The long wave limit for the water wave problem

I. The case of zero surface tension

Guido Schneider  
Mathematisches Institut  
Universität Bayreuth  
95440 Bayreuth  
Germany  
guido.schneider@uni-bayreuth.de

C. Eugene Wayne  
Department of Mathematics and Center for BioDynamics  
Boston University  
111 Cummington St.  
Boston, MA 02215  
USA  
cew@math.bu.edu

Abstract

The Korteweg–de Vries equation, Boussinesq equation, and many other equations can be formally derived as approximate equations for the 2D water wave problem in the limit of long waves. Here we consider the classical problem concerning the validity of these equations for the water wave problem in an infinitely long canal without surface tension. We prove that the solutions of the water wave problem in the long wave limit split up into two wave packets, one moving to the right and one to the left, where each of these wave packets evolves independently as a solution of a Korteweg–de Vries equation. Our result allows us to describe the nonlinear interaction of solitary waves.

Keywords: free boundary value problem, quasilinear system, KdV–equation, approximation, water waves

AMS–Classification: 35Q35, 35Q53, 35R35, 35B20
Contents

1 Introduction .................................................. 2
  1.1 The water wave problem .................................. 2
  1.2 The long wave limit ..................................... 4

2 The main ideas .................................................. 9

3 The operator $\mathcal{K}(X)$ .................................. 21
  3.1 The Riemann mapping .................................... 22
  3.2 The representation ...................................... 23
  3.3 The formal expansion .................................... 25
  3.4 Analytic properties of the expansion .................... 26

4 Formulation as evolutionary problem ....................... 34

5 Long wave analysis ............................................. 42
  5.1 Derivation of the KdV–equations ....................... 42
  5.2 The approximation and estimates for the residual .... 44

6 The error estimates ............................................. 48

1 Introduction

The introduction is separated into two parts. In the first we explain the water wave problem and then in the second we give our results for the long wave limit.

1.1 The water wave problem

We consider an irrotational, incompressible fluid in an infinitely long canal of fixed depth under the influence of gravity. The bottom is impermeable and surface tension is neglected. The coordinates are denoted with $x_1 \in \mathbb{R}$ and $x_2$ in the bounded direction. The fluid fills the domain $\Omega(t)$ in between the bottom $\{(\alpha, -1)| \alpha \in \mathbb{R}\}$ and the free surface

$$\Gamma(t) = \{(\bar{X}_1(\alpha, t), \bar{X}_2(\alpha, t)) = (\alpha + X_1(\alpha, t), X_2(\alpha, t)) \mid \alpha \in \mathbb{R}\}.$$ 

For fixed time $t$, $\Gamma$ is a Jordan–curve which remains bounded in the $x_2$–direction and which has no intersection with $\{(\alpha, -1) \mid \alpha \in \mathbb{R}\}$. 

2
The velocity field $u = (u_1, u_2)$ in $\Omega(t)$ is governed by Euler’s equations. We have the balance of forces

$$\partial_t u + (u \cdot \nabla) u = -\nabla p + g(0, -1), \tag{1}$$

with $u(x, t) \in \mathbb{R}^2$, $p(x, t) \in \mathbb{R}$, $g \in \mathbb{R}$, and the incompressibility condition

$$\partial_{x_1} u_1 + \partial_{x_2} u_2 = 0. \tag{2}$$

It is well known (cf. [LL91]) that a solution with an irrotational initial condition stays irrotational under the evolution of Euler’s equation, and so we also have

$$\partial_{x_2} u_1 - \partial_{x_1} u_2 = 0. \tag{3}$$

Moreover, we have the boundary condition at the bottom $u_2|_{x_2 = -1} = 0$. This system is completed with the boundary conditions for the free surface $\Gamma$. We assume for fixed time $t$ that the pressure $p$ is constant along the free surface $\Gamma$ and that $(1, u_1, u_2)$ is tangential to $(t, \Gamma(t))$.

Without loss of generality we assume $g = 1$, and we have assumed the depth of the fluid at rest to be one. By definition we have $\partial_t \bar{X} = \partial_x X = u$ on the free surface, and so (1) becomes

$$\partial^2_t \bar{X} = \partial_t u + (u \cdot \nabla) u = -\nabla p - (0, 1). \tag{4}$$

On the top surface we have $p(\bar{X}(\alpha, t), t) = \text{const.}(t)$, and so $\nabla p \cdot \partial_\alpha \bar{X} = 0$. Multiplying (4) with $\partial_\alpha \bar{X}$ gives

$$(\partial^2_t \bar{X}) \cdot (\partial_\alpha \bar{X}) + \partial_\alpha \bar{X}_2 = 0 \tag{5}$$

or equivalently

$$\partial^2_t X_1 (1 + \partial_\alpha X_1) + \partial_\alpha X_2 (1 + \partial^2_t X_2) = 0. \tag{6}$$
From the incompressibility (2) of the fluid and from the irrotationality (3) of the motion of the fluid it follows that the velocities $\partial_t X_1 = u_1$ and $\partial_t X_2 = u_2$ are related. It is advantageous to express $\partial_t X_2$ in terms of $\partial_t X_1$. The associated operator is called the Dirichlet-Neumann operator. We denote it by $\mathcal{K}$, and note that it is linear in $\partial_t X_1$, but it depends nonlinearly on $X$, i.e.

$$
\partial_t X_2 = \mathcal{K}(X)\partial_t X_1.
$$

(7)

Properties of the bounded operator $\mathcal{K}(X)$ are explained in Section 3. The evolution of the free surface $\Gamma(t)$ is governed by the equations (6) and (7). Using (2) and (3) the velocity field $u$ in the interior of the domain can be computed. From (1) the pressure $p$ follows up to a constant. Thus we see that the dynamics of the system can completely be described by the evolution of the free surface $\Gamma(t)$. By introducing $U_1 = \partial_t X_1$ equations (6) and (7) can be written for $\partial_\alpha X_1$ and $\partial_\alpha X_2$ sufficiently small as a first order system in $X_1$, $X_2$, and $U_1$

$$
\begin{align*}
\partial_t X_1 &= U_1 \\
\partial_t X_2 &= \mathcal{K}(X)U_1 \\
\partial_t U_1 &= -(1 + \partial_\alpha X_1 + (\partial_\alpha X_2)\mathcal{K}(X))^{-1}[(\partial_\alpha X_2)(1 + [\partial_t, \mathcal{K}(X)]U_1)]
\end{align*}
$$

(8)

with initial conditions $X_1|_{t=0} = \phi_0$, $X_2|_{t=0} = \phi_1$, and $U_1|_{t=0} = \phi_2$, where $[M, N] = MN - NM$ defines the commutator of two operators $M$ and $N$. The freedom in the parametrization of the top surface $\Gamma(t)$ is reflected by the system. For relatively flat surfaces without loss of generality we can assume an initial parametrization $X_1|_{t=0} = 0$. Thus, we have essentially two independent initial conditions $X_2|_{t=0}$ and $U_1|_{t=0}$.

