Validity of the nonlinear Schrödinger approximation for quasilinear dispersive systems with more than one derivative

Max Heß
Institut für Analysis, Dynamik und Modellierung, Universität Stuttgart, Stuttgart, Germany

Correspondence
Max Heß, Universität Stuttgart, Institut für Analysis, Dynamik und Modellierung; Pfaffenwaldring 57, D-70569 Stuttgart, Germany.
Email: max.hess@mathematik.uni-stuttgart.de

Communicated by: M. Groves

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: DU 1198/2

For nonlinear dispersive systems, the nonlinear Schrödinger (NLS) equation can usually be derived as a formal approximation equation describing slow spatial and temporal modulations of the envelope of a spatially and temporally oscillating underlying carrier wave. Here, we justify the NLS approximation for a whole class of quasilinear dispersive systems, which also includes toy models for the waterwave problem. This is the first time that this is done for systems, where a quasilinear quadratic term is allowed to effectively lose more than one derivative. With effective loss, we here mean the loss still present after making a diagonalization of the linear part of the system such that all linear operators in this diagonalization have the same regularity properties.

KEYWORDS
error estimates, NLS approximation, normal forms, quasilinear PDEs

MSC CLASSIFICATION
35A35; 35Q55; 35L77; 35Q35

1 INTRODUCTION

Nonlinear dispersive systems can be very difficult to solve as well analytically as numerically. Thus, a valid nonlinear Schrödinger (NLS) approximation can be a great tool for understanding the dynamics of such systems. An introduction to this theory can be found in Schneider and Uecker. For the sake of simplicity, we will restrict ourselves in this introduction to the basic prototype equation

\[ \partial_t^2 u = -\partial_x^4 u - \partial_x^2 (u^2), \]  

(1)

with \( u : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, t) \mapsto u(x, t) \), which is a quasilinear beam equation. Besides modeling deformations of an elastic beam, it appears for surface waves in shallow water, in the dislocation theory of crystals or for the interaction between waves guides and some external medium, compare, for example, This equation is part of the class of quasilinear dispersive systems for which we justify the NLS approximation in this paper, namely,\n
\[ \partial_t u = -i \omega v, \]
\[ \partial_t v = -iou - i\rho u^2, \]  

(2)
with $\rho$ and $\omega$ being differential operators. These systems are already diagonalized such that the operators acting on the linear terms both cause the same loss of regularity; thus, the effective loss caused by the quasilinear terms is the loss caused by $\rho$. For (1), the operators $\omega$ and $\rho$ are given in Fourier space by the multipliers $\omega(k) = \rho(k) = \text{sign}(k)k^2$, so (1) has a quasilinear quadratic term effectively losing two derivatives.

In order to derive the NLS equation for (2), we make an ansatz of the form

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

where

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

$$0 < \varepsilon \ll 1$$ is a small perturbation parameter, $A$ is a complex-valued amplitude, and c.c. the complex conjugate (Figure 1). The ansatz leads to a basic temporal wave number $\omega_0 = \omega(k_0)$ associated to the basic spatial wave number $k_0 > 0$ and a group velocity $c_g = \omega'(k_0)$. Most importantly, the NLS equation

$$\partial_T A = i \frac{\omega'(k_0)}{2} \partial_X^2 A + i v_2(k_0) A|A|^2$$

is obtained as a lowest order modulation equation, describing slow modulations in time and space of the envelope of the wave packet. $T = \varepsilon^2 t$ is the slow time scale; $X = \varepsilon (x - c_g t)$ the slow spatial scale and $v_2(k_0) \in \mathbb{R}$. A formula for $v_2(k_0)$ is given later. The NLS Equation (4) can be explicitly solved; see, for example.5

We prove the following result.

**Theorem 1.1.** Fix $\omega$, $\rho$ and $k_0 > 0$. For $s_A \geq 7$ and all $C_1$, $T_0 > 0$ there exists some $\varepsilon_0$, $C_2 > 0$ such that for all solutions $A \in C([0, T_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ of the NLS Equation (4) with

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

the following holds.

For all $\varepsilon \in (0, \varepsilon_0)$ there are solutions

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

of the original system (2) which satisfy

$$\sup_{T \in [0, T_0/\varepsilon^2]} \|u(\cdot, t) - \varepsilon \psi_{\text{NLS}}(\cdot, t)\|_{H^{s_A}(\mathbb{R}, \mathbb{C})} \leq \varepsilon^{3/2} C_2.$$
a quasilinear quadratic term in the presence of nontrivial resonances,
a nonlinearity that causes the loss of more than one derivative in the error estimates.

This is the first paper where an NLS justification result is given for a quasilinear system where the nonlinearity causes an effective loss of more than one derivative. In this case a qualitatively new analysis is needed since a loss of regularity in the error evolution can no longer be dodged through integration by parts like in previous articles.

A quadratic term yields in the equations of the error to terms of order $O(\varepsilon)$, which could potentially lead to an explosion on a time scale of order $O(\varepsilon^{-2})$. In numerous articles, a theory to handle quadratic terms by using normal form transformations was developed; see, for example, Kalyakin\textsuperscript{7} and for the case of resonances.\textsuperscript{8} However, quasilinear quadratic terms were explicitly excluded. Such terms make the closing of error estimates much harder. The first NLS validity results for systems with a quasilinear quadratic term (losing a half derivative) then were proven in Schneider and Wayne\textsuperscript{9} and Totz and Wu.\textsuperscript{10} Using a modified energy to handle the occurring loss of regularity in the normal form NLS validity results for systems with quadratic terms that effectively lose a whole derivative could be proven in previous studies.\textsuperscript{11,12} We here now for the first time prove a NLS validity result for a class of systems with quasilinear quadratic terms that can effectively lose an arbitrary amount of derivatives.

It has to be mentioned that the problem of proving NLS justification results has always been closely related to the water wave problem (WWP). Indeed, the first time the NLS equation was derived as a model equation was for the WWP in Zakharov.\textsuperscript{13} The (2-D) WWP is the problem of finding the irrational flow of an incompressible fluid in an infinitely long canal with flat bottom and a free surface under the influence of gravity. For the WWP the NLS equation was rigorously justified on the right time scale by Totz and Wu for the case of zero surface tension and infinite depth in Totz and Wu\textsuperscript{10} and by Düll, Schneider and Wayne for the case of zero surface tension and finite depth.\textsuperscript{14} In both cases quasilinear terms lose effectively a half derivative. In Ifrim and Tataru\textsuperscript{15} a result similar to the one of\textsuperscript{10} was proven by using a modified energy method. The NLS equation for the WWP in case of finite depth and possibly of surface tension was justified, uniformly with respect to the strength of the surface tension as the height of the wave packet and the surface tension go to zero, in Düll.\textsuperscript{16} In Düll\textsuperscript{16} it was also shown that due to cancellations, there only is an effective loss of one derivative, although one might expect a loss of one and a half derivatives. A claim that such cancellations always occur for physical systems of relevance would be quite optimistic. Our article now also hopes to contribute to a justification of the WWP in the case of finite depth and an ice cover, in which one expects an effective loss of up to two derivatives. We give more details on how system (2) can be viewed as a toy model for the WWP in the discussions section at the end of the article.

The plan of the paper is as follows.

We first quickly explain how the NLS approximation is derived and the residual estimates are proven for (2). Then we justify the NLS approximation by proving error estimates. In order to obtain the natural $O(\varepsilon^{-2})$-time scale of the NLS equation for the error, we use a modified energy method, that is, we use a modified energy based on normal form transformations that is equivalent to the squared Sobolev norm of the error. The evolution of this energy then still contains terms that cannot be estimated without a loss of regularity since an effective loss of more than one derivative is allowed. In order to control these terms, we recursively construct an expression of order $O(\varepsilon)$, whose time derivative cancels out with the problematic terms. The construction mainly exploits the time space relation given for the error and the strict concentration of the NLS approximation in Fourier space around certain multiples of the wavenumber $k_0$. By adding the constructed $O(\varepsilon)$-term to our energy, we can then close the energy estimates such that Theorem 1.1 follows with Gronwall’s inequality.

**Notation.** The Fourier transform of a function $u \in L^2(\mathbb{R}, \mathbb{K})$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, is denoted by $\mathcal{F}(u)(k) = \hat{u}(k) = \frac{1}{2\pi} \int_{\mathbb{R}} u(x) e^{-ikx} dx$. $H^s(\mathbb{R}, \mathbb{K})$ is the space of functions 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. The space $L^1(s)(\mathbb{R}, \mathbb{K})$ is defined by $u \in L^1(s)(\mathbb{R}, \mathbb{K}) \Leftrightarrow u \sigma^s \in L^1(\mathbb{R}, \mathbb{K})$, where $\sigma(x) = (1 + x^2)^{s/2}$. We use $|a| := \min\{\{z \in \mathbb{Z} : z \geq a\}$. We write $I \leq O(E)$ for expressions $I$ and $E$, when there exists some constant $C > 0$ such that $I \leq CE$. This constant can then always be chosen independently of $E$ and the small perturbation parameter $\varepsilon$.

## 2 | THE GENERAL CLASS OF SYSTEMS

The class of systems for which we consider the NLS approximation consists of the quasilinear dispersive first-order systems (2) where the pseudo-differential operators $\omega$ and $\rho$ can be expressed through some odd real-valued functions $\rho$ and
\(\omega\) in Fourier space. Such a first-order system is also equivalent to a quasilinear dispersive equation

\[
\partial_t^2 u = -\omega^2 u - \rho \omega u^2. \tag{5}
\]

We do not allow the quadratic term of the system to contain more derivatives than the linear one. We express this by demanding

\[
\deg^*(\rho) \leq \deg(\omega), \tag{6}
\]

where we write \(\deg^*(\gamma) \leq s\) for a function \(\gamma : \mathbb{R} \to \mathbb{R}\) when there exists some constant \(C\) such that \(|\gamma(k)| \leq C (1 + |k|)^s\) for large \(|k|\) and \(\deg(\gamma) = s\) when there also is some \(c > 0\) such that \(c (1 + |k|)^s \leq |\gamma(k)| \leq C (1 + |k|)^s\) for large \(|k|\). One of the functions \(\omega\) or \(\text{sign}(-\omega)\) as well as \(\rho\) or \(\text{sign}(-\rho)\) has to lie in \(C^{n_o}(\mathbb{R})\) for \(m_\omega := \max\{5, [\deg(\omega)] + 1\}\). In other words, we allow \(\omega\) and \(\rho\) to have a jump in \(k = 0\). We further demand that

\[
\deg^*(\rho^{(n)}) \leq \deg^*(\rho^{(n-1)}) - 1, \tag{7}
\]

\[
\deg(\omega^{(n)}) = \deg(\omega^{(n-1)}) - 1, \tag{8}
\]

for \(n = 1, \ldots, m_\omega\) as long as \(\rho^{(n)} \neq 0\), respectively, \(\omega^{(n)} \neq 0\), which is a behavior typical for most differential operators.

For the derivation of the NLS Equation (4), we require that for the wavenumber \(k_0 > 0\):

\[
\omega''(k_0) \neq 0, \tag{9}
\]

\[
m_\omega(k_0) \neq \pm \omega(mk_0) \quad \text{for} \quad m = \pm 2, \ldots, \pm 5 \tag{10}
\]

and at least one of the two conditions

\[
\omega'(k_0) \neq \pm \omega'(0) \quad \text{and} \quad \rho(0) = 0, \tag{H1}
\]

\[
\lim_{k \to 0^+} \omega(k) = \omega(0^+) \neq 0 \tag{H2}
\]

is fulfilled. These conditions are necessary for our derivation of the NLS equation and residual estimates; for example, (4) is not the NLS equation if (9) is not true. In this context, they arise naturally for the NLS approximation (3).

Real-valued solutions to the equations

\[
\omega(k) - j_1 j_2 \omega(k \mp k_0) + j_1 \omega(\pm k_0) = 0 \tag{11}
\]

with \(j_1, j_2 \in \{\pm 1\}\) are called resonances. Resonances are problematic in the presence of quadratic terms and have to be avoided for a well-defined normal-form transformation. A resonance in \(k = k_1\) is called trivial, if the quadratic term vanishes for the wavenumber \(k_1\) in Fourier space. We here restrict us to the case that only a trivial resonance in \(k = 0\) and nontrivial resonances in \(k = \pm k_0\) can occur. Allowing more resonances would be possible by making some adjustments as in previous studies,\(^\text{16,17}\) We naturally also forbid the left-hand side of (11) converging to zero for \(|k| \to \infty\). If (H2) is true, we additionally have to demand that

\[
\omega(0^+) \neq \pm 2\omega(k_0) \tag{H2a}
\]

or

\[
(H1) \quad \text{is true and} \quad \omega(0^+) \neq 2\omega(k_0) + j\omega(2k_0) \quad \text{for} \quad j \in \{\pm 1\} \tag{H2b}
\]

to rule out further resonances.

For this class of systems, we prove Theorem 1.1.

