^1 \mathbf{n}^1 \right)_{j_1,j_1 \ell}(k)
= \left( \hat{\mathcal{R}}_{j_1 \ell}(k, m) + \hat{\mathcal{R}}_{j_1 \ell}(k, m) \right) \left( \rho^1 \mathbf{n}^1 \right)_{j_1,j_1 \ell}(k)
\]
with
\[
\hat{\mathcal{R}}_{j_1 \ell}(k, m) = -\frac{1}{2} \text{sgn}(j_1) i \left( (\omega(m) - \omega(k - \ell k_0)) + (\omega(-k) - \omega(-m - \ell k_0)) \right),
\]
\[
\hat{\mathcal{R}}_{j_1 \ell}(k, m) = -\frac{1}{2} \text{sgn}(j_1) i \left( (\omega(m) - \omega(k - \ell k_0)) - (\omega(-k) - \omega(-m - \ell k_0)) \right).
\]

By the mean value theorem we have
\[
(\omega(m) - \omega(k - \ell k_0)) \pm (\omega(-k) - \omega(-m - \ell k_0))
= (\omega(m) - \omega(k - \ell k_0)) \mp (\omega(m) - \omega(m + \ell k_0))
= -\omega'(m + \theta_0(k, m, \ell)(k - m - \ell k_0)) (k - m - \ell k_0)
\mp \omega'(k - \theta_1(k, m, \ell)(k - m - \ell k_0)) (k - m - \ell k_0)
\]
\begin{equation}
(380)
\end{equation}

with \( \theta_0(k, m, \ell), \theta_1(k, m, \ell) \in [0, 1] \). Hence, we obtain
\[
\hat{\mathcal{R}}_{j_1 \ell}(k, m) \chi_{\ell}(k - m) = \mathcal{O}(|k|^{-1/2}(1 + bk^2)^{1/2}) (k - m - \ell k_0)
\]
\begin{equation}
(381)
\end{equation}
for \( |k| \to \infty \) uniformly with respect to \( m \in \mathbb{R} \) and \( b \lesssim 1 \) and, by using the mean value theorem once more,
\[
\hat{\mathcal{R}}_{j_1 \ell}(k, m) \chi_{\ell}(k - m) = \mathcal{O}(|k|^{-3/2}(1 + bk^2)^{1/2}) (k - m - \ell k_0)
\]
\begin{equation}
(382)
\end{equation}
for $|k| \to \infty$ uniformly with respect to $m \in \mathbb{R}$ and $b \lesssim 1$. Consequently, due to (280), (323) and Lemma 4.2, we conclude

\begin{align*}
I_{1,2}^1 &= -\frac{\varepsilon}{2} \sum_{\substack{j_l \in \{\pm 2\}, \xi_l \in \{\pm 1\}}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_l}(k) \hat{\gamma}_{j_l}(k, m) \left(\hat{\rho}^{a}_l(n^{a})_{j_l}(k) \hat{\psi}_{j_l}(k-m)\right) \times (im)^l \hat{R}_{j_l}(m) \, dmdk \\
&- \frac{\varepsilon}{2} \sum_{\substack{j_l \in \{\pm 2\}, \xi_l \in \{\pm 1\}}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_l}(k) \hat{\gamma}_{j_l}(k, m) \left(\hat{\rho}^{a}_l(n^{a})_{j_l}(k) \hat{\psi}_{j_l}(k-m)\right) \times (im)^l \hat{R}_{j_l}(m) \, dmdk \\
&+ \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1) \\
&= \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\end{align*}

as long as $\varepsilon^{5/2}\|R_1\|_{(L^2)^2}, \varepsilon^{5/2}\|R_2\|_{(H^1)^2} \ll 1$, uniformly on compact subsets of $B$.