### 1.2 The long wave limit

In 1834, Russell observed a solitary wave in a canal between Edinburgh and Glasgow. In the following years he made some experiments and in 1844 he published his observations ([Ru1844]). There was a controversy if such a wave of permanent homoclinic form can exist if dissipation is neglected. Airy ([Ai1845]) argued that dispersion will destroy this wave. It was accepted that such waves exist when Boussinesq ([Bo1871]) and Rayleigh ([Ra1876]) found approximations to such a wave by using perturbation analysis. The perturbation ansatz for unidirectional long waves

$$
\begin{pmatrix}
\partial_\alpha X_1(\alpha, t) \\
X_2(\alpha, t) \\
U_1(\alpha, t)
\end{pmatrix} = \varepsilon^2 A(\varepsilon(\alpha - t), \varepsilon^3 t) \begin{pmatrix}
-1 \\
1 \\
-1
\end{pmatrix} + \mathcal{O}(\varepsilon^3)
$$

(9)
with $0 < \varepsilon \ll 1$ shows that the approximate amplitude $A$ has to satisfy in lowest order a Korteweg–de Vries equation \(^1\) (named after [KdV1895]) which is given in its normalized form by

$$\partial_t A = -\partial_{\alpha}^3 A + 6A\partial_{\alpha} A,$$

(10)

where $A(\alpha, T) \in \mathbb{R}$, $c \geq 0$, $T = \varepsilon^3 t$, and $\alpha = \varepsilon(\alpha - t)$. The important point about (10) is that the nonlinear terms are in some equilibrium with the dispersion, and so (10) possesses a family of solitary waves

$$A(\alpha, T) = \tilde{A}_\beta(\alpha - \beta^2 T) = -\frac{\beta^2}{2} \text{sech}^2\left(\frac{1}{2} \beta(\alpha - \beta^2 T)\right)$$

as solutions. In 1954, Friedichs and Hyers ([FH54]) showed rigorously the existence of solitary waves for the water wave problem without surface tension. In addition, we refer the reader to [Io65], where an overview can be found about waves of permanent form if surface tension is included.

In 1965, Zabusky and Kruskal [ZK65] observed that the KdV-equation possesses $N$-solitons, i.e. waves of permanent form which go through each other without changing their shape asymptotically, i.e.

$$A_{N-Sd}(\alpha, T) \sim \sum_{j=1}^{N} \tilde{A}_{\beta_j}(\alpha - \beta_j^2 T \pm \gamma_j + \delta_j),$$

for $T \to \pm \infty$, with $\beta_1 > \ldots > \beta_N$ and $\gamma_j, \delta_j \in \mathbb{R}$ some phase-shifts.

In this paper we show that the water wave problem also possesses $N$-soliton like solutions, i.e. waves which go through each other without changing their shape very much, except at the time of nonlinear interaction. In order to do so we prove exact estimates between the approximations (9) obtained via the Korteweg–de Vries equation (10) and true solutions of the water wave problem (8).

We denote Fourier transform by $(\mathcal{F}u)(k) = \hat{u}(k) = \frac{1}{2\pi} \int u(x)e^{-ikx}dx$. The Sobolev space $H^s$ is equipped with the norm $\|u\|_{H^s} = \left(\int |\hat{u}(k)|^2(1 + |k|^2)^s dk\right)^{1/2}$. Moreover, let $\|u\|_{H^s(\mathbb{R})} = \|u\rho^n\|_{H^s}$, where $\rho(x) = (1 + x^2)^{1/2}$, and let $\|u\|_{C_0^0} = \sum_{j=0}^{n} \|\partial_x^j u\|_{C_0}$. Then the following holds.

**Theorem 1.1**

*Fix* $s \geq 4$. *For all* $C_1, T_0 > 0$ *there exist* $C_2, \varepsilon_0 > 0$ *such that for all* $\varepsilon \in (0, \varepsilon_0)$ *the following is true. Let* $A_1 \in C([0, T_0], H^{s+6}(\mathbb{R}))$ *be a solution of the KdV-equation*

$$\partial_t A_1 = -\frac{1}{6} \partial_{\alpha}^3 A_1 - \frac{3}{4} A_1 \partial_{\alpha} A_1$$

\(^1\)Boussinesq (cf. [Bo1877]) had also derived this equation, but made no direct use of it.
with \( \sup_{t \in [0, T_1]} \| A_1(T) \|_{H^{-s}(Q)} \leq C_1 \). Then for \( t \in [0, T_1/\varepsilon^3] \) with \( T_1 = T_0 \) we have a solution of the water wave problem (8) with \( \sup_{t \in [0, T_1/\varepsilon^3]} |X_1(0, t)| \leq C_2 \varepsilon^{1/2} \) and

\[
\sup_{t \in [0, T_1/\varepsilon^3]} \left\| \begin{pmatrix} \partial_t X_1 \\ X_2 \\ U_1 \end{pmatrix} - \psi_{KdV}(\varepsilon, \alpha, t) \right\|_{H^{-1} \times H^{-1} \times H^{-1/2}} \leq C_2 \varepsilon^{7/2},
\]

where

\[
\psi_{KdV}(\varepsilon, \alpha, t) = \varepsilon^2 A_1(\varepsilon - t, \varepsilon^3 t) \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.
\]

**Proof.** This follows directly from Theorem 6.1, Lemma 2.5 and Remark 6.2.

**Corollary 1.2**

If Theorem 1.1 is satisfied, then for the Euler variables \( u_1 = u_1(x, t) \) and \( \eta = \eta(x, t) \) defined by \( u_1(X_1(\alpha, t), t) = U_1(\alpha, t) \) and \( \eta(X_1(\alpha, t), t) = X_2(\alpha, t) \) we have the estimates

\[
\sup_{t \in [0, T_1/\varepsilon^3]} \left\| \begin{pmatrix} \eta \\ u_1 \end{pmatrix} - \varepsilon^2 A_1(\varepsilon(\cdot - t), \varepsilon^3 t) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\|_{C_0^{-3/2}} \leq C_2 \varepsilon^{5/2}
\]

**Proof.** See page 19 in section 2.

This means that we can find the dynamics known from the KdV-equation in the water wave problem, too.