For the quasilinear beam Equation (1) the conditions (6)–(10) and (H1) are obviously true for all \(k_0 > 0\), and (H2) is false. To verify that resonances can only occur in \(k = 0\) and \(k = \pm k_0\), one makes a case analysis of (11) with the quadratic formula.
3 | RESIDUAL ESTIMATES

All coming calculations get much easier by working with diagonalized system

\[\begin{align*}
\partial_t u_{-1}(x, t) &= -i\omega u_{-1}(x, t) - \frac{1}{2}i\rho(u_{-1} + u_1)^2(x, t), \\
\partial_t u_1(x, t) &= i\omega u_1(x, t) + \frac{1}{2}i\rho(u_{-1} + u_1)^2(x, t),
\end{align*}\]

(12)

with \(u_{-1}(x, t), u_1(x, t) \in \mathbb{R}\). One obtains this system from (2) via the invertible transformation

\[\left( \begin{array}{c}
u_{-1} \\ u_1
\end{array} \right) = \frac{1}{2} \left( \begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right) \left( \begin{array}{c} u \\ v
\end{array} \right),\]

(13)

With a simple ansatz of the form

\[\left( \begin{array}{c}
u_{-1} \\ u_1
\end{array} \right) = \left( \begin{array}{c} \varepsilon \Psi^{\text{NLS}}/\varepsilon \\ 0
\end{array} \right) + \mathcal{O}(\varepsilon^2) \left( \begin{array}{c} 1 \\ 1
\end{array} \right),\]

one can derive the NLS Equation (4) by expanding the operators \(\omega, \rho\) in Fourier space around integer multiples of the basic wave number \(k_0\) with Taylor’s theorem, compare.\(^{12}\) We obtain the NLS Equation (4) with

\[v_2(k_0) = -\rho(k_0) \left( \frac{\rho(2k_0)\omega(2k_0)}{4(\omega(k_0))^2 - (\omega(2k_0))^2} + \frac{2\rho'(0)\omega'(0)}{(\omega'(0))^2 - (\omega(0))^2} \right),\]

when (H2) is not true and with

\[v_2(k_0) = -\rho(k_0) \left( \frac{\rho(2k_0)\omega(2k_0)}{4(\omega(k_0))^2 - (\omega(2k_0))^2} - 2\frac{\rho(0^+)}{\omega(0^+)} \right),\]

when (H2) is true. The explicit computation done here can be found in Heß.\(^{18}\)

The residual \(\text{Res}_\varepsilon(\varepsilon \Psi)\) of an approximation \(\varepsilon \Psi\) denotes all terms that remain after plugging in an approximation \(\varepsilon \Psi\) into the equations of system (2). For the coming error estimates, a very small residual whose norm can be controlled in high Sobolev spaces is needed. For this reason an improved approximation \(\varepsilon \Psi\) is used. By exploiting that an ansatz like above is always strongly concentrated around a finite number of integer multiples of the basic wave number \(k_0 > 0\), cut-off functions can be used to restrict the support of an ansatz in Fourier space to small neighborhoods of these wave numbers \(jk_0\) with \(j \in \{-5, \ldots, 5\}\). This way, an approximation \(\varepsilon \Psi\) that is an analytic function and has a residual of the formal order \(\mathcal{O}(\varepsilon^6)\) is obtained, compare Section 2 of Düll, Schneider, & Wayne.\(^{14}\)

The approximation that we use is

\[\varepsilon \Psi = \varepsilon \Psi_c + \varepsilon^2 \Psi_q,\]

(14)

where

\[\varepsilon \Psi_c = \varepsilon \psi_c \left( \begin{array}{c} 1 \\ 0
\end{array} \right) = \varepsilon (\psi_{c_1} + \psi_{c_{-1}}) \left( \begin{array}{c} 1 \\ 0
\end{array} \right),\]

\[\varepsilon^2 \Psi_q = \varepsilon^2 \left( \begin{array}{c} \psi_{c_{+1}} \\ \psi_{c_{-1}}
\end{array} \right) = \varepsilon^2 \Psi_0 + \varepsilon^2 \Psi_2 + \varepsilon^2 \Psi_\text{R},\]

\[\varepsilon^2 \Psi_0 = \varepsilon^2 \left( \begin{array}{c} A_0(\varepsilon (x - c_{g_1}t), \varepsilon^2 t) \\ D_0(\varepsilon (x - c_{g_1}t), \varepsilon^2 t)
\end{array} \right),\]

\[\varepsilon^2 \Psi_2 = \varepsilon^2 \left( \begin{array}{c} A_2(\varepsilon (x - c_{g_1}t), \varepsilon^2 t) E^2 + \text{c.c.} \\ D_2(\varepsilon (x - c_{g_1}t), \varepsilon^2 t) E^2 + \text{c.c.}
\end{array} \right),\]
$$\varepsilon^2 \Psi_i = \sum_{n=1,2,3,4} \varepsilon^{1+n} \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E + c.c. \right) \left( D_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E + c.c. \right)$$

$$+ \sum_{n=1,2,3} \varepsilon^{2+n} \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) \right)$$

$$+ \sum_{n=1,2,3} \varepsilon^{2+n} \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) \right) + c.c. \left( D_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^3 + c.c. \right)$$

$$+ \sum_{n=0,1,2} \varepsilon^{3+n} \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^3 + c.c. \right) \left( D_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^3 + c.c. \right)$$

$$+ \sum_{n=0,1} \varepsilon^{4+n} \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^4 + c.c. \right) \left( D_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^4 + c.c. \right)$$

$$+ \varepsilon^5 \left( A_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^5 + c.c. \right) \left( D_0^n(\varepsilon(x - c_g t), \varepsilon^2 t) E^5 + c.c. \right),$$

where $E = e^{i(k_0 x - \omega_0 t)}$, $\omega_0 = \omega(k_0)$ and $c_g = c_g'(k_0)$. Here, $A_1(\varepsilon(-c_g t), \varepsilon^2 t)$ is the restriction of $A(\varepsilon(-c_g t), \varepsilon^2 t)$ in Fourier space to the interval $\{k \in \mathbb{R} : |k| \leq \delta < k_0/20 \}$ by some cut-off function, while $A$ is the solution of the NLS Equation (4) and $\delta > 0$. More precisely,

$$A_1(\varepsilon(-c_g t), \varepsilon^2 t) := F^{-1} [\chi_{[-\delta, \delta]}(\cdot)F \left[ A(\varepsilon(-c_g t), \varepsilon^2 t) \right] (\cdot)],$$

where $\chi_{[-\delta, \delta]}$ is the characteristic function on the interval $[-\delta, \delta]$, that is, $\chi_{[-\delta, \delta]}(k) = 1$ for $[-\delta, \delta]$ and $\chi_{[-\delta, \delta]}(k) = 0$ for $k \notin [-\delta, \delta]$. One can think of $\varepsilon \Psi$, as $\varepsilon \Psi_{\text{NLS}}$ with a support in Fourier space restricted to small neighborhoods of the wave numbers $\pm k_0$. The $A_0^n$ and $D_0^n$ are chosen suitably depending on $A_1$ such that the supports of $A_0^n E^j$ and $D_0^n E^j$ in Fourier space lie in small neighborhoods of the wave number $jk_0$.

Similarly as in Düll, Schneider, & Wayne, one obtains the following:

**Lemma 3.1.** Let $s_A \geq 7$ and $A \in C([0, T_0], H^{s_A}(\mathbb{R}, \mathbb{C}))$ be a solution of the NLS Equation (4) with

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

Then for all $s \geq 0$, there exist $C_{\text{Res}}, C_\Psi, \varepsilon_0 > 0$ depending on $C_A$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the approximation $\varepsilon \Psi = \varepsilon \Psi_c + \varepsilon^2 \Psi_q$ satisfies

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

$$\sup_{t \in [0, T_0/\varepsilon^2]} \|\varepsilon \Psi - \varepsilon \Psi_{\text{NLS}} \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] \|_{H^{s_A}} \leq C_\Psi \varepsilon^{3/2},$$

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

$$\|\partial_t \Psi_{\pm 1} + i \omega \Psi_{\pm 1}\|_{L^{s}} \leq C_\Psi \varepsilon^2.$$
4 | THE ERROR ESTIMATES

We write the error of the approximation \( \epsilon \Psi \) as

\[
\epsilon^{\theta} \left( \frac{\partial R_{-1}}{\partial R_1} \right) := \left( \frac{u_{-1}}{u_1} \right) - \epsilon \Psi, \tag{20}
\]

where \( R_{-1} \) and \( R_1 \) are our error functions, \( \beta = 5/2 \), and \( \theta \) is an invertible operator given in Fourier space either by the weight function

\[
\hat{\theta}(k) = \begin{cases} 
\epsilon + (1 - \epsilon) \frac{|k|}{\delta} & \text{for } |k| \leq \delta, \\
1 & \text{for } |k| > \delta,
\end{cases} \tag{21}
\]
or by \( \hat{\theta}(k) = 1 \) if (H2) and (H2a) are true. The fixed parameter \( \delta \) is as in (15).

The inclusion of the operator \( \theta \) is essential for our handling of the nontrivial resonances in \( k = \pm k_0 \). For this purpose \( \theta \) has also already been used, for example, in previous studies.\(^{12,17}\) When \( 0 \not= \pm \omega(0^+) \neq 2\omega(k_0) \), there is no resonance in \( k = k_0 \) such that setting \( \hat{\theta}(k) = 1 \) is better.

By plugging in the above definition into the diagonalized system, we obtain the following dynamics for the error

\[
\begin{align*}
\partial_t R_{-1} &= -i \omega R_{-1} - \epsilon i \rho \theta^{-1}(R_{-1} + \theta R_1) + \epsilon^{-\beta} \theta^{-1} \text{Res}_{-z_1}(\epsilon \Psi), \\
\partial_t R_1 &= -i \omega R_1 + \epsilon i \rho \theta^{-1}(R_{-1} + \theta R_1) + \epsilon^{-\beta} \theta^{-1} \text{Res}_{\eta_1}(\epsilon \Psi),
\end{align*} \tag{22}
\]

where

\[
R_{-1} := \psi + \frac{1}{2} \epsilon \theta^{-1}(R_{-1} + \theta R_1), \tag{23}
\]

\[
\psi := \psi_c + \epsilon (\psi_{q_+} + \psi_{q-}). \tag{24}
\]

Following the idea of Düll and Heß,\(^{12}\) we define the modified energy

\[
E_\epsilon = E_0 + E_\epsilon, \tag{25}
\]

\[
E_\epsilon(R) = \|R_{-1}\|_{L^2}^2 + \|R_1\|_{L^2}^2,
\]

where

\[
\hat{R}_j = R_j + \epsilon \sum_{j_2 \in \{\pm 1\}} \theta^{-1}N_{j_1j_2}(\psi_c, R_{j_2}) + \epsilon^2 \sum_{j_2, j_3 \in \{\pm 1\}} \theta^{-1}T_{j_1j_2j_3}(\psi_{j_1}, \psi_{j_2}, R_{j_3}),
\]

\[
\hat{N}_{j_1j_2}(\psi_c, R_{j_2})(k) = \int n_{j_1j_2}(k, k - m, m) \hat{\psi}_c(k - m) \hat{R}_{j_2}(m) \, dm,
\]

\[
\hat{T}_{j_1j_2j_3}(\psi_{j_1}, \psi_{j_2}, R_{j_3})(k) = \int t_{j_1j_2j_3}(k) \hat{\psi}_{j_1}(k - m) \hat{\psi}_{j_2}(m - n) \hat{R}_{j_3}(n) \, dn \, dm,
\]

\[
n_{j_1j_2}(k, k - m, m) = \frac{\rho(k) \hat{\theta}_{\epsilon, \infty}(m) \chi_c(k - m)}{\omega(k) - j_1 j_2 \omega(m) + j_1 \omega(k - m)},
\]

\[
t_{j_1j_2j_3}(k) = \frac{-j_2 \hat{p}_{0,\delta}(k) n_{j_1j_2}(k, j_3 k_0, k - j_4 k_0) \rho(k - j_3 k_0)}{(-j_1 \omega(k) - 2j_2 \omega(j_3 k_0) + j_3 \omega(k - 2j_4 k_0))},
\]

\[
\hat{\theta}_{\epsilon, \infty}(m) = \begin{cases} 
0 & \text{for } |m| \leq \epsilon, \\
\epsilon + (1 - \epsilon) \frac{|m|}{\delta} & \text{for } \epsilon < |m| \leq \delta, \\
1 & \text{for } |m| > \delta,
\end{cases}
\]

and \( \chi_c \) denotes the characteristic function on \( \text{supp} \hat{\psi}_c \). We set \( t_{j_1j_2j_3} = 0 \) and \( \hat{\theta}_{\epsilon, \infty} = 1 \) when \( 0 \not= \pm \omega(0^+) \neq 2\omega(k_0) \).
The final handling of the loss of regularity stemming from quasilinear quadratic terms losing more than one derivative is done in Section 4.3, by transforming this energy. After the transformation, the energy estimates close and the theorem is proven.

4.1 Natural time scale of the NLS

In order to achieve error estimates valid on the natural $O(\epsilon^{-2})$-time scale of the NLS equation, the evolution of the energy has to be of order $O(\epsilon^2)$. To met this goal, we chose the normal form transformations such that the $O(\epsilon)$-terms in the error evolution are eliminated.

The operator $\theta^{-1}$ was placed outside of the normal form transforms $N_{j_1j_2}$ and $T_{j_1j_2j_3j_4}$ since in general we only have

$$\hat{\theta}^{-1}(k) = \frac{1}{\hat{\theta}(k)} = O(\epsilon^{-1}).$$

(26)

However, when $\theta^{-1}$ appears next to a spatial derivative, better estimates become possible:

Lemma 4.1. We have

$$|k\hat{\theta}^{-1}(k)| \leq 1 + |k|.$$  \hspace{1cm} (27)

In particular,

$$\|i\rho \theta^{-1} f\|_{L^2} \leq O(\|f\|_{L^{\deg^*}}).$$  \hspace{1cm} (28)

Proof. The lemma is obviously true for $\hat{\theta}(k) = 1$. Otherwise, we have $|k\hat{\theta}^{-1}(k)| = |k|$ for $|k| > \delta$ and

$$|k\hat{\theta}^{-1}(k)| = \frac{|k|}{\epsilon + (1 - \epsilon)\frac{|k|}{\delta}} \leq \frac{1}{\epsilon} + \frac{1 - \epsilon}{\delta} \leq \delta$$

for $0 < |k| \leq \delta$ such that (27) is true. When $\hat{\theta}^{-1} \neq 1$ we are in the case (H1). Thus $\rho(k) = O(k)$ for $|k| \to 0$ such that (28) follows.