Furthermore, because of (148), (345), (381)–(382) and Lemma 4.2, we have

\begin{align*}
I_{1,2}^2 &= -\frac{\varepsilon}{2} \sum_{\substack{j_l \in \{\pm 2\}, \xi_l \in \{\pm 1\}}} \int_{\mathbb{R}} \int_{\mathbb{R}} (ik)^l \hat{R}_{j_l}(k) \hat{\gamma}_{j_l}(k, m) \left(\hat{\rho}^{a}_l(n^{a})_{j_l}(k) \hat{\psi}_{j_l}(k-m)\right) \times (im)^l \hat{R}_{j_l}(m) \, dmdk \\
&= \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1).
\end{align*}

Hence, using

\begin{align*}
(\hat{\rho}^{a}_l(n^{a})_{j_l}(k, m)) &= \frac{1}{2} \left((ik)^l (\hat{\rho}^{a}_l(n^{a})_{j_l}(k)) - (i(-m))^l (\hat{\rho}^{a}_l(n^{a})_{j_l}(-m))\right)im \\
&\quad + \frac{1}{2} (ik)^l (\hat{\rho}^{a}_l(n^{a})_{j_l}(k)) i(k-m)
\end{align*}

as well as (112), (118), (282), (381) and Young’s inequality for convolutions, we obtain

\begin{align*}
I_{1,2}^2 &= \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\end{align*}

as long as $\varepsilon^{5/2}\|R_1\|_{(L^2)^2}, \varepsilon^{5/2}\|R_2\|_{(H^1)^2} \ll 1$, uniformly on compact subsets of $B$.

Analogously to the case of $I_{3,4}^2$ and $I_{4,4}^3$, we deduce that

\begin{align*}
I_{1,3}^2 &= \frac{\varepsilon^2}{4} \frac{d}{d\alpha} \int_{\mathbb{R}} \partial_{\alpha}^l (R_2 - R_2) L_1(\psi_{\alpha}) (1 - b \hat{\rho}^{2}_a)^{-1} \partial_{\alpha}^l (R_2 - R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1), \\
I_{1,3}^3 &= -\frac{\varepsilon^2}{4} \frac{d}{d\alpha} \int_{\mathbb{R}} \partial_{\alpha}^l (R_2 - R_2) L_2(\psi_{\alpha}) bK_0 \partial_{\alpha}(1 - b \hat{\rho}^{2}_a)^{-1} \partial_{\alpha}^l (R_2 - R_2) \, d\alpha \\
&\quad + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),
\end{align*}
as long as $\varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{H^2}^2 \ll 1$, uniformly on compact subsets of $B$, where

$$\hat{L}_1(\psi_c)(k) = \sum_{\ell \in \{\pm 1\}} \varepsilon^{-1}(\hat{K}_0(k) \sigma^{-1}(k) i k - \hat{K}_0(\ell k_0) \sigma^{-1}(\ell k_0) i \ell k_0) \hat{\psi}_\ell(k),$$

$$\hat{L}_2(\psi_c)(k) = - \sum_{\ell \in \{\pm 1\}} \varepsilon^{-1}(\sigma^{-1}(k) k^2 - \sigma^{-1}(\ell k_0)(\ell k_0)^2) \hat{\psi}_\ell(k).$$

To bound $I_{1,3}^4$ we split the integral kernel into

$$\tilde{q}^{2,y}_{j_1j_2}(k, k - m, m) - \tilde{q}^{2,y}_{j_1j_2}(k, \ell k_0, m)$$

$$= \frac{1}{2} \text{sgn}(j_1) i k i \frac{\sigma(k) - \sigma(k - m)}{k - (k - m)} (\sigma^{-1}(k - m)(k - m)^2 - \sigma^{-1}(\ell k_0)(\ell k_0)^2) (im)^{-1}$$

$$+ \frac{1}{2} \text{sgn}(j_1) i k i \left( \frac{\sigma(k) - \sigma(k - m)}{k - (k - m)} - \frac{\sigma(k) - \sigma(\ell k_0)}{k - \ell k_0} \right) \sigma^{-1}(\ell k_0)(\ell k_0)^2 (im)^{-1}.$$

With the help of Taylor’s theorem, we obtain

$$\frac{\sigma(k) - \sigma(k - m)}{k - (k - m)} - \frac{\sigma(k) - \sigma(\ell k_0)}{k - \ell k_0} \leq \frac{1}{2} \|\hat{q}_k^2\|_{L^\infty} (k - m - \ell k_0)$$

$$= \mathcal{O}(1) (k - m - \ell k_0),$$

uniformly with respect to $k, m \in \mathbb{R}$ and $b \lesssim 1$. Because of Lemma 4.2 we conclude that

$$I_{1,3}^4 = - \frac{\varepsilon^2}{2} \int_\mathbb{R} ((\sigma, \partial^{l-1}_\alpha (R_{-2} + R_2)) L_2(\psi_c)) \partial^{l}_\alpha (R_{-2} - R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1).$$