In the literature there exist two related results. In [KN86] estimates are shown on an \( O(1/\varepsilon) \)-time scale in the original system. However, a timescale of \( O(1/\varepsilon) \) in the original variables represents times \( T \approx O(\varepsilon^2) \) in the KdV variables – i.e. far too short for physical interest. But see also [Sch96]. Moreover, analytic initial conditions are required. Results closer to our own are contained in the work of W. Craig [Cr85], one part of which can be interpreted as an approximation theorem in the above sense. In particular, his work considers initial conditions in Sobolev spaces and gives estimates over the right qualitative times scale (i.e. \( O(\varepsilon^{-3}) \)). However, Craig’s results have two drawbacks which make them unsuitable for our purposes. First the approximation time \( T_1 \) in Theorem 1.1 may be smaller than \( T_0 \), i.e. smaller than the interaction time of the solitons. Second, for technical reasons Craig assumes that \( X_1 \in L^2 \). This implies that the initial conditions of the KdV-equation must be assumed to have vanishing mean.
value: \( \int A(\alpha)|_{T=0} d\alpha = 0 \). This excludes soliton solutions a priori. We avoid these restrictions by working with \( \partial_a X_1 \in L^2 \) and by taking \( \psi_{KdV} \) as center of the coordinate system. See Remark 2.1.

Our result goes beyond this approximation theorem. We consider the general long wave limit, i.e. we consider initial conditions \( X_1|_{t=0}(\alpha) = 0 \), \( X_2|_{t=0}(\alpha) = \varepsilon^2 \Phi_1(\varepsilon \alpha) \) and \( \partial_t X_1|_{t=0}(\alpha) = \varepsilon^2 \Phi_2(\varepsilon \alpha) \) with \( 0 < \varepsilon \ll 1 \). There exist a number of formal models and a scientific discussion (cf. [Lo64], [BS71], [BBM72], [Me79], [Ma91]) going back to [Bo1877] about the question which of these models provide a good description of the dynamics of the water wave problem in the long wave limit. Our answer is as follows:

**Theorem 1.3**

The solutions of the water wave problem in the long wave limit split up into two wave packets, one moving to the right and one to the left, where each of these wave packets evolves independently as a solution of a Korteweg–de Vries equation. More precisely, we have: Fix \( s \geq 4 \). For all \( C_1, T_0 > 0 \) there exist \( C_2, \varepsilon_0 > 0 \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) the following is true. Let \( X_1|_{t=0} = 0 \), \( X_2|_{t=0}(\alpha) = \varepsilon^2 \Phi_1(\varepsilon \alpha) \) and \( \partial_t X_1|_{t=0}(\alpha) = \varepsilon^2 \Phi_2(\varepsilon \alpha) \) with \( \| (\Phi_1, \Phi_2) \|_{H^{s+6}(\mathbb{R}) \cap H^{s+11}(0)} \leq C_1 \). Then there exists a unique solution of the water wave problem (8) with \( \sup_{t \in [0,T_0/\varepsilon^3]} |X_1(0, t)| \leq C_2 \varepsilon^{1/2} \) and

\[
\sup_{t \in [0,T_0/\varepsilon^3]} \left\| \begin{pmatrix} \partial_t X_1 \\ X_2 \\ U_1 \end{pmatrix} - \psi_{LWL}(\varepsilon, \cdot, t) \right\|_{H^{s-1} \times H^s \times H^{s-1/2}} \leq C_2 \varepsilon^{7/2},
\]

where

\[
\psi_{LWL}(\varepsilon, \alpha, t) = \varepsilon^2 A_1(\varepsilon(\alpha - t), \varepsilon^3 t) \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} + \varepsilon^2 A_2(\varepsilon(\alpha + t), \varepsilon^3 t) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.
\]

The amplitudes \( A_1 = A_1(\alpha, T) \) and \( A_2 = A_2(\alpha, T) \) satisfy

\[
\partial_T A_1 = -\frac{1}{6} \partial_\alpha^3 A_1 - \frac{3}{4} A_1 \partial_\alpha A_1 \quad \text{and} \quad \partial_T A_2 = \frac{1}{6} \partial_\alpha^3 A_2 + \frac{3}{4} A_1 \partial_\alpha A_2,
\]  

(11)

with \( 2A_1|_{T=0} = \Phi_2 + \Phi_1 \) and \( 2A_2|_{T=0} = \Phi_2 - \Phi_1 \).

**Proof.** See page 50 in Section 6.

**Remark 1.4** Note that in order to prove the approximation result in Theorem 1.3, we have to take advantage of the freedom to change the parameterization \( x = \alpha + X_1(\alpha, t) \). The way in which we do this is explained below in Lemma 2.6.
Corollary 1.5
Assume that Theorem 1.3 holds. Then we have for the Eulerian variables \( u_1 = u_1(x, t) \) and \( \eta = \eta(x, t) \) defined by \( u_1(\bar{X}_1(\alpha, t), t) = U_1(\alpha, t) \) and \( \eta(\bar{X}_1(\alpha, t), t) = X_2(\alpha, t) \) the estimate

\[
\sup_{t \in [0, T_0/\varepsilon^3]} \left\| \begin{pmatrix} \eta \\ u_1 \end{pmatrix} - \varepsilon^2 A_1(\varepsilon^3 t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \varepsilon^2 A_2(\varepsilon^3 t) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\|_{C_t^{3/2}} \leq C_2 \varepsilon^{5/2}.
\]

Proof. See page 19 in section 2. □

Remark 1.6 This means that to the order of approximation considered in the long-wave approximation, it is sufficient to consider two decoupled KdV-equations to describe the dynamics completely. While other formal, more complicated, models could also be expected to be correct in the sense of Theorem 1.1, we feel that the KdV description has the advantage of being simple, exactly solvable, and independent of \( \varepsilon \), while other models still depend on \( \varepsilon \). See for instance (47) which is one such a model.