Lemma 4.2. The normal form transforms $N_{j_1j_2}$ were constructed such that for all $f \in H^{\deg^*(\rho) + 1}(\mathbb{R})$:

$$-j_1i\omega N_{j_1j_2}(\psi_c, f) - N_{j_1j_2}(i\omega \psi_c, f) + j_2N_{j_1j_2}(\psi_c, i\omega f) = -j_1i\rho(\psi_c \theta_{\epsilon, oo} f),$$

(29)

where

$$\epsilon \|j_1i\rho \theta^{-1}(\psi \theta f) - j_1i\rho \theta^{-1}(\psi_c \theta_{\epsilon, oo} f)\|_{L^2} = O(\epsilon^2)\|f\|_{L^{\deg^*(\rho)}}.$$  \hspace{1cm} (30)

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

$$\|N_{j_1j_2}(h, g)\|_{L^2} \leq C\|g\|_{H^1},$$  \hspace{1cm} (31)

$$\|N_{j_1j_2}(h, g)\|_{L^2} \leq C\|g\|_{L^2}.$$  \hspace{1cm} (32)

Proof. In order prove that the $N_{j_1j_2}$ are well-defined, we have to look at the zeros of the denominator of $N_{j_1j_2}$, that is, of

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

for $|k - m + k_0| \leq \delta$. Due to the assumption for (11) in Section 2, we can choose $\delta$ such small that for $|k - m + k_0| \leq \delta$ the equation

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

(33)

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

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

$$
\begin{align*}
\omega(k) - j_1 j_2 \omega(m) + j_1 \omega(k - m) \\
= \omega(\text{sign}(k) \cdot 0^+) - j_1 j_2 \omega(m) + j_1 \omega(-m) \\
+ \omega'(\text{sign}(k) \cdot 0^+) k + j_1 \omega'(-m) k + O(k^2)
\end{align*}
$$

Thus, if

$$
\omega(0^+) \neq (j_2 + 1) \omega(k_0), \tag{34}
$$

and we choose $\delta$ small enough, $N_{j_1 j_2}$ has no resonance in $k = 0$.

If (34) is not true but

$$
\pm \omega'(0) \neq \omega'(k_0), \tag{35}
$$

we can choose $\delta$ small enough such that

$$
\omega(k) - j_1 j_2 \omega(m) + j_1 \omega(k - m) = O(k) \quad \text{for} \quad k \to 0.
$$

When (H2) and (H2a) are true, we have (34) and thus $N_{j_1 j_2}$ has no resonance in $k = 0$.

Otherwise (H1) is true, such that we have (35) and $\rho(k) = O(k)$ for $k \to 0$. Thus $N_{j_1 j_2}$ can at worst have a trivial resonance in $k = 0$.

The case $m = 0$ works analogously due to symmetry and the choice of $\theta$. To give more details, in the problematic case

$$
\omega(k) - j_1 j_2 \omega(m) + j_1 \omega(k - m) = O(m) \quad \text{for} \quad m \to 0,
$$

there occur no nontrivial resonances or a loss of $\epsilon$-powers, since

$$
|m^{-1} \hat{\theta}_{\epsilon, \infty}(m)| \leq 1 + \delta^{-1}.
$$

Resonances for $|k|, |m| \to \infty$ were excluded in Section 2.

The property (29) can be easily checked in Fourier space.

Concerning estimate (30),

$$
\|j_1 i \rho \theta^{-1}(\psi \theta f) - j_1 i \rho \theta^{-1}(\psi_c \theta \epsilon \infty f)\|_{L^2} = \|j_1 i \rho \theta^{-1}(\epsilon (\psi_{q_1} + \psi_{q_2}) \theta f) + i \rho \theta^{-1}(\psi_c (\theta - \theta \epsilon \infty) f)\|_{L^2} = O(\epsilon) \|f\|_{H^{\infty}(\theta, \omega)}.
$$

particularly due to (28) and $(\theta - \hat{\theta}_{\epsilon, \infty}) \leq O(\epsilon)$.

We now will show that the $N_{j_1 j_2}(h, \cdot)$ are continuous linear operators.

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

We first look at $N_{jj}$.

For $|k| \to \infty$, we have

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

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

$$
\omega(k) - \omega(m) = \omega'(m)(k - m) + r(k, k - m, m),
$$

where

$$
r(k, k - m, m) = \left( \sum_{l=2}^{p} \frac{1}{l!} \omega^{(l)}(m)(k - m)^l + O\left(\omega^{(l+1)}(m)\right) \right) \chi_c(k - m),
$$

and

$$
\omega'(0) \neq \omega'(k_0).
$$
for some sufficiently large chosen \( p \geq \lceil \deg^*(\rho) \rceil \). Then we use the expansion

\[
\frac{a}{b + c} = \sum_{l=0}^{n} (-1)^l \frac{a^l}{b^{l+1}} + (-1)^{n+1} \frac{a^{n+1}}{b^{n+1}(b + c)} \quad (b + c \neq 0, b \neq 0).
\]

(37)

Due to the presence of the characteristic function \( \chi_c(k - m) \) in (36), \((k - m)\) is bounded away from zero uniformly, and \( |k| \to \infty \) implies \( |m| \to \infty \). We distinguish the three cases \( \deg(\omega) > 1 \), \( \deg(\omega) = 1 \) and \( \deg(\omega) < 1 \).

If \( \deg(\omega) > 1 \) (i.e., \( \deg(\omega') > 0 \)), the term \( (\omega(k) - \omega(m)) \) is unlimited, we have for \( |k| \to \infty \):

\[
\frac{\chi_c(k - m)}{\omega(k) - \omega(m) + j \omega(k - m)} = \frac{\chi_c(k - m)}{\omega'(m)(k - m) + r(k, k - m, m) + j \omega(k - m)}
\]

\[
= \left( \frac{1}{\omega'(m)(k - m)} - \frac{r(k, k - m, m) + j \omega(k - m)}{\omega'(m)^2 (k - m)^2} \right)
\]

\[
+ \frac{(r(k, k - m, m) + j \omega(k - m))^2}{\omega'(m)^3 (k - m)^3} - \frac{(r(k, k - m, m) + j \omega(k - m))^3}{\omega'(m)^4 (k - m)^4}
\]

\[
\pm \ldots + O(|m|^{-\deg(\omega')}) \chi_c(k - m).
\]

(38)

If \( \deg(\omega) = 1 \) (i.e., \( \deg(\omega') = 0 \)), the term \( (\omega(k) - \omega(m)) \) is bounded; we have

\[
\frac{\chi_c(k - m)}{\omega(k) - \omega(m) + j \omega(k - m)} = \left( \frac{1}{\omega'(m)(k - m) + j \omega(k - m)} + O(|m|^{-1}) \right) \chi_c(k - m), \text{ for } |k| \to \infty.
\]

(39)

If \( \deg(\omega) < 1 \) (i.e., \( \deg(\omega') < 0 \)), the term \( (\omega(k) - \omega(m)) \) tends to zero and there is some \( N = N(\omega') \in \mathbb{N} \) such that

\[
\frac{\chi_c(k - m)}{\omega(k) - \omega(m) + j \omega(k - m)} = \left( \sum_{n=0}^{N} (-1)^n j^{n+1} \frac{(\omega'(m)^n)(k - m)^n}{((\omega(k - m))^n + 1) + O(|m|^{-1})} \right) \chi_c(k - m), \text{ for } |k| \to \infty.
\]

(40)

Due to (6), (7), and (8), we now get that the \( N_j(h, \cdot) \) map \( H^1(\mathbb{R}) \) on \( L^2(\mathbb{R}) \) by taking advantage of Plancherel’s theorem and Young’s inequality for convolutions

\[
\| N_{j}(h, g) \|_{L^2} \lesssim \| \hat{N}_{j}(h, g) \|_{L^2} = \| \int_{\mathbb{R}} n_{j}(\cdot, - m, m) \hat{h}(\cdot - m) \hat{g}(m) dm \|_{L^2}
\]

\[
\leq O \left( \sup_{k, m \in \mathbb{R}} \frac{|n_{j}(k, k - m, m)|}{(|m|^2 + 1)^{1/2}} \right) \| \int_{\mathbb{R}} |\hat{h}(\cdot - m) \chi_c(\cdot - m) (|m|^2 + 1)^{1/2} \hat{g}(m)| dm \|_{L^2}
\]

\[
\leq O \left( \| \hat{h}(\cdot) \chi_c(\cdot) \|_{L^2} \right) \| g \|_{H^1}.
\]

Now, we look at \( N_{j, -j} \).

Using Taylor, we get for \( |k| \to \infty \):

\[
n_{j, -j}(k, k - m, m) = \frac{\rho(k) \chi_c(k - m)}{\omega(k) + \omega(m) + j \omega(k - m) - \rho(k) \chi_c(k - m)}
\]

\[
= \frac{\rho(k) \chi_c(k - m)}{2\omega(k) + r(k, k - m, m) + j \omega(k - m)},
\]

where \( r(k, k - m, m) \) is now given by
\[ r(k, k-m, m) = \sum_{j=1}^{p} \frac{(-1)^j}{l!} \omega^{(j)}(k-m) + O(\omega^{(p+1)}(k)) \]

for some sufficiently large chosen \( p \geq \lceil \text{deg}^* (\rho) \rceil \). Using expansion (37), we obtain

\[ n_{j,-j}(k, k-m, m) = \left( \frac{\rho(k)}{2\omega(k)} - \frac{\rho(k)(r(k, k-m, m) + j\omega(k-m))}{4\omega(k)^2} \right. \]
\[ + \left. \frac{\rho(k)(r(k, k-m, m) + j\omega(k-m))^2}{8\omega(k)^3} \right) \quad \vDash \ldots + O(|k|^{-\text{deg}^* (\rho)}) \chi_c(k-m), \text{ for } |k| \to \infty. \quad (41) \]

We can now see that the \( N_{j,-j}(h, \cdot) \) map \( L^2(\mathbb{R}) \) on \( L^2(\mathbb{R}) \) by exploiting Young’s inequality for convolutions. Finally, since

\[ n_{j, j}(k, -k-m, -m) = n_{j, j}(k, k-m, m) \in \mathbb{R}, \]

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

**Lemma 4.3.** The normal form transforms \( T_{j_1 j_2 j_3 j_4} \) were constructed such that for all \( j_1, j_2, j_3, j_4 \in \{ \pm 1 \} \), we have

\[ \epsilon^2 \| \theta^{-1} Y_{j_1 j_2 j_3 j_4} \|_{L^2} \leq \epsilon^2 \mathcal{O} \left( \| R_{j_1} \|_{H^{\text{reg} + 1}} \right). \quad (42) \]

where

\[ Y_{j_1 j_2 j_3 j_4} = N_{j_1, j_2}(\psi c, j_3 \theta^{-1} i\rho(\psi \theta R_{j_1})) \]
\[ + \sum_{j_i = \pm 1} (-j_i \rho \mathcal{T}_{j_1 j_2 j_3 j_4}(\psi_{j_4}, \psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1 j_2 j_3 j_4}(-i\rho \psi_{j_4}, \psi_{j_4}, R_{j_3}) \]
\[ + \mathcal{T}_{j_1 j_2 j_3 j_4}^{\prime}(\psi_{j_4}, -i\rho \psi_{j_4}, R_{j_3}) + \mathcal{T}_{j_1 j_2 j_3 j_4}^{\prime}(\psi_{j_4}, \psi_{j_4}, j_3 \rho \psi_{j_3})(j_3 \rho R_{j_3}) \). \quad (43) \]

Furthermore, for fixed functions \( g, h \) with \( \tilde{g}, \tilde{h} \in L^1(\mathbb{R}, \mathbb{C}) \), the mapping \( f \mapsto T_{j_1}(g, h, f) \) defines a continuous linear map from \( L^2(\mathbb{R}, \mathbb{C}) \) into \( L^2(\mathbb{R}, \mathbb{C}) \) and there is a constant \( C = C \left( \| \tilde{g} \|_{L^1}, \| \tilde{h} \|_{L^1} \right) \) such that for all \( f \in L^2(\mathbb{R}, \mathbb{C}) \), we have

\[ \| T_{j_1 j_2 j_3 j_4}(g, h, f) \|_{L^2} \leq C \| f \|_{L^2}. \quad (44) \]

**Proof.** If (H2) and (H2a) are true, the lemma is trivial. Otherwise, we have to first look at the zeros of the denominator of \( t_{j_1 j_2 j_3 j_4}(k) \), that is, the zeros of

\[ (\omega(k) - j_1 j_2 \omega(k - j_4 k_0) + j_3 \omega(j_4 k_0)) (-j_1 \omega(k) - 2\omega(j_4 k_0) + j_3 \omega(k - 2 j_4 k_0)) \]

for \( |k| \leq \delta \). For the first factor, we have (11), so the only possible zero of the first factor is \( k = 0 \). For the second factor, we get by expanding the expression \( \omega(k) \) in the point sign \( k \cdot 0^+ \) and \( \omega(k - 2 j_4 k_0) \) in the point \(-2 j_4 k_0 \):

\[ -j_1 \omega(k) - 2\omega(j_4 k_0) + j_3 \omega(k - 2 j_4 k_0) = -j_1 \omega(\text{sign}(k) \cdot 0^+) - 2\omega(j_4 k_0) + j_3 \omega(-2 j_4 k_0) + \mathcal{O}(k). \]

We can choose \( \delta \) such small that this expression has no zeros. For \( \omega(0) = 0 \) this is possible due to (10) and for \( \omega(0) \neq 0 \) due to (H2b). Summing up, there can only occur a trivial resonance in \( k = 0 \). We now obtain (44) by using Young’s inequality for convolutions and the fact that \( \| \tilde{T}_{j_1 j_2 j_3 j_4} \|_{L^\infty} \) can be uniformly bounded.