Hence, we can proceed analogously to the case of $I_{4}^4$ to deduce

$$I_{1,3}^4 = - \frac{\varepsilon^2}{4} \frac{d}{dt} \int_\mathbb{R} \partial^{l}_\alpha (R_{-2} - R_2) L_2(\psi_c)(-K_0 \partial^{l-1}_\alpha (R_{-2} - R_2) \, d\alpha + \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),$$

as long as $\varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{H^2}^2 \ll 1$, uniformly on compact subsets of $B$.

Finally, due to the mean value theorem, we have

$$(\tilde{q}^{1}_{j_1j_2}(k, \ell k_0, m) - \tilde{q}^{1}_{j_1j_2}(k, \ell k_0, k - \ell k_0)) \chi(k - m) = \mathcal{O}(|k|^{-1/2}) (k - m - \ell k_0)$$

for $|k| \to \infty$ uniformly with respect to $m \in \mathbb{R}$ and $b \lesssim 1$, such that with the help of Lemma 4.2 we obtain

$$I_{1,4}^4 = \varepsilon^2 \mathcal{O}(\mathcal{E}_s + 1),$$

as long as $\varepsilon^{5/2}\|R_1\|_{L^2}^2, \varepsilon^{5/2}\|R_2\|_{H^2}^2 \ll 1$, uniformly on compact subsets of $B$.

Now, we define our final energy $\tilde{\mathcal{E}}_s$ by

$$\tilde{\mathcal{E}}_s := \mathcal{E}_s + \frac{1}{4} \varepsilon^2 \sum_{i=1}^7 \sum_{l=1}^s h_i^l,$$

(383)
with

\[ h_1^1 = -\int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) L_1(\psi_c) (1 - b\partial_\alpha^2)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^2 = \int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) L_2(\psi_c) b K_0 \partial_\alpha (1 - b\partial_\alpha^2)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^3 = -\int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) L_2(\psi_c) (-K_0 \partial_\alpha)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^4 = -\int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) (K_0 \sigma^{-1} \partial_\alpha \psi_c) (S_{-2}^l + S_2^l + G_{-2}^l + G_2^l) (\psi_c) \times (1 - b\partial_\alpha^2)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^5 = -\int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) (K_0 \sigma^{-1} \partial_\alpha^{-1} g_{-}(\Psi^h_2, \mathcal{R}_2) + c(\Psi, \mathcal{R})) \times (1 + \varepsilon(S_{-2}^l + S_2^l + G_{-2}^l + G_2^l) (\psi_c)) (1 - b\partial_\alpha^2)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^6 = \int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) (\sigma^{-1} g_{-}(\Psi^h_2, \mathcal{R}_2)) b K_0 \partial_\alpha (1 - b\partial_\alpha^2)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha, \]

\[ h_1^7 = -\int_{\mathbb{R}} \partial_\alpha^t (R_{-2} - R_2) \sigma^{-1} g_{-}(\Psi^h_2, \mathcal{R}_2) (-K_0 \partial_\alpha)^{-1} \partial_\alpha^t (R_{-2} - R_2) \, d\alpha. \]

Then, for sufficiently small \( \varepsilon > 0 \), the energy \( \tilde{E}_\alpha \) satisfies the estimates (250)–(252), as long as \( \varepsilon^{5/2} \| R_1 \|_{(L^2)^2}, \varepsilon^{5/2} \| R_2 \|_{(H^1)^2} \ll 1 \), uniformly on compact subsets of \( \mathcal{B} \). Hence, we have proven Lemma 4.3.

**Proofs of Theorems 4.1 and 1.1.** If \( \| R_1 \|_{t=0} \|_{(L^2)^2}, \mathcal{R}_2 \|_{t=0} \|_{(H^1)^2} \lesssim 1 \), then Lemma 4.3 allows us to use Gronwall’s inequality to obtain for sufficiently small \( \varepsilon > 0 \) the \( \mathcal{O}(1) \)-boundedness of \( \tilde{E}_\alpha \) for all \( t \in [0, \tau_0/\varepsilon^2] \) uniformly on compact subsets of \( \mathcal{B} \). Due to (122) and (252), Theorem 4.1 follows. Transferring the assertions of Theorem 4.1 into Eulerian coordinates finally yields Theorem 1.1.