The long-time behavior for the water wave problem in the long wave limit is as follows. On a time scale \( O(1/\varepsilon) \) the solutions split up into two wave packets one moving to the right and one to the left. These wave packets evolve independently as solutions of the KdV-equations. The long-time behavior for the solutions of the KdV-equations can be computed explicitly with the help of the inverse scattering transform. Some solitons which are ordered with respect to their height evolve out of a dispersive remainder (cf. [EvH81]). As drawn in Figure 2 the same behavior is predicted by Corollary 1.5 for the water wave problem in the long wave limit.

Thus, Theorem 1.3 can be seen as an analytic confirmation of the behavior of tsunamis in the Pacific Ocean (cf. [Pe96]). These large waves caused by earthquakes appear as wavetrains (time scale \( O(1/\varepsilon) \)) or in a number of single waves (time scale \( O(1/\varepsilon^3) \)).

Our analysis is based on [Sch98], where a similar result has been shown for the Boussinesq equation (48). The special difficulty for the water wave problem comes from the fact that the method that we use to obtain error estimates on these amplitude or modulation equations does not fit together very well with the existence theory for the water wave problem. Due to this reason and due to the fact that we can not allow \( X_1 \in L^2 \) if we are interested in solitary water waves we have to improve the local existence and uniqueness theory and to combine it with the method for obtaining the error estimates.

\[2\]The KdV-equation is a completely integrable infinite-dimensional Hamiltonian system
The following analysis can be transferred to the water wave problem with surface tension. This will be the subject of a forthcoming paper ([SW99]).

Throughout this paper, we assume $0 < \varepsilon \ll 1$. The $j$-th component of a vector $v$ is denoted by $(v)_{(j)}$. The commutator of two operators $L$ and $M$ is defined as $[L, M] = LM - ML$.

**Acknowledgements:** The authors are grateful for discussions with Walter Craig. Guido Schneider would like to thank for the Mathematics Departments of Boston University and the of the Pennsylvania State University for their kind hospitality. Eugene Wayne thanks the Mathematical Institute of the University of Hannover and the Mathematical Institute of the University of Stuttgart for their hospitality during a visit there, and also thanks R. Pierce for discussions about water waves. The work of E. Wayne was supported in part by the NSF through the grant DMS-9803164.

## 2 The main ideas

The character of this section is again introductory. The first thing that we do is explain why, in the context of the linearized problem, the method we use for obtaining error estimates does not fit together very well with the existence and uniqueness theory for these equations. This in turn leads to a discussion of some technical details which are
essential for the understanding of the next steps. Moreover, we prove Corollary 1.2 and 1.5 and the less complicated statements from Theorem 1.1 and 1.3. The main proof starts in Section 4.

The main aim of this section is to explain why two decoupled KdV–equations are sufficient to describe the dynamics of the water wave problem in the long wave limit. The formal derivation (see Section 5) of modulation equations to describe the dynamics on unbounded domains goes back to the last century (cf. [Bo1877]), but only in the last decade has a systematic mathematical theory been developed.

The method of modulation equations is a weakly nonlinear theory since only small solutions are considered. Therefore the dispersion relation plays a crucial role. Thus, we first consider the linearization

\[ \begin{align*}
\partial_t X_1 &= U_1 \\
\partial_t X_2 &= \mathcal{K}_0 U_1 \\
\partial_t U_1 &= -\partial_n X_2
\end{align*} \]  

of (8). In Section 3 we show that \( \mathcal{K}_0 \) is a multiplication operator in Fourier space, with \( \mathcal{K}_0(k) = -i \tanh(k) \). In Section 1 we already explained that we cannot choose \( X_1 \) in a Sobolev space if we want to describe solitary water waves. In the following paragraph, we show that \( X_1 \) grows linearly in time if \( \phi_1 \) and \( \phi_2 \) have non-vanishing mean value. For (12) and also for the nonlinear system (8) we have three initial conditions

\[ X_{1|t=0} = \phi_0, \ X_{2|t=0} = \phi_1, \ \text{and} \ \partial_t X_{1|t=0} = \phi_2. \]

In the situation we are interested in we can choose without loss of generality the parametrization

\[ X_{1|t=0} = 0, \ X_{2|t=0} = \phi_1, \ \text{and} \ \partial_t X_{1|t=0} = \phi_2. \]

which we call reference initial conditions. Note that in (13) and (15) the initial condition \( \phi_0 \) only appears as an additive constant. As a consequence, we have essentially two independent initial conditions \( \phi_1 \) and \( \phi_2 \) for the water wave problem (8).

Taking Fourier transforms, the solution of (12) is given explicitly by:

\[ \begin{align*}
\hat{X}_1 &= \hat{\phi}_0 + \frac{1}{\mathcal{K}_0} \frac{e^{\lambda t} + e^{-\lambda t} - 2}{2} \hat{\phi}_1 + \frac{1}{\lambda} \frac{e^{\lambda t} - e^{-\lambda t}}{2} \hat{\phi}_2 \\
\hat{X}_2 &= \frac{e^{\lambda t} + e^{-\lambda t}}{2} \hat{\phi}_1 + \frac{\mathcal{K}_0 e^{\lambda t} - e^{-\lambda t}}{2} \hat{\phi}_2 \\
\hat{U}_1 &= \frac{(e^{\lambda t} - e^{-\lambda t})}{2 \mathcal{K}_0} \frac{\lambda}{\phi_1} + \frac{e^{\lambda t} + e^{-\lambda t}}{2} \hat{\phi}_2,
\end{align*} \]
where \( \hat{\lambda} = \hat{\lambda}(k) = i \sqrt{k \tanh k} \). We obtain the estimates

\[
\|X_1\|_{H^\vartheta} \leq \|\phi_0\|_{H^\vartheta} + C(\min(\|\frac{\dot{\phi}_1}{K_0}\|_{H^0(s)}^2, t\|\phi_1\|_{H^\vartheta}) + \min(\|\frac{\dot{\phi}_2}{\lambda}\|_{H^0(s)}^2, t\|\phi_2\|_{H^\vartheta-1/2}))
\]

\[
\|X_2\|_{H^\vartheta} \leq C(\|\phi_1\|_{H^\vartheta} + \|\phi_2\|_{H^\vartheta-1/2})
\]

\[
\|U_1\|_{H^\vartheta-1/2} \leq C(\|\phi_1\|_{H^\vartheta} + \|\phi_2\|_{H^\vartheta-1/2})
\]