The estimate (42) is obtained similarly as in Düll and Heß,\(^{12}\) the details can be found in Heß.\(^{16}\) \( \square \)
Lemma 4.4. For \( m \geq \deg^{*}(\rho) + 1 \), we have
\[
\partial_t E_0 \leq \epsilon^2 \Theta \left( \epsilon^{1/2} \left( \| R_{-1} \|_{H^m}^2 + \| R_1 \|_{H^m}^2 \right)^{3/2} + \| R_{-1} \|_{H^m}^2 + \| R_1 \|_{H^m}^2 + 1 \right). \tag{45}
\]

Proof. The statement follows by the construction of the normal form transforms, that is, with (19) and the Lemma 4.2 and 4.3. A complete proof can be found in Heß.\(^{18}\) \( \Box \)

### 4.2 Energy equivalence

In order to obtain a result for the error, the energy \( E_{\varepsilon} \) has to be equivalent to the \( H^{\varepsilon} \)-energy of the error functions. A loss of regularity caused by the normal form transforms is avoided by using the modified energy. The first time a modified energy was used to overcome a loss of regularity was in Hunter et al.\(^{19}\)

**Lemma 4.5.** There are constants \( C_0, \tilde{C}_0 \) such that the following estimates hold
\[
\sqrt{E_0} \leq C_0 (\| R_1 \|_{H^m} + \| R_{-1} \|_{H^m}), \tag{46}
\]
\[
\| R_1 \|_{L^2} + \| R_{-1} \|_{L^2} \leq \tilde{C}_0 \sqrt{E_0} + \epsilon \Theta (\| R_{-1} \|_{L^2} + \| R_1 \|_{L^2}). \tag{47}
\]

Proof. The proof is similar to the one in Düll and Heß\(^{12}\) and can be found in Heß.\(^{18}\) \( \Box \)

**Lemma 4.6.** Let \( f, g, h \in L^2(\mathbb{R}, \mathbb{R}) \) be real-valued functions and \( K : \mathbb{R}^3 \to \mathbb{C} \).

If
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \left| K(k, k - m, m) \overline{f(k)} \hat{h}(k - m) \hat{g}(m) \right| dm dk < \infty,
\]
then we have
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} K(k, k - m, m) \overline{f(k)} \hat{h}(k - m) \hat{g}(m) dm dk = \int_{\mathbb{R}} \int_{\mathbb{R}} K(-m, k - m, -k) \overline{\hat{g}(k)} \hat{h}(k - m) \hat{f}(m) dm dk. \tag{48}
\]

Proof. The result is obtained by exploiting the fact that \( \overline{f(k)} = f(-k) \) and \( \hat{g}(m) = \overline{\hat{g}(-m)} \), making a change of variables and using Fubini’s theorem. \( \Box \)

**Corollary 4.7.** Let \( \epsilon < \epsilon_0 \) and \( \epsilon_0 \) be sufficiently small. For \( \epsilon' \geq 1 \), the energy \( E_{\epsilon'} \) is equivalent to \( (\| R_{-1} \|_{H^m} + \| R_1 \|_{H^m})^2 \), that is, there are constants \( C_1, C_2 > 0 \) such that
\[
(\| R_{-1} \|_{H^m} + \| R_1 \|_{H^m})^2 \leq C_1 E_{\epsilon'} \leq C_2 (\| R_{-1} \|_{H^m} + \| R_1 \|_{H^m})^2.
\]

Proof. The fact that \( \hat{h}(k) = 1 \) for \( k \) outside of the compact set \( [-\delta, \delta] \) gives the a priori estimate
\[
\int_{\mathbb{R}} \partial_x^\alpha f \partial_x^\beta g dx = \int_{\mathbb{R}} \partial_x^\alpha f \partial_x^\beta g dx + O(\| f \|_{L^2} \| g \|_{L^2}). \tag{49}
\]

With (27), we also get
\[
\int_{\mathbb{R}} \partial_x^\alpha f \partial_x^{m+1} g dx = \int_{\mathbb{R}} \partial_x^\alpha f \partial_x^{m+1} g dx + O(\| f \|_{L^2} \| g \|_{L^2}). \tag{50}
\]
Due to (50) and (31) and (32), we have

\[ E_\varepsilon = \frac{1}{2} \sum_{j_1 \in \{\pm 1\}} \| \partial_x R_{j_1} \|^2_{L^2} + \varepsilon \sum_{j_1, j_2 \in \{\pm 1\}} \int \partial_x R_{j_1} \partial_x N_{j_1 j_2}(\psi_c, R_{j_2}) \, dx + \varepsilon \mathcal{O} \left( \| R_{j_1} \|^2_{H^1} + \| R_{j_2} \|^2_{H^1} \right). \]

For \((j_1, j_2) = (j, -j)\), using Cauchy–Schwarz and (32) yields

\[ \varepsilon \int \partial_x R_j \partial_x N_{j-j}(\psi_c, R_{-j}) \, dx = \varepsilon \mathcal{O}(\| R_{j-1} \|^2_{H^1} \| R_j \|^2_{H^1}). \]

For \((j_1, j_2) = (j, j)\), there is up to one additional derivative falling on \(\partial_x R_j\). Using Leibniz’s rule and (31), we get

\[ \varepsilon \int \partial_x R_j \partial_x N_{j+j}(\psi_c, R_j) \, dx = \varepsilon \int \partial_x R_j N_{j+j}(\psi_c, \partial_x R_j) \, dx + \varepsilon \mathcal{O}(\| R_j \|^2_{H^1}). \]

Lemma 4.6 gives us

\[ \varepsilon \int \partial_x R_j N_{j+j}(\psi_c, \partial_x R_j) \, dx = \frac{1}{2} \varepsilon \int \partial_x R_j \left( N_{j+j}(\psi_c, \partial_x R_j) + N^*_{j+j}(\psi_c, \partial_x R_j) \right) \, dx \]

where

\[ \hat{N}^*_{j,j}(h, f)(k) := \int \hat{n}_{j,j}(-m, k-m, -k) \hat{h}(k-m) \hat{f}(m) \, dm. \]

Due to the skew symmetry of \(\rho\) and \(\omega\), we have for \(|k| \to \infty\):

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

Using Taylor to expand \(\rho(k)\) in the point \(m\) and exploiting (38), (39), and (40) yields

\[ n_{jj}(k, k-m, m) + n_{jj}(-m, k-m, -k) \]

\[ = \frac{\rho'(m)(k-m) + \mathcal{O}(\rho''(m))}{\omega(k) - \omega(m) + j \omega(k-m)} \chi_c(k-m) = \mathcal{O}(\chi_c(k-m)) \text{ for } |k| \to \infty. \]

Due to (6), (7), and (8). With Cauchy–Schwarz, the Plancherel theorem, and Young’s inequality, one now obtains

\[ \varepsilon \int \partial_x R_j N_{j+j}(\psi_c, \partial_x R_j) \, dx = \varepsilon \mathcal{O}(\| R_j \|^2_{H^1}) \]

such that the statement follows with Lemma 4.5.

\[ \square \]

### 4.3 Closing the energy estimates

For the Gronwall argument we are aiming for, the evolution of the energy has to be estimated against the energy itself. More precisely, we need to estimate the evolution of \(E_\varepsilon\) against terms involving no higher Sobolev norm of the error than the \(H^s\)-norm.

**Lemma 4.8.** For \(\ell \geq 1\), we have

\[ \partial_t E_\varepsilon = \sum_{i=0}^4 l_i + \varepsilon^2 \mathcal{O}(E_\varepsilon + 1), \]

(51)
where

\[
I_0 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \int \partial_x^\varepsilon R_{j_1} i \rho \partial_x^\varepsilon \theta^{-1} (R_Q \theta R_{j_1}) \, dx,
\]

\[
I_1 + I_2 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \left( \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} (R_Q \theta R_{j_1}) \, dx 
+ \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, \theta R_{j_1}) \, dx \right),
\]

\[
I_3 + I_4 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \left( \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} (R_Q \theta R_{j_1}) \, dx 
- \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, \theta R_{j_1}) \, dx \right),
\]

\[
R_Q = (\psi_{q_{-1}} + \psi_{q_1}) + \frac{1}{2} \varepsilon^{\beta - 2} (\theta R_{-1} + \theta R_1).
\]

**Proof.** Using the error Equations (22) and exploiting \( R_Q = \psi_c + \varepsilon R_Q \), we get

\[
\partial_t E_\varepsilon = \sum_{j_1 \in \{\pm 1\}} j_1 \int \partial_x^\varepsilon R_{j_1} i \omega \partial_x^\varepsilon R_{j_1} \, dx
+ \varepsilon \sum_{j_1,j_2 \in \{\pm 1\}} \left( j_1 \int \partial_x^\varepsilon R_{j_1} i \rho \partial_x^\varepsilon \theta^{-1} (\psi_c, \theta R_{j_1}) \, dx 
+ j_1 \int i \omega \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, \theta R_{j_1}) \, dx 
+ j_2 \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, i \omega R_{j_1}) \, dx 
- \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (i \omega \psi_c, R_{j_1}) \, dx 
+ \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\partial \psi_c + i \omega \psi_c, R_{j_1}) \, dx \right)
+ \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \int \partial_x^\varepsilon R_{j_1} i \rho \partial_x^\varepsilon \theta^{-1} (R_Q \theta R_{j_1}) \, dx 
+ \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} \left( j_1 \int i \omega \partial_x^\varepsilon \theta^{-1} (R_Q \theta R_{j_1}) \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, R_{j_1}) \, dx 
+ j_2 \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, i \rho \theta^{-1} (R_Q \theta R_{j_1})) \, dx \right)
+ \sum_{j_1 \in \{\pm 1\}} \int \partial_x^\varepsilon R_{j_1} e^{-\beta} \partial_x^\varepsilon \theta^{-1} \text{Res}_{j_1} (\varepsilon \Psi) \, dx
+ \varepsilon \sum_{j_1,j_2 \in \{\pm 1\}} \left( \int e^{-\beta} \partial_x^\varepsilon \theta^{-1} \text{Res}_{j_1} (\varepsilon \Psi) \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, R_{j_1}) \, dx 
+ \int \partial_x^\varepsilon R_{j_1} \partial_x^\varepsilon \theta^{-1} N_{j_1,j_2} (\psi_c, e^{-\beta} \theta^{-1} \text{Res}_{j_2} (\varepsilon \Psi)) \, dx \right).
\]
Exploiting the skew symmetry of $i \omega$ in the third integral and then using (29), we get

$$
\partial_t E_\epsilon = \epsilon \sum_{j_1, j_2 \in \{ \pm 1 \}} \left( \int_\mathbb{R} \partial_x R_{j_1} \partial_\theta \theta^{-1} (\psi_c(\theta - \theta_{\epsilon, \infty}) R_{j_2}) \, dx 
+ \int_\mathbb{R} \partial_x R_{j_1} \partial_\xi \xi^{-1} N_{j_1 j_2} (\partial_\xi \psi_c + i \epsilon \psi_c, R_{j_2}) \, dx \right) 
+ \sum_{i=0}^{4} I_i + \sum_{j_i \in \{ \pm 1 \}} \int_\mathbb{R} \partial_x R_{j_1} \epsilon^{-\beta} \partial_\xi \xi^{-1} \text{Res}_{u_1} (\epsilon \Psi) \, dx 
+ \epsilon \sum_{j_1, j_2 \in \{ \pm 1 \}} \left( \int_\mathbb{R} \epsilon^{-\beta} \partial_\xi \xi^{-1} \text{Res}_{u_1} (\epsilon \Psi) \partial_\xi \xi^{-1} N_{j_1 j_2} (\psi_c, R_{j_2}) \, dx 
+ \int_\mathbb{R} \partial_x R_{j_1} \partial_\xi \xi^{-1} N_{j_1 j_2} (\psi_c, \epsilon^{-\beta} \partial_\xi \xi^{-1} \text{Res}_{u_1} (\epsilon \Psi)) \, dx \right).
$$

We now show that all remaining terms can be estimated against $\epsilon^2 \mathcal{O}(E_\epsilon + 1)$. Thereby, we will especially take advantage of corollary 4.7 and (18).

Using (50), Cauchy–Schwarz and the fact that $\hat{\theta} - \hat{\theta}_{\epsilon, \infty} = \mathcal{O}(\epsilon)$ has compact support, the first integral is bounded by $\epsilon^2 \mathcal{O}(E_\epsilon)$.

The second integral in the above evolution equality is $\epsilon^3 \mathcal{O}(E_\epsilon)$ due to the estimate (19). In order to see this, we first use (50), then we proceed as in the proof of (4.7) in order to estimate without losing regularity.

The last three integrals are $\epsilon^2 \mathcal{O}(E_\epsilon + 1)$ due to (16). To see this, we use first (50), then integration by parts to shift some derivatives away from $R_{\pm 1}$, and finally Cauchy–Schwarz together with (31) and (32). We also exploit the estimate $\sqrt{x} \leq |x| + 1$.\]

\[\Box\]

The terms $I_0 - I_4$ contain integrals, in which too many additional derivatives are falling on the error functions such that a $H^r$-energy estimate is impossible for $\deg^*(\rho) > 0$.

In the following, we will assume $\epsilon_0$ to be chosen such small that

$$
\epsilon E_\epsilon \leq 1, \tag{53}
$$

for $0 < \epsilon < \epsilon_0$. This assumption is possible since there is some $T(\epsilon) > 0$ such that the $H^r$-norms of $R_{\pm 1}(t)$ and $R_{1}(t)$ can be uniformly bounded for $0 \leq t \leq T(\epsilon)$. When the energy estimates do close, Gronwall’s inequality provides $T(\epsilon) \geq \epsilon_0 \epsilon^{-2}$.

**Lemma 4.9.** Let $0 < \epsilon < \epsilon_0$ and $r \geq \lceil \deg(\omega) \rceil + \lceil \deg^*(\rho) \rceil + 1$.

Let $\gamma$ be a pseudo-differential operator given by its symbol in Fourier space such that

$$
\deg^*(\gamma^{(0)}) \leq \deg^*(\gamma^{(r-1)}) - 1, \tag{54}
$$

as long as $\gamma^{(0)} \neq 0$. We assume $\gamma$ to be independent of $\epsilon$ and to be either given by $\gamma = i \sigma$ or $\gamma = \nu$, for some odd function $\sigma \in C^{\lceil \deg(\sigma) \rceil}([\mathbb{R}, \mathbb{R}])$ with $\deg^*(\sigma) \leq \deg(\omega)$ or some even function $\nu \in C^{\lceil \deg^*(\nu) \rceil}([\mathbb{R}, \mathbb{R}])$ with $\deg^*(\nu) \leq \deg(\omega)$.