**Acknowledgements.** This work was supported by the Deutsche Forschungsgemeinschaft DFG under the grant DU 1198/2. The author thanks Max Heß for discussions and the referee for useful comments.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**References**

1. Ablowitz, M.J., Segur, H.: Solitons and the inverse scattering transform. In: SIAM Studies in Applied Mathematics, vol. 4, SIAM, 1981
2. Alazard, T., Delort, J.-M.: Global solutions and asymptotic behavior for two dimensional gravity water waves. *Ann. Sci. Éc. Norm. Supér. (4)* **48**(5), 1149–1238, 2015
3. Ambrose, D.M., Masmoudi, N.: The zero surface tension limit of two-dimensional oscillatory long waves for equations with quasilinear quadratic terms. *Commun. Pure Appl. Math.* **58**(10), 1287–1315, 2005
4. 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
5. Craig, W., Sulem, C., Sulem, P.L.: Nonlinear modulation of gravity waves: a rigorous approach. *Nonlinearity* 5(2), 497–522, 1992
6. Cummings, P., Wayne, C.E.: Modified energy functionals and the NLS approximation. *Discrete Contin. Dyn. Syst.* 37(3), 1295–1321, 2017
7. Deng, Y., Ionescu, A.D., Pausader, B., Pusateri, F.: Global solutions of the gravity-capillary water-wave system in three dimensions. *Acta Math.* 219(2), 213–402, 2017
8. Düll, W.-P.: Validity of the Korteweg-de Vries approximation for the two-dimensional water wave problem in the arc length formulation. *Commun. Pure Appl. Math.* 65(3), 381–429, 2012
9. Düll, W.-P.: Justification of the Nonlinear Schrödinger approximation for a quasilinear Klein–Gordon equation. *Commun. Math. Phys.* 355(3), 1189–1207, 2017
10. Düll, W.-P.: On the mathematical description of time-dependent surface water waves. *Jahresber. Dtsch. Math.-Ver.* 120(2), 117–141, 2018
11. Düll, W.-P., Herrmann, 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(2), 847–867, 2016
12. Düll, W.-P., Hess, M.: Existence of long time solutions and validity of the Nonlinear Schrödinger approximation for a quasilinear dispersive equation. *J. Differ. Equ.* 264(4), 2598–2632, 2018
13. 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
14. 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
15. Gaison, J., Moskow, S., Wright, J.D., Zhang, Q.: Approximation of polyatomic FPU lattices by KdV equations. *Multiscale Model. Simul.* 12(3), 953–995, 2014
16.Germain, P., Masmoudi, N., Shatah, J.: Global solutions for the gravity water waves equation in dimension 3. *Ann. Math.* (2) 175(2), 691–754, 2012
17. Harrop-Griffiths, B., Ifrim, M., Tataru, B.: Finite depth gravity water waves in holomorphic coordinates. *Ann. PDE* 3(1), Art. 4, 102 pp., 2017
18. Hess, M.: Validity of the Nonlinear Schrödinger approximation for quasilinear dispersive systems. PhD thesis, Universität Stuttgart, 2019.
19. Hisako, M.: Breather trapping mechanism in piecewise homogeneous DNA. *Phys. Lett. A* 227(1–2), 87–93, 1997
20. Hunter, J.K., Ifrim, M., Tataru, D.: Two dimensional water waves in holomorphic coordinates. *Commun. Math. Phys.* 346(2), 483–552, 2016
21. Hunter, J.K., Ifrim, M., Tataru, D., Wong, T.K.: Long time solutions for a Burgers-Hilbert equation via a modified energy method. *Proc. Am. Math. Soc.* 143(8), 3407–3412, 2015
22. Ifrim, M., Tataru, D.: Two dimensional water waves in holomorphic coordinates II: global solutions. *Bull. Soc. Math. France* 144(2), 369–394, 2016
23. Ifrim, M., Tataru, D.: The lifespan of small data solutions in two dimensional capillary water waves. *Arch. Ration. Mech. Anal.* 225(3), 1279–1346, 2017
24. Ifrim, M., Tataru, D.: The NLS approximation for two dimensional deep gravity waves. *Sci. China Math* 62(6), 1101–1120, 2019
25. Ionescu, A.D., Pusateri, F.: Global solutions for the gravity water waves system in 2d. *Invent. Math.* 199(3), 653–804, 2015
26. Ionescu, A.D., Pusateri, F.: Global regularity for 2D water waves with surface tension. *Mem. Amer. Math. Soc.* 256(1227), v+124pp., 2018
27. Janssen, P., Onato, M.: The intermediate water depth limit and the Zakharov equation and consequences for wave prediction. *J. Phys. Oceanogr.* 37, 2389–2400, 2007
28. Leblond, H.: Direct derivation of a macroscopic NLS equation from quantum theory. *J. Phys. A Math. Gen.* 34, 3109–3123, 2001
29. Kalyakin, L.A.: Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. *Sb. Math.* 60, 457–483, 1988
30. **Klainerman**, S., **Ponce**, G.: Global, small amplitude solutions to nonlinear evolution equations. *Commun. Pure Appl. Math.*, 36(1), 133–141, 1983