Remark 2.1 It is easy to see that \( X_1 \) grows in \( H^s \) linearly in time if \( \phi_1 \) and \( \phi_2 \) have not vanishing mean value, i.e. more precisely if \( k \mapsto \dot{\phi}_2(k)/k \) does not lie in a suitable Banach space \( H^0(s) \). However, we can control \( Z_1 = K_0X_1 \)

\[
\|Z_1\|_{H^\vartheta} \leq \|K_0\phi_0\|_{H^\vartheta} + C(\|\phi_1\|_{H^\vartheta} + \|\phi_2\|_{H^\vartheta-1/2})
\]

It turns out that also in the nonlinear system (8) we can replace the variable \( X_1 \) by the variable \( Z_1 \). Moreover, we can allow \( Z_1 \) to be an element of the Sobolev space \( H^s \). The variable \( X_1 \) is then controlled a posteriori in Lemma 2.5 and in the proof of Lemma 2.6.

This new collection of variables is denoted by \( \mathcal{W} = (Z_1, X_2, U_1) \). In order to prove local existence and uniqueness for (8) we have to extend it to a bigger system with another variable \( V_1 = \partial_t U_1 \). The extended system is quasilinear in the sense of [Kat75] so the local existence and uniqueness of solutions \( \mathcal{W}_c = (Z_1, X_2, U_1, V_1) \) will follow.

We consider the linearization

\[
\partial_t \mathcal{W}_c = L_c \mathcal{W}_c
\]

with

\[
L_c \mathcal{W}_c = \begin{pmatrix}
K_0 U_1 \\
K_0 U_1 \\
V_1 \\
-K_0 \partial_t U_1
\end{pmatrix}.
\]

of the extended system to explain why the local existence and uniqueness theory does not fit together very well with the theory for the error estimates.

For (14) we have the freedom of another initial condition \( V_1|_{t=0} = \phi_3 \). But not all solutions of (14) give solutions of (12), only those for which we have

\[
\mathcal{W}_c|_{t=0} \in \mathcal{C}_{p,l} = \{(K_0\phi_0, \phi_1, \phi_2, \phi_3) \mid \phi_3 = -\partial_t \phi_1\}.
\]
For (14) from $\mathcal{W}_{e}|_{t=0} \in \mathcal{C}_{p,t}$ it follows $\mathcal{W}_{e}(t) \in \mathcal{C}_{p,t}$ for all $t \geq 0$. Similarly, we will have that not all solutions of the quasilinear system (33) give solutions of (8).

For (14) in Fourier space we have the solutions

$$
\hat{Z}_1 = \hat{K}_0 \hat{\phi}_0 + \frac{\hat{K}_0 e^{\lambda t} - e^{-\lambda t}}{\lambda} \hat{\phi}_2 + \frac{\hat{K}_0 e^{\lambda t} + e^{-\lambda t}}{\lambda^2} \hat{\phi}_3
$$

$$
\hat{X}_2 = \hat{\phi}_1 + \frac{\hat{K}_0 e^{\lambda t} - e^{-\lambda t}}{\lambda} \hat{\phi}_2 + \frac{\hat{K}_0 e^{\lambda t} + e^{-\lambda t}}{\lambda^2} \hat{\phi}_3
$$

$$
\hat{U}_1 = \frac{e^{\lambda t} + e^{-\lambda t}}{2} \hat{\phi}_2 + \frac{e^{\lambda t} - e^{-\lambda t}}{2} \hat{\phi}_3
$$

$$
\hat{V}_1 = \frac{e^{\lambda t} - e^{-\lambda t}}{2} \hat{\phi}_2 + \frac{e^{\lambda t} + e^{-\lambda t}}{2} \hat{\phi}_3.
$$

We obtain the estimates

$$
\|Z_1\|_{H^s} \leq \|K_0 \phi_0\|_{H^s} + C(\|\phi_2\|_{H^{s-1/2}} + \min(\|\frac{\hat{\phi}_3}{\lambda}\|_{H^0(s-1/2), t\|\phi_3\|_{H^{s-1}}}))
$$

$$
\|X_2\|_{H^s} \leq \|\phi_1\|_{H^s} + C(\|\phi_2\|_{H^{s-1/2}} + \min(\|\frac{\hat{\phi}_3}{\lambda}\|_{H^0(s-1/2), t\|\phi_3\|_{H^{s-1}}}))
$$

$$
\|U_1\|_{H^{s-1/2}} \leq C(\|\phi_2\|_{H^{s-1/2}} + \min(\|\frac{\hat{\phi}_3}{\lambda}\|_{H^0(s-1/2), t\|\phi_3\|_{H^{s-1}}}))
$$

$$
\|V_1\|_{H^{s-1}} \leq C(\|\phi_2\|_{H^{s-1/2}} + \|\phi_3\|_{H^{s-1}}).
$$

and so again for arbitrary initial conditions there is a secular growth of the solutions. If $\mathcal{W}_{e}|_{t=0} \in \mathcal{C}_{p,t}$, i.e. if the linear compatibility condition $\phi_3 = -\partial_\alpha \phi_1$ holds, the solutions stay uniformly bounded in $\mathcal{H}_e^s = H^s \times H^s \times H^{s-1/2} \times H^{s-1}$, in detail

$$
\|Z_1\|_{H^s} \leq \|K_0 \phi_0\|_{H^s} + C(\|\phi_2\|_{H^{s-1/2}} + \|\phi_1\|_{H^s})
$$

$$
\|X_2\|_{H^s} \leq C(\|\phi_2\|_{H^{s-1/2}} + \|\phi_1\|_{H^s})
$$

$$
\|U_1\|_{H^{s-1/2}} \leq C(\|\phi_2\|_{H^{s-1/2}} + \|\phi_1\|_{H^s})
$$

$$
\|V_1\|_{H^{s-1}} \leq C(\|\phi_2\|_{H^{s-1/2}} + \|\phi_1\|_{H^s}),
$$

where we have used that $-i k \hat{K}_0(k)/\hat{\lambda}(k)^2 = 1$ is a bounded operator.