Let $h$ be a sum consisting of products of $\epsilon, \psi_c, \psi_{q_{\pm 1}}, R_{\pm 1}$ and their derivatives, for which

$$
\|h\|_{H^{\lceil \deg^*(\nu) \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_\xi h\|_{H^{\lceil \deg^*(\nu) \rceil - 1}} = \mathcal{O}(\epsilon^{-1/2}),
\|h\|_{C^{\lceil \deg^*(\nu) \rceil - 1}} + \|\partial_\xi h\|_{C^{\lceil \deg^*(\nu) \rceil - 1}} = \mathcal{O}(1). \tag{55}
$$

Let $g$ be a sum consisting of products of $\epsilon, \psi_c, \psi_{q_{\pm 1}}, R_{\pm 1}$ and their derivatives, for which $\|\partial_\xi^{-1} g\|_{\infty} = \mathcal{O}(\epsilon^{-1})$,

$$
\|g\|_{H^{\lceil \deg^*(\nu) \rceil + \lceil \deg(\omega) \rceil}} + \|\partial_\xi g\|_{H^{\lceil \deg^*(\nu) \rceil}} = \mathcal{O}(\epsilon^{-1/2}),
\|g\|_{C^{\lceil \deg^*(\nu) \rceil - 1}} + \|\partial_\xi^{-1} g\|_{C^{\lceil \deg^*(\nu) \rceil}} = \mathcal{O}(1). \tag{56}
$$
If $\gamma = i\sigma$ set $f = h$, and if $\gamma = \nu$ set $f = g$. Suppose

$$\|f\|_{H^{\lfloor \deg(\nu) \rfloor + \lfloor \deg(\sigma) \rfloor}} = \mathcal{O}(1) \quad \text{or} \quad \|\hat{f}\|_{L^1(\lfloor \deg(\nu) \rfloor + \lfloor \deg(\sigma) \rfloor)} = \mathcal{O}(1),$$

(57)

then there exists an $\epsilon_0 > 0$ such that for all $j_1, j_2 \in \{\pm 1\}$ there is an expression $D$ with

$$\epsilon^2 D = \epsilon \mathcal{O}(\mathcal{E}_\varepsilon)$$

and

$$\epsilon^2 \int_R \gamma \frac{\partial_x^\gamma R_{j_1}}{\partial_x^\gamma R_{j_2}} f \, dx = \epsilon^2 \partial_t D + \epsilon^2 \mathcal{O}(\mathcal{E}_\varepsilon + 1).$$

(58)

**Remark 4.10.** To make a choice for $\partial_x^{-1}$, we set

$$\partial_x^{-1} g := \int_{-\infty}^{(\cdot)} g \, dx.$$  

(59)

We will will show the proof of this lemma after the following corollary.

**Corollary 4.11.** Let $\ell' \geq \lfloor \deg(\omega) \rfloor + \lfloor \deg(\rho) \rfloor + 1$.

For $\epsilon_0$ sufficiently small and $0 < \varepsilon < \epsilon_0$, there exists a functional $\tilde{\mathcal{E}}_{\varepsilon}$ and some constants $c, C > 0$ such that

$$(\|R_{-1}\|_{H'\ell'} + \|R_1\|_{H'\ell'})^2 \leq c\tilde{\mathcal{E}}_{\varepsilon} \leq C(\|R_{-1}\|_{H'\ell'} + \|R_1\|_{H'\ell'})^2$$

(60)

and

$$\partial_t \tilde{\mathcal{E}}_{\varepsilon} \leq \epsilon^2 \mathcal{O}(\tilde{\mathcal{E}}_{\varepsilon} + 1).$$

**Proof.** According to lemmas 4.4 and 4.8, we have

$$\partial_t \mathcal{E}_\varepsilon = \sum_{i=0}^4 I_i + \epsilon^2 \mathcal{O}(\mathcal{E}_\varepsilon + 1).$$

First, we analyze the term $I_0$.

For $N \in \mathbb{N}$ and $\ell' \geq 2N + 1$, Leibniz's rule yields

$$\partial_x^{\ell'} (R_{\psi} \vartheta(R_1 + R_{-1})) = \sum_{n=0}^{N} \binom{\ell'}{n} \partial_x^{\ell'} (\psi + \varepsilon^{\beta-1} \vartheta(R_1 + R_{-1})) \partial_x^{\ell' - n} \vartheta(R_1 + R_{-1})$$

$$+ \sum_{n=N+1}^{\ell'-N-1} \binom{\ell'}{n} \partial_x^{\ell'} R_{\psi} \partial_x^{\ell' - n} \vartheta(R_1 + R_{-1})$$

$$+ \sum_{n=-N}^{\ell'} \binom{\ell'}{n} \partial_x^{\ell'} \psi \partial_x^{\ell' - n} \vartheta(R_1 + R_{-1}).$$

(61)

Only the terms of the first sum are problematic. A similar result is obtained when $R_{\psi}$ is replaced by $R_Q$. So, to easier keep track of those terms containing the highest derivatives of the error, we here introduce the notations

$$\tilde{\psi} := \psi + \varepsilon^{\beta-1} \vartheta(R_1 + R_{-1}), \quad \tilde{R}_Q := (\psi_{q_{-1}} + \psi_{q_1}) + \varepsilon^{\beta-2} \vartheta(R_{-1} + R_1).$$

(62)
Using (50), the skew symmetry of $i\rho$, Leibniz’s rule and (49), we get

$$I_0 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \int_{\mathbb{R}} \partial_x^\varepsilon R_{j_1} i\rho \partial_x^\varepsilon \theta^{-1}(R_{Q}\theta R_{j_1}) \, dx$$

$$= \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} \sum_{n=0}^N \left( \frac{\varepsilon}{n} \right) \int_{\mathbb{R}} \partial_x R_{j_1} \partial_x^- R_{j_1} \partial_x R_{Q} \, dx + \varepsilon^2 \mathcal{O}(\varepsilon) + 1), \quad (63)$$

where $N := \lfloor \text{deg}(\rho) \rfloor - 1$. With integration by parts and Lemma 4.9, one obtains

$$I_0 = \varepsilon^2 \partial_1 D_0 + \varepsilon^2 \mathcal{O}(\varepsilon)$$

for some $D_0$ with $\varepsilon^2 D_0 = \mathcal{O}(\varepsilon)$. Now, we analyze the term $I_1 + I_2$.

Using (50) and (31), we get

$$I_1 + I_2 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \left( \int_{\mathbb{R}} \partial_x R_{j_1} i\rho \partial_x \theta^{-1}(R_{Q}\theta R_{j_1}) \partial_x N_{j_1,j_1}(\psi_c, R_{j_1}) \, dx \right.$$

$$\left. + \int_{\mathbb{R}} \partial_x R_{j_1} \partial_x N_{j_1,j_1}(\psi_c, i\rho \theta^{-1}(R_{Q}\theta R_{j_1})) \, dx \right) + \varepsilon^2 \mathcal{O}(\varepsilon) + \varepsilon^{3/2} \mathcal{O}(\varepsilon^{3/2}). \quad (64)$$

Due to (31), applying Leibniz’s rule gives

$$I_1 + I_2 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \left( \int_{\mathbb{R}} \partial_x R_{j_1} i\rho \partial_x \theta^{-1}(R_{Q}\theta R_{j_1}) N_{j_1,j_1}(\psi_c, \partial_x R_{j_1}) \, dx \right.$$

$$\left. + \sum_{m=1}^{\lfloor \text{deg}(\rho) \rfloor} \left( \frac{\varepsilon}{m} \right) \int_{\mathbb{R}} \partial_x R_{j_1} i\rho \partial_x \theta^{-1}(R_{Q}\theta R_{j_1}) N_{j_1,j_1}(\partial_x^m \psi_c, \partial_x^{-m} R_{j_1}) \, dx \right.$$

$$\left. + \sum_{m=1}^{\lfloor \text{deg}(\rho) \rfloor} \left( \frac{\varepsilon}{m} \right) \int_{\mathbb{R}} \partial_x R_{j_1} N_{j_1,j_1}(\psi_c, i\rho \partial_x \theta^{-1}(R_{Q}\theta R_{j_1})) \, dx \right.$$

$$\left. + \sum_{m=1}^{\lfloor \text{deg}(\rho) \rfloor} \left( \frac{\varepsilon}{m} \right) \int_{\mathbb{R}} \partial_x R_{j_1} N_{j_1,j_1}(\partial_x^m \psi_c, i\rho \partial_x \theta^{-1}(R_{Q}\theta R_{j_1})) \, dx \right) + \varepsilon^2 \mathcal{O}(\varepsilon) + \varepsilon^{3/2} \mathcal{O}(\varepsilon^{3/2}).$$

By using Lemma 4.6, the skew symmetry of $i\rho$ and (50) we get

$$I_1 + I_2 = \varepsilon^2 \sum_{j_1,j_2 \in \{\pm 1\}} j_1 \left( \int_{\mathbb{R}} \partial_x^\varepsilon (R_{Q}\theta R_{j_1}) i\rho \left( N_{j_1,j_1}(\psi_c, \partial_x^\varepsilon R_{j_1}) + N_{j_1,j_1}(\psi_c, \partial_x^\varepsilon R_{j_1}) \right) \, dx \right.$$

$$\left. + \sum_{m=1}^{\lfloor \text{deg}(\rho) \rfloor} \left( \frac{\varepsilon}{m} \right) \int_{\mathbb{R}} \partial_x^\varepsilon (R_{Q}\theta R_{j_1}) i\rho N_{j_1,j_1}(\partial_x^m \psi_c, \partial_x^{-m} R_{j_1}) \, dx \right.$$

$$\left. + \sum_{m=1}^{\lfloor \text{deg}(\rho) \rfloor} \left( \frac{\varepsilon}{m} \right) \int_{\mathbb{R}} \partial_x^{-m}(R_{Q}\theta R_{j_1}) i\rho N_{j_1,j_1}(\partial_x^m \psi_c, \partial_x^m R_{j_1}) \, dx \right) + \varepsilon^2 \mathcal{O}(\varepsilon) + \varepsilon^{3/2} \mathcal{O}(\varepsilon^{3/2}),$$

where

$$\tilde{N}_{j_1,j_2}^{\varepsilon}(\psi_c, f)(k) := \int_{\mathbb{R}} n_{j_1,j_2}(-m, k - m, -k)\tilde{\psi}_c(k - m)\tilde{f}(m) \, dm.$$
If we now look at
\[ i\rho(k) \left( n_{j}((k, k - m, m)) + n_{j}((-m, k - m, -k)) \right) \]
\[ = i\rho(k)(\rho(k) - \rho(m)) \frac{\chi(k - m)}{\omega(k) - \omega(m) + j \omega(k - m)} \]
(for \( |k| \to \infty \))
and use Taylor's theorem, the same cancelation as in the proof of corollary 4.7 occurs. By now exploiting (38) or respectively (39) or (40) (with integration by parts for the third term), we get
\[ I_1 + I_2 = \epsilon^2 \sum_{j_1, j_2 \in \{\pm 1\}} j_1 \sum_{n=1}^{N} \int_{\mathbb{R}} \delta_x^n (R_{\omega} \theta R_{j_1}) \beta_n \psi_c \alpha_n \delta_x^n R_{j_2} \, dx + \epsilon^2 O(\mathcal{E}_\epsilon + 1) \]
for some \( N \in \mathbb{N} \) and some pseudo-differential operators \( \beta_n \) and \( \alpha_n \), where \( \alpha_n \) is either skew symmetric with \( \deg^*(\alpha_n) \leq \deg^*(\rho) \) or symmetric with \( \deg^*(\alpha_n) \leq \deg^*(\rho) - 1 \). With the help of (61), (49) and integration by parts, we can now apply Lemma 4.9 to obtain
\[ I_1 + I_2 = \epsilon^2 \beta_1 D_{1,2} + \epsilon^2 O(\mathcal{E}_\epsilon + 1) \]
for some \( D_{1,2} \) with \( \epsilon^2 D_{1,2} = O(\mathcal{E}_\epsilon) \). To apply Lemma 4.9, one splits
\[ \delta_x^{-1} \overline{R_{\omega} \delta_x^m \beta_n \psi_c} = \delta_x^{-1} (\psi_c + \epsilon \overline{R_Q}) \delta_x^m \beta_n \psi_c \]
such that \( \|f_1\|_{L^1(p)} = \mathcal{O}(1) \) and \( \|f_2\|_{H^p} = \mathcal{O}(1) \). The estimate \( \|\delta_x^{-1} f_1\|_\infty = \mathcal{O}(1) \) is obtained by
\[ \|\delta_x^{-1} f_1\|_\infty = \|\delta_x^{-1} (\delta_x^m \psi_c \delta_x^m \beta_n \psi_c)\|_\infty \leq \int_{\mathbb{R}} |\delta_x^m \psi_c \delta_x^m \beta_n \psi_c| \, dx \]
\[ \leq \|\delta_x^m \psi_c\|_{L^2} \|\delta_x^m \beta_n \psi_c\|_{L^2}. \]
To obtain the estimate \( \|\partial_i \delta_x^{-1} f_1\|_\infty = \mathcal{O}(1) \), one has to proceed more carefully. To obtain this estimate, one splits \( \psi_c = \psi_{-1} + \psi_1 \) with \( \psi_{\pm 1} \) as in (14). Then one exploits the fact that the products \( \psi_j \psi_j \) are strictly concentrated around \( k = \pm 2k_0 \) in Fourier space and \( \partial_i (\psi_j \psi_{-j}) = \mathcal{O}(\epsilon) \) such that \( \|f^i \partial_i \delta_x^{-1} (\psi_j \psi_{-j})\|_{L^1(p)} = \mathcal{O}(1) \) and \( \|\partial_i \delta_x^{-1} (\psi_j \psi_{-j})\|_\infty = \mathcal{O}(1) \) can be obtained. The other estimates are straightforward.