31. **Masmoudi**, N., **Nakanishi**, K.: Multifrequency NLS scaling for a model equation of gravity-capillary waves. *Commun. Pure Appl. Math.*, 66(8), 1202–1240, 2013

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

33. **Schneider**, G.: Approximation of the Korteweg-de Vries equation by the Nonlinear Schrödinger equation. *J. Differ. Equ.*, 147(2), 333–354, 1998

34. **Schneider**, G.: Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. *J. Differ. Equ.*, 216(2), 354–386, 2005

35. **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.

36. **Schneider**, G.: Justification of the NLS approximation for the KdV equation using the Miura transformation. *Adv. Math. Phys.*, Art. ID 854719, 4 pp., 2011

37. **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. Dyn. Differ. Equ.*, 27(3), 1077–1099, 2015

38. **Schneider**, G., **Wayne**, C.E.: The long wave limit for the water wave problem. I. The case of zero surface tension. *Commun. Pure Appl. Math.*, 53(12), 1475–1535, 2000

39. **Schneider**, G., **Wayne**, C.E.: The rigorous approximation of long-wavelength capillary-gravity waves. *Arch. Ration. Mech. Anal.*, 162(3), 247–285, 2002

40. **Schneider**, G., **Wayne**, C.E.: Justification of the NLS approximation for a quasilinear water wave model. *J. Differ. Equ.*, 251(2), 238–269, 2011

41. **Shatah**, J.: Normal forms and quadratic nonlinear Klein–Gordon equations. *Commun. Pure Appl. Math.*, 38(5), 685–696, 1985

42. **Sulem**, C., **Sulem**, P.L.: The nonlinear Schrödinger equation. Self-focusing and wave collapse. *Applied Mathematical Sciences* 139, Springer, New York, 1999.

43. **Shatah**, J., **Zeng**, C.: Local well-posedness for fluid interface problems. *Arch. Ration. Mech. Anal.*, 199(2), 653–705, 2011

44. **Totz**, N.: A justification of the modulation approximation to the 3d full water wave problem. *Commun. Math. Phys.*, 335(1), 369–443, 2015

45. **Totz**, N., **Wu**, S.: A rigorous justification of the modulation approximation to the 2D full water wave problem. *Commun. Math. Phys.*, 310(3), 817–883, 2012

46. **Wang**, X.: Global regularity for the 3D finite depth capillary water waves. *arXiv:1611.05472*, 2016

47. **Wang**, X.: Global solution for the 3D gravity water waves system above a flat bottom. *Adv. Math.*, 346, 805–886, 2019

48. **Wu**, S.: Global well-posedness of the 3-D full water wave problem. *Invent. Math.*, 184(1), 125–220, 2011

49. **Yoshihara**, H.: Capillary-gravity waves for an incompressible ideal fluid. *J. Math Kyoto Univ.*, 23(4), 649–694, 1983

50. **Zakharov**, V.E.: Stability of periodic waves of finite amplitude on the surface of a deep fluid. *Sov. Phys. J. Appl. Mech. Tech. Phys.*, 4, 190–194, 1968