**Remark 2.2** This fact that the form of the equations for which local existence can be proven leads to secular growth of the size of the solutions (which is inimical to the longtime estimates we need to control the approximation scheme) is what we mean when we say that the local existence theory does not fit together well with the theory for the error estimates. To avoid this difficulty in the nonlinear case we will not work directly with the compatibility condition $\mathcal{W}_e \in \mathcal{C}_p$ (See Definition 4.5 and Definition 4.8). The secular growth of $\mathcal{W}_e$ which would come from $\phi_3$ is avoided by taking the
original system (8) and by using system (33) for \( \mathcal{W}_c \) only for local existence and uniqueness. Since spatial derivatives of \( \mathcal{W}_c \) do not show a secular growth we obtain the error estimates and local existence and uniqueness by taking (8) for the control of \( \mathcal{W} \) and (33) for the control of the derivatives of \( \mathcal{W} \) and the variable \( V_1 \).

Let us now start to analyse

\[
\partial_t \mathcal{W} = L \mathcal{W} \tag{16}
\]

with

\[
L \mathcal{W} = \begin{pmatrix}
\mathcal{K}_0 U_1 \\
\mathcal{K}_0 U_1 \\
-\partial_\alpha X_2
\end{pmatrix}
\]

in the long wave limit, where we have the initial conditions

\[
\phi_0 = 0, \quad \phi_1 = \varepsilon^2 \Phi_1(\varepsilon \alpha), \quad \phi_2 = \varepsilon^2 \Phi_2(\varepsilon \alpha)
\]

for (8). We look for solutions \( \mathcal{W} \in \mathcal{H}^s = H^s \times H^s \times H^{s-1/2} \). In Fourier space the solutions of (16) can be written as

\[
\begin{align*}
\hat{Z}_1 &= \hat{c}_1 + \hat{c}_2 - (\hat{c}_1 + \hat{c}_2)|_{t=0} \\
\hat{X}_2 &= \hat{c}_1 + \hat{c}_2 \\
\hat{U}_1 &= \frac{\lambda_1}{\kappa_0} \hat{c}_1 + \frac{\lambda_2}{\kappa_0} \hat{c}_2
\end{align*}
\tag{17}
\]

with

\[
\hat{c}_j(k, t) = \hat{c}_j|_{t=0}(k)e^{\lambda_j(k)t} \quad \text{and} \quad \hat{\lambda}_{1,2}(k) = \mp i \sqrt{k \tanh k},
\]

where

\[
\begin{pmatrix}
\hat{c}_1|_{t=0}(k) \\
\hat{c}_2|_{t=0}(k)
\end{pmatrix} = \begin{pmatrix}
\frac{1}{2} & \frac{\kappa_0}{2\lambda_1} \\
\frac{1}{2} & \frac{\kappa_0}{2\lambda_2}
\end{pmatrix} \begin{pmatrix}
\hat{\phi}_1 \\
\hat{\phi}_2
\end{pmatrix}.
\]

The KdV-equation is derived when the wavelength of the considered wave of small amplitude is much larger than the depth of the fluid, i.e. when we consider so called long waves. See (9). The ratio of depth to wavelength defines the small parameter \( \varepsilon \). Recent work, starting with [CE90, vH91, Eck93], emphasized the importance of the Fourier transform for the mathematical theory of modulation equations. In the present case this manifests itself as a concentration of Fourier modes at the wave number \( k = 0 \) with width \( \mathcal{O}(\varepsilon) \). Concentration of Fourier modes at the wave number zero, \( k = \varepsilon K \), coincides with the smallness of spatial derivatives, \( \partial_\alpha = \varepsilon \partial_\alpha \). In detail,
we have \( \mathcal{F}(S_{\varepsilon}u) = \frac{1}{\varepsilon}S_{1/\varepsilon}(\mathcal{F}u) \), where \( (S_{\varepsilon}u)(x) = u(\varepsilon x) \).

Due to the concentration of the Fourier modes at the wavenumber \( k = 0 \) for the evolution of \( \hat{c}_{1,2} \) the form of \( \hat{\lambda}_{1,2} \) at the wavenumber \( k = 0 \) plays an essential role.

\[
\begin{aligned}
x &\longrightarrow x + i \lambda_1(0)t \\
\lambda & \\
k & \\
\lambda & \\
k & \\
\end{aligned}
\]

Figure 3: Expansion of the curve of eigenvalues at the wavenumber \( k = 0 \)

Expansion of \( \hat{\lambda}_{1,2} \) gives

\[
\hat{\lambda}_{1,2}(k) = \mp ik \pm \frac{i}{6} k^3 + \mathcal{O}(|k|^5)
\]

and so

\[
\hat{c}_j(k, t) = e^{(\mp ik \pm \frac{i}{6} k^3 + \mathcal{O}(|k|^5))t} \hat{c}_j(k)|_{t=0}.
\]

As a consequence, solutions with initial conditions \( c_{1,2}|_{t=0} = \varepsilon^2 S_{\varepsilon}A_{1,2}|_{t=0} \) behave as follows. Since we have a concentration of the Fourier modes in an \( \mathcal{O}(\varepsilon) \)-neighborhood, i.e. since \( \hat{A} \) is a function of \( K = k/\varepsilon \), the order \( \mathcal{O}(k^m) \)-terms of \( \hat{\lambda}_{1,2} \) determine the dynamics on the \( \mathcal{O}(1/\varepsilon^m) \) time scale. Outside the \( \mathcal{O}(\varepsilon) \)-neighborhood of wavenumbers the Fourier modes are small in terms of \( \varepsilon \) depending on the differentiability of \( A \).