Now, we analyze the term \( I_3 + I_4 \).
Using (50) and (32), we have
\[ I_3 + I_4 = \epsilon^2 \sum_{j_1, j_2 \in \{\pm 1\}} j_1 \left( \int_{\mathbb{R}} \rho \delta_x^\epsilon \theta^{-1} (R_{\omega} \theta R_{j_1}) \delta_x^\epsilon N_{j_1, -j_1} (\psi_c, R_{-j_1}) \, dx \right. \]
\[ - \left. \int_{\mathbb{R}} \delta_x^\epsilon R_{j_1} \delta_x^\epsilon N_{j_1, -j_1} (\psi_c, \rho \theta^{-1} (R_{\omega} \theta R_{j_1})) \, dx \right) + \epsilon^2 O(\mathcal{E}_\epsilon + \epsilon^\beta \mathcal{E}_\epsilon^{3/2}). \]  
(65)
According to Lemma 4.2 the \( N_{j_1, -j_1} (\delta_x^\epsilon \psi_c, \cdot) \) always map \( L^2(\mathbb{R}) \) on \( L^2(\mathbb{R}) \). With the help of (41), we can thus proceed as before for \( I_1 + I_2 \) after the cancelation was achieved. We apply Lemma 4.9 and obtain
\[ I_3 + I_4 = \epsilon^2 \partial_i D_{3,4} + \epsilon^2 O(\mathcal{E}_\epsilon + 1), \]
for some \( D_{3,4} \) with \( \epsilon^2 D_{3,4} = \mathcal{O}(\mathcal{E}_\epsilon) \).
Choosing \( \epsilon_0 \) small enough and summing up the results for \( I_0 - I_4 \), we can define a modified energy
\[ \tilde{\mathcal{E}}_\epsilon := \mathcal{E}_\epsilon - \epsilon^2 (D_0 + D_{1,2} + D_{3,4}), \]
such that

$$\partial_t \tilde{E}_\varepsilon \leq \varepsilon^2 (1 + E_\varepsilon).$$

Since $\tilde{E}_\varepsilon = E_\varepsilon + \varepsilon \mathcal{O}(E_\varepsilon)$, the statement follows with corollary 4.7.

For the proof of Lemma 4.9 we use the notation $[\gamma, f, g] := \gamma(fg) - f \gamma g$ for an operator $\gamma$ and functions $g$ and $f$. Further, we need the following lemma.

**Lemma 4.12.** Let $n \in \mathbb{N}$, and $\gamma$ be a function of $C^{n+1}(\mathbb{R})$ with $\deg^\ast(\gamma) \in \mathbb{R}$ for which

$$\deg^\ast(\gamma^{(l)}) \leq \deg^\ast(\gamma^{(l-1)}) - 1 \quad \text{for all} \quad 1 \leq l \leq n + 1. \quad (66)$$

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

$$[\gamma, g] f = \sum_{l=1}^{n} \frac{(-1)^l}{l!} \partial_x^l g \hat{\gamma}^{(l)} f + R(f, g). \quad (67)$$

For the remainder $R(f, g)$, given through

$$R(f, g) = \int_\mathbb{R} \left( \frac{c_{\gamma}}{n!} \right)^{n+1} \int_0^1 \gamma^{(n+1)} (m + (-m)x)(1-x) dx \hat{g}(-m) \hat{f}(m) dm,$$

we have the estimate

$$\|R(f, g)\|_{L^2} \leq \mathcal{O}(1) \|f\|_{F^{-1}} \left[ \left( 1 + |\cdot|^{2p/2} \partial^p g(\cdot) \right) \right] \|f\|_{H^p} \quad (68)$$

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

**Remark 4.13.** Estimate (68) implies $\|R(f, g)\|_{L^2} \leq \mathcal{O}(\|\partial_x^{n+1} g\|_{L^1(\mathbb{R})}) \|f\|_{H^p}$ and, with Sobolev’s embedding theorem $\|R(f, g)\|_{L^2} \leq \mathcal{O}(\|g\|_{H^{p+1}}) \|f\|_{H^p}$ for $q > 1/2$.

**Proof.** We have

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

Using Taylor, we get

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

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

where

$$r(k, k-m, m) = \frac{(k-m)^{n+1}}{n!} \int_0^1 \gamma^{(n+1)} (m + (k-m)x)(1-x) dx$$

$$\leq \frac{(k-m)^{n+1}}{n!} \max_{x \in [0,1]} \gamma^{(n+1)} (m + (k-m)x)$$

$$\leq \mathcal{O}(\|k-m\|^{n+1}) \left( 1 + 1 + |k-m|^{\deg(\gamma) - n-1} + 1 + |m|^{\deg(\gamma) - n-1} \right) .$$

We now get

$$\|R(f, g)\|_{L^2} \leq \int_\mathbb{R} |r(k, k-m, m)\hat{g}(k-m) \hat{f}(m)| dm \|_{L^2}$$

$$\leq \mathcal{O}(1) \int_\mathbb{R} \left[ (1 + |k-m|^{2p/2} \partial^p g(k-m)(1 + |m|^{2p/2} \hat{f}(m)| dm \right] \|_{L^2} .$$
with \( p = \max\{\deg^*(\gamma) - n - 1, 0\} \). With Plancherel's theorem, we obtain

\[
\| R(f, g) \|_{L^2} \leq \mathcal{O}(1) \left\| F^{-1} \left[ (1 + |\cdot|^2)^{p/2} \partial_x^p g(\cdot) \right] \right\|_{L^2} \\
\leq \mathcal{O}(1) \left\| F^{-1} \left[ (1 + |\cdot|^2)^{p/2} \partial_x^p g(\cdot) \right] \right\|_{\infty} \| f \|_{H^r}.
\]

\( \square \)

**Proof of Lemma 4.9.** If \( \deg^*(\gamma) \leq 0 \), the lemma is trivially true. So we will in the following assume \( \deg^*(\gamma) > 0 \). Since \( \deg(o) \geq \deg^*(\gamma) > 0 \), there exist some constants \( D_o, d_o > 0 \) such that \( |\omega(k)| \geq d_o > 0 \) for \( |k| \geq D_o \). For \( \gamma = \nu \), we can on top of that find \( D_o, d_o > 0 \) such that \( |\omega'(k)| \geq d_o > 0 \) for \( |k| \geq D_o \) due to \( \deg(o') \geq \deg^*(\nu) > 0 \).

There is some \( D \geq D_o \) and some function \( \tilde{\gamma} \in C^{[\deg^*(\gamma)]}(\mathbb{R}, \mathbb{R}) \) with (54) such that \( \tilde{\gamma}(k) = \gamma(k) \) for \( |k| \geq D \) and \( \tilde{\gamma}(k) = 0 \) for \( |k| \leq D_o \). Since

\[
\epsilon^2 \int_\mathbb{R} \gamma\partial_x^p R_{j_1} \partial_x^p R_{j_2} f \ dx = \epsilon^2 \int_\mathbb{R} \tilde{\gamma}\partial_x^p R_{j_1} \partial_x^p R_{j_2} f \ dx + \epsilon^2 \int_\mathbb{R} (\gamma - \tilde{\gamma})\partial_x^p R_{j_1} \partial_x^p R_{j_2} f \ dx
\]

we can in the following assume that we have \( \gamma(k) = 0 \) for \( |k| \leq D_o \). This makes the operators \( \frac{\partial_x}{\omega} \) and \( \frac{\partial_x}{\omega'} \) well-defined.

As a first step of the proof, we show

**Lemma 4.14.** There is an expression \( D \) with

\[
\epsilon^2 D = \epsilon \mathcal{O}(\mathcal{E}_{\epsilon}),
\]

and there exist some skew symmetric or symmetric operators \( \zeta_k \) and \( \gamma_k \) given by their symbol in Fourier space and independent of \( \epsilon \), \( m_{\zeta} = m_{\gamma} \left( \deg^*(\gamma) \right) \in \mathbb{N} \), functions \( f_k \) and \( p_k, q_k \in [-1, 1] \) such that

\[
\epsilon^2 \int_\mathbb{R} \gamma\partial_x^p R_{j_1} \partial_x^p R_{j_2} f \ dx = \epsilon^2 \partial_x D + \epsilon^2 \sum_{k=1}^{[\deg^*(\gamma) - 1]} \int_\mathbb{R} \zeta_k \partial_x^p R_{j_1} \partial_x^p R_{j_2} \partial_x^p f \ dx
\]

\[
+ \epsilon^2 \sum_{k=1}^{m_{\gamma}} \int_\mathbb{R} \gamma_k \partial_x^p R_{p_k} \partial_x^p R_{q_k} f_k \ dx + \epsilon^2 \mathcal{O}(\mathcal{E}_{\epsilon} + 1).
\]

Moreover, the functions \( \zeta_k \in C^{[\deg^*(\zeta_k)]}(\mathbb{R}, \mathbb{R}) \) and \( \gamma_k \in C^{[\deg^*(\gamma_k)]}(\mathbb{R}, \mathbb{R}) \) share the property (54). For the skew symmetric operators \( \gamma_k \) we have \( \deg^*(\gamma_k) \leq \deg^*(\gamma) \), while for the symmetric \( \gamma_k \) it is even \( \deg^*(\gamma_k) \leq \deg^*(\gamma) - 1 \). Furthermore,

\[
\deg^*(\zeta_k) \leq \deg^*(\gamma) - k.
\]

\[
\| F^{-1} \left[ (1 + |\cdot|^2)^{p/2} \partial_x^p f(\cdot) \right] \|_{\infty} = \mathcal{O}(1),
\]

\[
\| f_k \|_{H^{[\deg^*(\gamma_k)]} + [\deg(o)]} + \| \partial_x f_k \|_{H^{[\deg^*(\gamma_k)]}} \leq \epsilon C_1 \left( \| f \|_{H^{[\deg^*(\gamma_k)]} + [\deg(o)]} + \| \partial_x f \|_{H^{[\deg^*(\gamma_k)]}} \right)
\]

\[
+ \epsilon^1/2 C_2 \left( \| \partial_x^{-1} f \|_{C^{[\deg^*(\gamma_k)]}} + \| \partial_x \partial_x^{-1} f \|_{C^{[\deg^*(\gamma_k)]}} \right), \| \partial_x^{-1} f_k \|_{\infty} + \| \partial_x \partial_x^{-1} f_k \|_{C^{[\deg^*(\gamma_k)]}} \leq \epsilon^1/2 C_1 \left( \| f \|_{L^2} + \| \partial_x f \|_{L^2} \right) + \epsilon C_2 \left( \| \partial_x^{-1} f \|_{C^{[\deg^*(\gamma_k)]}} + \| \partial_x \partial_x^{-1} f \|_{C^{[\deg^*(\gamma_k)]}} \right),
\]

where the constants \( C_1, C_2 \) depend on \( \tilde{R}_\gamma, f, \gamma \) but are independent of \( \epsilon \). We set \( \epsilon C_2 := 0 \), when \( \gamma \) is skew symmetric, that is, when \( f = h \).
**a)** Handling the case $j_1 = -j_2$, i.e. integrals of the form

$$
\epsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx.
$$

(74)

By exploiting the skew symmetry of $io \omega$ and (22), we have

$$
\epsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx = \frac{1}{2} j \epsilon^2 \partial_i \int_{\mathbb{R}} \frac{\gamma}{io} \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx
$$

$$
- \frac{1}{2} \epsilon^2 \int_{\mathbb{R}} [io, f] \frac{\gamma}{io} \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx
$$

$$
- \frac{1}{2} \epsilon^3 \int_{\mathbb{R}} \frac{\gamma}{io} i \rho \partial_x^\epsilon \theta^{-1} (R_{\bar{\psi}} \theta(R_1 + R_{-1})) \partial_x^\epsilon R_{-j} f \, dx
$$

$$
+ \frac{1}{2} \epsilon^3 \int_{\mathbb{R}} \frac{\gamma}{io} \partial_x^\epsilon R_j i \rho \partial_x^\epsilon \theta^{-1} (R_{\bar{\psi}} \theta(R_1 + R_{-1})) f \, dx
$$

$$
- \frac{1}{2} j \epsilon^2 \int_{\mathbb{R}} \frac{\gamma}{io} \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} \partial_i f \, dx
$$

$$
- \frac{1}{2} j \epsilon^{2-\beta} \int_{\mathbb{R}} \gamma \partial_x^\epsilon \theta^{-1} \text{Res}_{\bar{\omega}}(\epsilon \Psi) \partial_x^\epsilon R_{-j} f \, dx
$$

$$
- \frac{1}{2} j \epsilon^{2-\beta} \int_{\mathbb{R}} \gamma \partial_x^\epsilon R_j \partial_x^\epsilon \theta^{-1} \text{Res}_{\bar{\omega}}(\epsilon \Psi) f \, dx.
$$

The first term is the time derivative of an integral, which can be estimated against $\epsilon^2 \mathcal{O}(\mathcal{E}_\epsilon)$ by using Cauchy–Schwarz. The last three integrals can be estimated against $\epsilon^2 \mathcal{O}(\mathcal{E}_\epsilon + 1)$ since $\| \partial_t f \|_\infty = \mathcal{O}(1)$ and due to (16).

For the second integral, applying (67) gives us

$$
- \frac{1}{2} \epsilon^2 \int_{\mathbb{R}} [io, f] \frac{\gamma}{io} \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx = -\frac{1}{2} \epsilon^2 \sum_{n=1}^{[\text{deg}(\nu)]-1} \frac{1}{n!} \int_{\mathbb{R}} (-i)^n \omega^{(n)} \frac{\gamma}{io} \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f \, dx
$$

$$
+ \mathcal{O}(\epsilon^2) \| R \left( \frac{\gamma}{io} \partial_x^\epsilon R_j, f \right) \|_{L^2} \| \partial_x^\epsilon R_{-j} \|_{L^2},
$$

where with (68) we can estimate $\| R \left( \frac{\gamma}{io} \partial_x^\epsilon R_j, f \right) \|_{L^2} \| \partial_x^\epsilon R_{-j} \|_{L^2} = \mathcal{O}(\mathcal{E}_\epsilon + 1)$.

The integrals in the third and the fourth place can be written as a sum of some $\epsilon^3 \mathcal{O}(\mathcal{E}_\epsilon + 1)$-terms and $m$ many integrals of the form

$$
\epsilon^2 \int_{\mathbb{R}} \gamma \partial_x^\epsilon R_j \partial_x^\epsilon R_{-j} f_k \, dx
$$

with $m$, $\gamma$, $f_k$, $p_k$, and $q_k$ just as in Lemma 4.14. One sees this by exploiting (50), Leibniz’s rule, (61), (67) and (49). Since we have by assumption

$$
\| \tilde{R}_\psi \|_{H^l(\text{deg}(\nu))} + \| \partial_t \tilde{R}_\psi \|_{H^{l+1}(\text{deg}(\nu))} \leq \epsilon^{-1/2} c_R,
$$

(75)

$$
\| \tilde{R}_\psi \|_{C^l(\text{deg}(\nu)+1)} + \| \partial_t \tilde{R}_\psi \|_{C^{l+1}(\text{deg}(\nu))} \leq c_R,
$$

for some $c_R \in \mathbb{R}$, straightforward estimates confirm that the functions having the form $f_k = \epsilon \partial_x^p \tilde{R}_\psi f$ with $p \geq 0$ indeed fulfill (72), (73) and hence (71).