The lowest order part of the dynamics, \( e^{\mp ikL} \), with \( t = \varepsilon t \), corresponds in physical space to translations with velocities \( \pm 1 \), i.e. to solutions of \( (\partial_{\alpha} \pm \partial_{\lambda})u = 0 \). Therefore, we have the two wave packets \( \hat{c}_{1,2} \), one moving to the right and one to the left. On the time scale \( T = \varepsilon^2 t = \varepsilon^3 t \) the third order terms of \( \hat{\lambda}_{1,2} \) become relevant. Writing \( \hat{c}_{1,2} = e^{\mp ikX} \varepsilon^{-1} S_{1/\varepsilon} \hat{A}_{1,2} \) shows that in physical space \( A_1 \) and \( A_2 \) must satisfy in lowest order the equations

\[
\partial_t A_1 = -\frac{1}{6} \partial_X^3 A_1, \quad \partial_t A_2 = +\frac{1}{6} \partial_X^3 A_2, \quad (X = \varepsilon (\alpha \mp t))
\]

which are the linearized KdV-equations. Therefore, in physical space on the \( \mathcal{O}(1/\varepsilon^3) \) time scale we obtain in the long wave limit

\[
c_{1,2}(\alpha, t, \varepsilon) = \varepsilon^2 A_{1,2}(\varepsilon (\alpha \mp t), \varepsilon^3 t) + h.o.t.,
\]

14
where $A_{1,2}$ solves (18).

This formal analysis can be made rigorous with the following lemma.

**Lemma 2.3**

Let $A \in H^{m+n}$ and $|f(k)| \leq C|k|^n$ with $f \in C^{n+1}$, $m, n \in \mathbb{N}$ for all $k \in \mathbb{R}$ and a $C > 0$. Then

$$
\|f(i\partial_x)(S_\varepsilon A)(x)\|_{H^m} \leq C\varepsilon^{n-1/2}\|A\|_{H^{m+n}}.
$$

**Proof.** For completeness we recall the proof for instance from [Sch98, Lemma A.2].

$$
\|f(i\partial_x)(S_\varepsilon A)(x)\|_{H^m} \leq C\|f(k)(\varepsilon^{-1}S_{1/\varepsilon}A)(k)\|_{H^0(m)} \\
\leq C\sup_{k \in \mathbb{R}} |f(k)(1 + |k/\varepsilon|^2)^{-n/2}| \|\varepsilon^{-1}S_{1/\varepsilon}A\|_{H^0(m+n)} \\
\leq C\varepsilon^n \|S_\varepsilon A\|_{H^{m+n}} \leq C\varepsilon^{n-1/2}\|A\|_{H^{m+n}},
$$

where we denoted different constants with the same symbol $C$.  

Applying Lemma 2.3 with $f(k) = e^{\hat{\lambda}_1 t} - e^{(i\varepsilon k^{1/3} t)}$ and using the fact that $\hat{\lambda}_{1,2}$ is odd immediately shows an approximation result for the linear case similar to Theorem 6.1, the approximation theorem for the nonlinear case.

**Theorem 2.4**

Fix $s > 1$. For all $C_1, T_0 > 0$ there exist $C_2, \varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following holds. For $j = 1, 2$ let $A_j \in C([0, T_0], H^{s+6})$ be solutions of (18) with $\sup_{T \in [0, T_0]} \|A_j(T)\|_{H^{s+6}} \leq C_1$. Let

$$
\varepsilon^2 \psi = \varepsilon^2 \begin{pmatrix}
\psi_{Z_1} \\
\psi_{X_2} \\
\psi_{U_1}
\end{pmatrix} = \varepsilon^2 A_1(\varepsilon(\alpha - t), \varepsilon^3 t) \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + \varepsilon^2 A_2(\varepsilon(\alpha + t), \varepsilon^3 t) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.
$$

Then there exists a unique solution $\mathcal{W} = \varepsilon^2 \psi + \varepsilon^{7/2} R$ of (16) with

$$
\sup_{t \in [0, T_0/\varepsilon^3]} \|R\|_{H^s} \leq C_2
$$

and $\mathcal{W}|_{t=0} = \varepsilon^2 \psi|_{t=0} + \varepsilon^{7/2} R|_{t=0} \in H^s$, where $\|R|_{t=0}\|_{H} \leq C_1$.

**Proof.** We introduce another approximation $\psi_1$ given in Fourier space by

$$
\varepsilon^2 \hat{\psi}_1 = \varepsilon^2 e^{-ik\varepsilon(\varepsilon^{-1}S_{1/\varepsilon}A_1(\varepsilon^3 t) + \varepsilon^2 e^{ik\varepsilon^{-1}S_{1/\varepsilon}A_2(\varepsilon^3 t) + \varepsilon^2 \hat{\lambda}_1/\hat{K}_0}
$$

15
By construction we have
\[ \sup_{t \in [0,T_0/\varepsilon^3]} \| \varepsilon^2 \psi_1 - \varepsilon^2 \hat{\psi} \|_{\mathcal{H}} \leq C \varepsilon^{7/2}. \]

The error \( \varepsilon^{7/2} R_1 = \mathcal{W} - \varepsilon^2 \psi_1 \) satisfies
\[ \partial_t R_1 = LR_1 + \varepsilon^{-7/2} \text{Res}(\varepsilon^2 \psi_1), \]
where \( \text{Res}(\varepsilon^2 \psi_1) = -\partial_t \varepsilon^2 \psi_1 + L \varepsilon^2 \psi_1 \). The variation of constant formula yields
\[ R_1(t) = e^{Lt} R_1_{|t=0} + \varepsilon^{-7/2} \int_0^t e^{L(t-s)} \text{Res}(\varepsilon^2 \psi_1) ds. \]