**b)** Handling the case $j_1 = j_2$ and $(\gamma, f) = (i \sigma, h)$, that is, integrals of the form

$$
\epsilon^2 \int_{\mathbb{R}} i \sigma \partial_x^\epsilon R_j \partial_x^\epsilon R_j h \, dx.
$$

(76)
Since $i\sigma$ is skew symmetric and due to (67) and (68), we have

$$\epsilon^2 \int_{\mathbb{R}} i\sigma \partial_x^2 R_j \partial_x^2 R_j h \, dx = -\frac{1}{2} \epsilon^2 \int_{\mathbb{R}} [i\sigma, h] \partial_x^2 R_j \partial_x^2 R_j \, dx,$$

$$= \epsilon^2 \sum_{k=1}^{\deg^*(e)-1} \int_{\mathbb{R}} \zeta_k \partial_x^2 R_j \partial_x^2 R_j \partial_x^2 h \, dx + \epsilon^2 O(\mathcal{E}_\varepsilon + 1),$$

with $\zeta_k$ just as in Lemma 4.14.

c) Handling the case $j_1 = j_2$ and $(\gamma, f) = (u, g)$, that is, integrals of the form

$$\epsilon^2 \int_{\mathbb{R}} u \partial_x^2 R_j \partial_x^2 R_j g \, dx.$$  \hspace{1cm} (77)

By using (67), we can write

$$\epsilon^2 \int_{\mathbb{R}} u \partial_x^2 R_j \partial_x^2 R_j g \, dx = \epsilon^2 \int_{\mathbb{R}} i\omega, \partial_x^{-1} g \frac{d}{d\omega} \partial_x \partial_x^2 R_j \partial_x^2 R_j \, dx$$

$$+ \epsilon^2 \sum_{n=2}^{\deg^*(\omega)} \frac{(-1)^n}{(n)!} \int_{\mathbb{R}} \partial_x^{n+1} \omega \partial_x^{n+1} \partial_x \partial_x^2 R_j \partial_x^2 R_j \partial_x^{-1} g \, dx$$

$$+ O(\epsilon^2) \| R \left( \frac{d}{d\omega} \partial_x \partial_x^2 R_j, \partial_x^{-1} g \right) \|_{L^2} \| \partial_x^2 R_j \|_{L^2},$$

where with (68) we can estimate $\| R \left( \frac{d}{d\omega} \partial_x \partial_x^2 R_j, \partial_x^{-1} g \right) \|_{L^2} \| \partial_x^2 R_j \|_{L^2} = O(\mathcal{E}_\varepsilon + 1)$.

Now, the second term already has the desired form and the last term is $\epsilon^2 O(\mathcal{E}_\varepsilon + 1)$ such that we only have to look at the first term.

By exploiting the skew symmetry of $i\omega$ and (22) (and (16)), we have

$$\epsilon^2 \int_{\mathbb{R}} i\omega, \partial_x^{-1} g \frac{d}{d\omega} \partial_x^2 R_j \partial_x^2 R_j \, dx$$

$$= \epsilon^2 \int_{\mathbb{R}} i\omega \left( \partial_x^{-1} g \frac{d}{d\omega} \partial_x^2 R_j \right) \partial_x^2 R_j \partial_x^2 R_j \, dx - \epsilon^2 \int_{\mathbb{R}} \partial_x^{-1} g i\omega \frac{d}{d\omega} \partial_x^2 R_j \partial_x^2 R_j \, dx$$

$$= -j \epsilon \int_{\mathbb{R}} \frac{d}{d\omega} \partial_x^2 R_j \partial_x^2 R_j \partial_x^{-1} g \, dx$$

$$+ \epsilon^3 \int_{\mathbb{R}} \frac{d}{d\omega} \partial_x^2 R_j i\rho \partial_x \partial_x^2 \theta^{-1} \left( R_\psi \theta (R_{-1} + R_1) \right) \partial_x^2 R_j \partial_x^{-1} g \, dx$$

$$+ \epsilon^3 \int_{\mathbb{R}} \frac{d}{d\omega} \partial_x^2 R_j \delta \partial_x \partial_x^2 \theta^{-1} \left( R_\psi \delta (R_{-1} + R_1) \right) \partial_x^{-1} g \, dx$$

$$+ j \epsilon^2 \int_{\mathbb{R}} \frac{d}{d\omega} \partial_x^2 R_j \partial_x \partial_x^2 R_j \partial_x \partial_x^{-1} g \, dx$$

$$+ \epsilon^2 O(\mathcal{E}_\varepsilon + 1).$$
The last integral can be estimated against $\epsilon^2 \mathcal{O}(\mathcal{E}_\epsilon + 1)$ since $\|\partial_x \partial_x^{-1} g\|_\infty = \mathcal{O}(1)$. Due to (50), the skew symmetry of $i\rho$ and the symmetry of $\omega'$ and $\nu$, we get

$$
\epsilon^2 \int_{\mathbb{R}} \left[ [\omega, \partial_x^{-1} g] \frac{\partial}{\partial \omega'} \partial_x^\prime \partial_x R_j \partial_x R_j \right] dx = -j \epsilon^2 \partial_j \int_{\mathbb{R}} \partial_x \partial_x^\prime \partial_x R_j \partial_x R_j \partial_x^{-1} g dx
- 2\epsilon^3 \int_{\mathbb{R}} [i\rho \frac{\partial}{\partial \omega'}, \partial_x^{-1} g] \partial_x^\prime \partial_x R_j \partial_x R_j (R_{\psi}(\theta(R-1) + R_1)) \partial_x^{-1} g dx
- \epsilon^3 \int_{\mathbb{R}} [i\rho, \partial_x^{-1} g] \frac{\partial}{\partial \omega'} \partial_x^\prime \partial_x R_j \partial_x R_j (R_{\psi}(\theta(R-1) + R_1)) \partial_x^{-1} g dx
+ \epsilon^2 \mathcal{O}(\mathcal{E}_\epsilon + 1).
$$

(78)

The first term is a time derivative of an expression $\epsilon^2 \mathcal{D}$, which can be estimated against $\epsilon \mathcal{O}(\mathcal{E}_\epsilon)$ since $\epsilon \|\partial_x^{-1} g\|_\infty = \mathcal{O}(1)$. By using (67) and Leibniz's rule, we can write the third and the fourth integral as a sum of some $\epsilon^3 \mathcal{O}(\mathcal{E}_\epsilon + 1)$ terms and integrals of the form

$$
\epsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x \partial_x^\prime \partial_x R_k \partial_x R_k f_k dx
$$

with $\gamma_k, f_k, p_k, q_k$ just as in Lemma 4.14. Making straightforward estimates by using (75) shows that the functions of the form $f_k = \epsilon \partial_x^n \bar{R}_{\psi} \partial_x^m g$ with $n, m \geq 0$ here fulfill (71), (72), and (73).

What remains to be analyzed is the second term of the right-hand side from Equation (78). Using Leibniz's rule and afterwards (61) and (49), we obtain

$$
-2\epsilon^3 \int_{\mathbb{R}} [i\rho \frac{\partial}{\partial \omega'}, \partial_x^{-1} g] \partial_x^\prime \partial_x R_j \partial_x R_j (R_{\psi}(\theta(R-1) + R_1)) \partial_x^{-1} g dx
= -2\epsilon^3 \int_{\mathbb{R}} [i\rho \frac{\partial}{\partial \omega'}, \partial_x^{-1} g] \partial_x^\prime \partial_x R_j \partial_x R_j (R_{\psi}(\theta(R-1) + R_1)) \partial_x^{-1} g dx
- 2\epsilon^3 \sum_{m=1}^M \left( \begin{array}{c} \ell \\ m \end{array} \right) \int_{\mathbb{R}} [i\rho \frac{\partial}{\partial \omega'}, \partial_x^{-1} g] \partial_x^\prime \partial_x R_j \partial_x^{-m}(R-1 + R_1) \partial_x^m \bar{R}_{\psi} \partial_x^{-1} g dx
+ \epsilon^3 \mathcal{O}(\mathcal{E}_\epsilon + 1),
$$

where $M := [\text{deg}(\rho \omega) - \text{deg}(\omega')] - 1$.

The second term here consists (after integration by parts) of integrals having the form

$$
\epsilon^2 \int_{\mathbb{R}} \gamma_k \partial_x \partial_x^\prime \partial_x R_k \partial_x R_k f_k dx
$$

with $\gamma_k, f_k, p_k, q_k$ just as in Lemma 4.14. The functions of the form $f_k = \epsilon \partial_x^n \bar{R}_{\psi} \partial_x^m g$ with $m > 0$ fulfill (72) and (73): For $m \geq 1$ and $n$ as required, one gets

$$
\|\epsilon \partial_x^n \bar{R}_{\psi} \partial_x^{-1} g\|_{H^n} \leq \epsilon \|\partial_x^n \bar{R}_{\psi}\|_{H^n} \|\partial_x^{-1} g\|_{C^n},
\|\epsilon \partial_x^n \bar{R}_{\psi} \partial_x^{-1} g\|_{C^n} \leq \epsilon \|\partial_x^n \bar{R}_{\psi}\|_{C^n} \|\partial_x^{-1} g\|_{C^n}.
$$

The estimates for $\|\epsilon \partial_x (\partial_x^n \bar{R}_{\psi} \partial_x^{-1} g)\|_{H^n}$ and $\|\epsilon \partial_x (\partial_x^n \bar{R}_{\psi} \partial_x^{-1} g)\|_{C^n}$ are similarly straightforward. Concerning the other estimates, we estimate

$$
\|\epsilon \partial_x^{-1} (\partial_x^n \bar{R}_{\psi} \partial_x^{-1} g)\|_{\infty} = \epsilon \|\partial_x^{n-1} \bar{R}_{\psi} \partial_x^{-1} g - \partial_x^{-1} (\partial_x^{n-1} \bar{R}_{\psi} g)\|_{\infty}
\leq \epsilon \|\partial_x^{n-1} \bar{R}_{\psi}\|_{\infty} \|\partial_x^{-1} g\|_{\infty} + \epsilon \|\partial_x^{n-1} \bar{R}_{\psi}\|_{L^2} \|g\|_{L^2},
$$

and similar $\|\epsilon \partial_x \partial_x^{-1} (\partial_x^n \bar{R}_{\psi} \partial_x^{-1} g)\|_{\infty}$. The estimate $\|\epsilon \partial_x \partial_x^n \bar{R}_{\psi} \partial_x^{-1} g\|_{H^n} = \mathcal{O}(1)$ is not implied by the above estimates such that (71) is not trivially obtained. We confirm (71) by exploiting that the supremum over all $x \in \mathbb{R}$ is the same as the supre-
mum over all \((\epsilon^{-1}x) \in \mathbb{R}\) such that the loss of \(\epsilon\)-powers caused by the slow spatial scale of the NLS present in estimate (75) can here be avoided.

Thus, we now only have to examine the term

\[
-2\epsilon^3 \int_\mathbb{R} i \rho \frac{D}{\omega'} \partial_x^2 R_j \partial_x^2 (R_{-1} + R_1) R_w \partial_x^{-1} g \, dx = -2\epsilon^3 \int_\mathbb{R} i \rho \frac{D}{\omega'} \partial_x^2 R_j \partial_x^2 R_{-1} R_w \partial_x^{-1} g \, dx
\]

For the above first integral we can proceed as in paragraph a) and for the second integral as in paragraph b). The required estimates for the function \(\hat{h} = \epsilon \hat{R}_w \partial_x^{-1} g\) work as above; however, one has to be aware that \(\hat{h}\) only meets the conditions for a) and b) since \(\deg^a \left( \frac{\partial}{\omega'} \right) \leq \deg^a (\omega) + 1\) and \(\deg^a \left( \frac{\partial}{\omega'} \right) \leq \deg^a (\rho) \leq \deg^a (\omega)\) due to (6). Also note that for a) and b) no estimate for \(\partial_x^{-1} \hat{h}\) is needed.

With this Lemma 4.14 is proven, and we can move on with the proof of Lemma 4.9.

When Lemma 4.14 is applied:

- the number of additional derivatives falling on \(\partial_x^2 R_{-1}\) or \(\partial_x^2 R_1\) does not increase,
- we do not generate new problematic terms for which Lemma 4.14 cannot be applied,
- the new emerging problematic integrals get much smaller in size,
- the number of the emerging integrals only depends on \(\deg^a (\rho)\) and \(\deg(\omega)\).

We exploit this and use Lemma 4.14 repeatedly, that is, we apply it again to every integral on the right hand side of (69) that contains additional derivatives. So we can use (69) and exploit (70) in order to get

\[
\epsilon^2 \int_\mathbb{R} \gamma \partial_x^2 R_j \partial_x^2 R_{j_2} f \, dx = \epsilon^2 \partial_t \tilde{D} + \epsilon^2 \sum_{k=1}^{m} \gamma_k \partial_x^2 R_{p_k} \partial_x^2 R_{q_k} f_k \, dx + \epsilon^2 \mathcal{O}(\mathcal{E}_r + 1),
\]

where \(\tilde{D} = \epsilon \mathcal{O}(\mathcal{E}_r)\), \(\tilde{m} \leq m = m(\deg(\omega)) \in \mathbb{N}\) and due to (72) and (73):

\[
|| \tilde{f_k} ||_{H^{\left( \deg^a (\omega_{j_2}) \right)}} + || \partial_t \tilde{f_k} ||_{H^{\left( \deg^a (\omega_{j_2}) \right)}} \leq \epsilon^{-1/2} C_f,
\]

\[
|| \partial_x^{-1} \tilde{f_k} ||_{C_{\left( \deg^a (\omega_{j_2}) \right)}} + || \partial_t \partial_x^{-1} \tilde{f_k} ||_{C_{\left( \deg^a (\omega_{j_2}) \right)}} \leq C_f,
\]

for some constant \(C_f = C_f (\hat{R}_w, f, \gamma) > 1\).

By using (69) and exploiting (70) again for every integral on the above right-hand side, we can obtain an expression \(\epsilon^2 \tilde{D} = \epsilon \mathcal{O}(\mathcal{E}_r)\) such that we have

\[
\epsilon^2 \int_\mathbb{R} \gamma \partial_x^2 R_j \partial_x^2 R_{j_2} f \, dx = \epsilon^2 \partial_t \tilde{D} + \epsilon^2 \sum_{k=1}^{m} \gamma_k \partial_x^2 R_{p_k} \partial_x^2 R_{q_k} f_k \, dx + \epsilon^2 \mathcal{O}(\mathcal{E}_r + 1),
\]

where \(\tilde{m} \leq m^2\),

\[
|| \tilde{f_k} ||_{H^{\left( \deg^a (\omega_{j_2}) \right)}} + || \partial_t \tilde{f_k} ||_{H^{\left( \deg^a (\omega_{j_2}) \right)}} \leq \epsilon^{1/2} C^2_f,
\]

\[
|| \partial_x^{-1} \tilde{f_k} ||_{C_{\left( \deg^a (\omega_{j_2}) \right)}} + || \partial_t \partial_x^{-1} \tilde{f_k} ||_{C_{\left( \deg^a (\omega_{j_2}) \right)}} \leq C^2_f,
\]

By repeating the last step \(N + 1\) times, we get

\[
\epsilon^2 \int_\mathbb{R} \gamma \partial_x^2 R_j \partial_x^2 R_{j_2} f \, dx = \epsilon^2 \sum_{p=0}^{N} \epsilon^{p/2} \partial_t D_p + \epsilon^2 \sum_{k=1}^{m_N} \gamma_{k,N} \partial_x^2 R_{p_k} \partial_x^2 R_{q_k} f_{k,N} \, dx + \epsilon^2 \sum_{p=0}^{N} \epsilon^{p/2} C_p,
\]
for some expressions $D_p$ with $\epsilon^2 D_p = \Theta(\mathcal{E}_\epsilon)$, some $C_p = \Theta(\mathcal{E}_\epsilon + 1)$, $m_N \leq m^{3+N}$ and

\[
\|f_{kN}\|_{H^{|\deg^*(\gamma_{kN})|+|\deg(\nu)|}} + \|\partial_t f_{kN}\|_{H^{|\deg^*(\gamma_{kN})|}} \leq \epsilon^{1/2} C_f^{3+N},
\]

\[
\|\partial^2_x f_{kN}\|_{\infty} + \|\partial_t \partial^2_x f_{kN}\|_{C(|\deg^*(\gamma_{kN})|)} \leq C_f^{3+N}.
\]

Moreover, we have $\deg^*(\gamma_{kN}) \leq \deg^*(\rho)$. We will now show that

\[
D^\infty := \sum_{p=0}^{\infty} \epsilon^{p/2} D_p
\]

does exist, $\epsilon^2 D^\infty = \Theta(\mathcal{E}_\epsilon)$ and

\[
\epsilon^2 \int_{\mathbb{R}} \gamma \partial^2_x R_{i_1} \partial^2_x R_{i_2} f \, dx = \epsilon^2 \partial_t D^\infty + \epsilon^2 \Theta(\mathcal{E}_\epsilon + 1).
\]

By taking a close look at the proof of (69), we find that

\[
\epsilon^2 \epsilon^{1/2} D_p \leq \epsilon \epsilon^{1/2} m^{p+3} \Theta^{0+3} \mathcal{E}_\epsilon^{1+3},
\]

\[
\epsilon^\frac{p}{2} C_p \leq \epsilon^\frac{p}{2} m^{p+3} \Theta^{0+3} (\mathcal{E}_\epsilon + 1),
\]

for some $c > 1$ as long as $f, i_p, i_\alpha$, and $\mathcal{E}$ are fixed. We emphasize that this is in particular possible due to the fact that $\deg^*(\gamma_{kN})$ is always uniformly bounded by $\deg^*(\rho)$.

By now choosing $\epsilon_0$ small enough, for instance, such that

\[
\epsilon_0^{1/4} mc \leq 1,
\]

we get the following. There is a $c \in \mathbb{R}$ such that

\[
\epsilon^2 D^\infty = \epsilon^2 \sum_{p=0}^{\infty} \epsilon^{p/2} D_p \leq \epsilon^2 \sum_{p=0}^{\infty} \epsilon^{p/2} |D_p| \leq \epsilon \sum_{p=0}^{\infty} \epsilon^{p/2} c \mathcal{E}_\epsilon = \epsilon c \mathcal{E}_\epsilon \sum_{p=0}^{\infty} \epsilon^{p/4} = \epsilon \Theta(\mathcal{E}_\epsilon);
\]

analogously, we get

\[
\sum_{p=0}^{\infty} \epsilon^{p/2} C_p \leq \sum_{p=0}^{\infty} \epsilon^{p/2} |C_p| = \Theta(\mathcal{E}_\epsilon + 1).
\]

Moreover,

\[
\epsilon^{-N/2} \sum_{k=1}^{m_N} \int_{\mathbb{R}} \gamma_{kN} \partial^2_x R_{i_1} \partial^2_x R_{i_2} f_{kN} \, dx
\]

\[
\leq \epsilon^{-N/2} C_f \left( \|R_1\|_{H^r} \|R_1\|_{C^{0,|\deg^*(\nu)|}} + \|R_1\|_{H^r} \|R_{-1}\|_{C^{0,|\deg^*(\nu)|}} 
\right.
\]

\[
+ \|R_{-1}\|_{H^r} \|R_1\|_{C^{0,|\deg^*(\nu)|}} + \|R_{-1}\|_{H^r} \|R_{-1}\|_{C^{0,|\deg^*(\nu)|}} )
\]

\[
= 0, \text{ for } N \to \infty.
\]
We now obtain
\[\varepsilon^2 \int_{\mathbb{R}} \gamma \partial_x^2 R_j \partial_x^2 R_j f \, dx = \varepsilon^2 \partial_t D^\infty + \varepsilon^2 \sum_{p=0}^\infty \tilde{C}_p^{p/2} \sup_{b_0, \omega_1} \left| \frac{R_{-1}^b}{R_1^b} \right|^{H} \leq C_R,\]

For \( \mathrm{deg}^a(\rho) \leq 1 \) or \( \mathrm{deg}^a(\rho) < \mathrm{deg}(\omega) \), one can modify the above proof by exploiting the special structure of (2), that is,\[\frac{\partial_t (R_1 + R_{-1})}{i \omega (R_1 - R_{-1})} = \varepsilon^{-\beta} \partial_t \{|(\varepsilon \Psi) + \text{Res}_{u_1}(\varepsilon \Psi)|\} \]

to obtain a terminating algorithm that gives out an explicit expression \( \varepsilon^2 D \) consisting of a finite sum of integrals.

Corollary 4.11 now allows us to prove Theorem 1.1.

**Proof of Theorem 1.1.** For \( \ell' \geq \lceil \mathrm{deg}(\omega) \rceil + \lceil \mathrm{deg}^a(\rho) \rceil + 1 \), we can use corollary 4.11 together with Gronwall’s inequality in order to obtain the \( \mathcal{O}(1) \)-boundedness of \( \hat{E}_\varepsilon \) for all \( t \in [0, T_\varepsilon/\varepsilon^2] \) as long as \( \varepsilon_0 > 0 \) is chosen sufficiently small. For sufficiently small \( \varepsilon_0 > 0 \) there thus is some constant \( C_R \) such that

\[\sup_{[0, T_\varepsilon/\varepsilon^2]} \| u - \varepsilon \Psi_NL_{\varepsilon} \|_{H^s} \leq C_R,\]

due to corollary 4.11. Choosing \( \ell' \geq s_A \), estimate (17) now allows to conclude

\[\sup_{[0, T_\varepsilon/\varepsilon^2]} \| u - \varepsilon \Psi_NL_{\varepsilon} \|_{H^s} \leq \sup_{[0, T_\varepsilon/\varepsilon^2]} \| (u_{-1} - u_1) - \varepsilon \left( \begin{array}{c} \Psi_NL_{\varepsilon} \\ 0 \end{array} \right) \|_{H^s} \]
\[\leq \sup_{[0, T_\varepsilon/\varepsilon^2]} \| \varepsilon^{\beta} \left( \begin{array}{c} \partial R_{-1} \\ \partial R_1 \end{array} \right) \|_{H^s} + \sup_{[0, T_\varepsilon/\varepsilon^2]} \| \varepsilon^{\Psi_NL_{\varepsilon}} - (\varepsilon \Psi_NL_{\varepsilon}) \|_{H^s} \]
\[\leq \mathcal{O}(\varepsilon^{3/2}).\]

\[\square\]

## 5 DISCUSSION

As model problem for the 2D WWP with finite depth and surface tension \( b \geq 0 \), one can look at (2) with

\[\omega(k) = \rho(k) = \mathrm{sign}(k) \sqrt{k \tanh(k)(1 + bk^2)},\]

compare previous studies\(^9,20\) for the case without surface tension \( b = 0 \). Theorem 1.1 grants us the validity of the NLS approximation for all \( b \geq 0 \) and \( k_0 > 0 \), excluding some special pairs \((b, k_0)\) with \( 0 < b < 1/3 \). Indeed, the validity of the NLS approximation for the full 2D WWP with finite depth and surface tension was recently proven in Düll.\(^16\)

System (2) with \( \omega \) and \( \rho \) given in Fourier space by

\[\omega(k) = \rho(k) = \mathrm{sign}(k) \sqrt{\frac{k \tanh(k)}{1 + k \tanh(k)}(1 + k^2 + k^4)}\]

can be considered as a model problem for the 2D WWP with finite depth and ice cover. This model has the same linear dispersion relation as the full problem; see, for example,\(^21\) Moreover, its quasilinear quadratic term shares principle difficulties with the ones of the full problem regarding the construction of the normal form transformations and the loss of regularity in the error estimates. We omit an analysis of the possible resonances that can occur but for some \( k_0 \), for example, \( k_0 = 1 \), the resonance condition of this paper is fulfilled such that Theorem 1.1 grants the validity of the NLS approximation.
approximation for these wavenumbers. Thus, the techniques of this paper might be useful for proving a NLS validity result for the full 2D WWP with ice cover.

Our result could also be interesting for double dispersion equations. With (1), we already gave one example but Theorem 1.1 also applies to other quasilinear double dispersion equations, like, for example, \( \partial_t^2 u = \partial_x^6 u + \partial_x^2 u^2 \) or \( \partial_t^2 u = -\partial_x^4 u + \partial_x^2 u + \partial_x^2 u^2 - \partial_x^2 u^2 \).

**ACKNOWLEDGEMENTS**

This work was supported by the Deutsche Forschungsgemeinschaft DFG under the grant DU 1198/2. The author thanks Guido Schneider and Wolf-Patrick Düll for discussions. The author thanks the referees for their useful comments. This work does not have any conflicts of interest. Open Access funding enabled and organized by Projekt DEAL.

**ORCID**

Max Heß https://orcid.org/0000-0002-3079-998X

**REFERENCES**

1. Schneider G, Uecker H. *Nonlinear PDEs. A Dynamical Systems Approach*, Graduate Studies in Mathematics, vol. 182. Providence, RI: American Mathematical Society; 2017.
2. Kolkovska N, Vucheva V. Invariant preserving schemes for double dispersion equations. *Adv Difference Equ*. 2019;216:1-16.
3. Li Y, Gao Y. The method of lower and upper solutions for the cantilever beam equations with fully nonlinear terms. *J Inequal Appl*. 2019;136:1-16.
4. Wang S, Chen G. Cauchy problem of the generalized double dispersion equation. *Nonlinear Anal*. 2006;64(1):159-173.
5. Ablowitz MJ, Segur H. Solitons and the inverse scattering transform. *SIAM Studies in Applied Mathematics*, Vol. 4: SIAM; 1981. https://doi.org/10.1137/1.9781611970883
6. Schneider G, Sunny DA, Zimmermann D. The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions. *J Dynam Differential Equations*. 2015;27(3):1077-1099.
7. Kalyakin LA. Asymptotic decay of a one-dimensional wave packet in a nonlinear dispersive medium. *Sb Math*. 1988;60:457-483.
8. Schneider G. Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. *J Differential Equations*. 2005;216:354-386.
9. Schneider G, Wayne CE. Justification of the NLS approximation for a quasilinear water wave model. *J Differential Equations*. 2011;251:238-269.
10. Düll W-P. Existence of long time solutions and validity of the nonlinear Schrödinger approximation for a quasilinear dispersive equation. *J Differential Equations*. 2018;264(4):2598-2632.
11. Zakharov VE. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J Appl Mech Techn Phys*. 1968;9(2):190-194.
12. Düll W-P, Schneider G, Wayne CE. Justification of the nonlinear Schrödinger approximation for the evolution of gravity driven 2D surface water waves in a canal of finite depth. *Arch Ration Mech Anal*. 2016;220(2):543-602.
13. Il’ichev AT. Envelope solitary waves and dark solitons at a water-ice interface. *Proc Steklov Inst Math*. 2015;289(1):1751-1768.
14. Hunter JK, Ifrim M, Tataru D. The NLS approximation for two dimensional deep gravity waves. *Proc Amer Math Soc*. 2015;143(3):1295-1321.
15. Cummings P, Wayne CE. Modified energy functionals and the NLS approximation. *Discrete Contin Dyn Syst*. 2017;37(3):1295-1321.
16. Il’ichev AT. Envelope solitary waves and dark solitons at a water-ice interface. *Proc Steklov Inst Math*. 2015;289(1):152-166.

**How to cite this article:** Heß M. Validity of the nonlinear Schrödinger approximation for quasilinear dispersive systems with more than one derivative. *Math Meth Appl Sci*. 2022;45:1725-1751. doi:10.1002/mma.7885