Due to (17) the linear semigroup \( e^{Lt} \) is uniformly bounded. Due to the construction of the approximation \( \varepsilon^2 \psi_1 \) we have with the help of Lemma 2.3 that
\[ \sup_{t \in [0,T_1/\varepsilon^3]} \| \text{Res}(\varepsilon^2 \psi_1) \|_{\mathcal{H}} \leq C \varepsilon^{13/2} \]
with a constant \( C \) independent of \( \varepsilon \). Thus,
\[ \sup_{t \in [0,T_1/\varepsilon^3]} \| R_1(t) \|_{\mathcal{H}} \leq \sup_{t \in [0,T_0/\varepsilon^3]} \| e^{Lt} \|_{\mathcal{H} \rightarrow \mathcal{H}} \| R_1_{|t=0} \|_{\mathcal{H}} \]
\[ + \varepsilon^{-7/2} (T_0/\varepsilon^3) \sup_{t \in [0,T_0/\varepsilon^3]} \| e^{Lt} \|_{\mathcal{H} \rightarrow \mathcal{H}} \]
\[ \times \sup_{t \in [0,T_0/\varepsilon^3]} \| \text{Res}(\varepsilon^2 \psi_1) \|_{\mathcal{H}} = O(1) \]
for \( \varepsilon \rightarrow 0 \) and consequently for \( R(t) = R_1(t) + \varepsilon^{-3/2}(\psi_1(t) - \hat{\psi}(t)) \) that
\[ \sup_{t \in [0,T_1/\varepsilon^3]} \| R(t) \|_{\mathcal{H}} \leq C_2 \]
with a constant \( C_2 \) independent of \( \varepsilon \). \( \blacksquare \)

The only thing remaining to prove a linear version of Theorem 1.1 is to control the term \( \sup_{t \in [0,T_0/\varepsilon^3]} |X_1|_{\alpha=0} \).

**Lemma 2.5**

Assume that Theorem 2.4 holds and that additionally \( A_j \in C([0,T_0], H^{s+6}(2)) \) with \( \sup_{T \in [0,T_0]} \| A_j(T) \|_{H^{s+6}(2)} \leq C_1 \) for \( j = 1, 2 \). Then we have
\[ \sup_{t \in [0,T_0/\varepsilon^3]} |X_1|_{\alpha=0} \leq C_2 \varepsilon^{1/2}. \]
Proof. We have
\[ X_1|_{\alpha=0}(t) = X_1|_{\alpha=0}(0) + \int_0^t U_{1|_{\alpha=0}}(s) ds \]
\[ = X_1|_{\alpha=0}(0) + \int_0^t (\varepsilon^2 \psi U_{1|_{\alpha=0}}(s) + \varepsilon^{7/2} R_{U_{1|_{\alpha=0}}}(s)) ds. \]
The right hand side can be estimated as follows

a) \[ \sup_{t \in [0,T_0/\varepsilon^3]} |X_1|_{\alpha=0}(0) = |X_1|_{\alpha=0}(0). \]
Without loss of generality we can assume \( X_1|_{\alpha=0}(0) = 0 \), so that this term vanishes.

b) Next we use
\[ |u| \leq |(S_{\varepsilon} u)(S_{\varepsilon} \rho)(S_{\varepsilon} \rho)^{-1}| \leq |(S_{\varepsilon} \rho)^{-1}| \| (S_{\varepsilon} (u \rho)) \|_{L^\infty} \]
\[ \leq |(S_{\varepsilon} \rho)^{-1}| \| u \rho \|_{L^\infty} \leq |(S_{\varepsilon} \rho)^{-1}| \| u \|_{H^2(\Omega)} \]
which is based on Sobolev’s embedding theorem to obtain
\[ \sup_{t \in [0,T_1/\varepsilon^3]} \int_0^t |\varepsilon^2 \psi U_{1|_{\alpha=0}}(t)| dt \]
\[ \leq 2 \max_{j=1,2} \int_0^{T_1/\varepsilon^3} |\varepsilon^2 A_j(\varepsilon^3 (0 + \tau))| d\tau \]
\[ \leq 2 \varepsilon^2 \int_0^{T_0/\varepsilon^3} (1 + |\varepsilon^2 \tau|^3)^{-1} d\tau \max_{j=1,2} \sup_{T \in [0,T_1]} \| A_j(T) \|_{H^{s+2}} \]
\[ \leq 2 \varepsilon^2 \int_0^{T_1/\varepsilon^2} \varepsilon^{-1} (1 + |y|^2)^{-1} dy \ C_1 \]
\[ \leq C \varepsilon \]
for a constant \( C \) independent of \( \varepsilon \).

c) Finally, Theorem 2.4 allows us to prove
\[ \sup_{t \in [0,T_0/\varepsilon^3]} \int_0^t |\varepsilon^{7/2} R_{U_{1|_{\alpha=0}}}(t)| dt \]
\[ \leq \int_0^{T_0/\varepsilon^3} |\varepsilon^{7/2} R_{U_{1|_{\alpha=0}}}(\tau)| d\tau \]
\[ \leq (T_0/\varepsilon^3)\varepsilon^{7/2} \sup_{t \in [0,T_0]} \| R_{U_{1}} \|_{H^{-1/2}} \]
\[ \leq C \varepsilon^{1/2} \]
for a constant \( C \) independent of \( \varepsilon \) again due to Sobolev’s embedding theorem, since \( \| \cdot \|_{L^\infty} \leq C \| \cdot \|_{H^{s-1/2}} \) for \( s > 1 \).
We now show how to prove a linear version of Theorem 1.3. The idea is as follows. Assume that we are given long-wavelength initial conditions \( \phi_1 \) and \( \phi_2 \) for (12). We show that by an appropriate reparameterization of the initial surface, the resulting solution of (12) can be well approximated by solutions of the linearized KdV equations and that the initial conditions of the KdV equations are simply given in terms of \( \Phi_1 \) and \( \Phi_2 \). More precisely, we have:

**Lemma 2.6**

Assume that the hypotheses of Theorem 1.3 hold. For all \( C_1 > 0 \) there exist \( \varepsilon_0 > 0, C_2 > 0 \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) the following holds. Assume initial conditions

\[
\phi_0(x_1) = 0, \quad \phi_1(x_1) = \varepsilon^2 \Phi_1(\varepsilon x_1), \quad \phi_2(x_1) = \varepsilon^2 \Phi_2(\varepsilon x_1),
\]

for (12) and let \( 2A_1|_t=0 = \Phi_1 + \Phi_2, 2A_2|_t=0 = \Phi_2 - \Phi_1 \). Then there exists a parametrization of \( \Gamma|_{t=0} = \{ \overline{X}_1|_{t=0}(\alpha), X_2|_{t=0}(\alpha) | \alpha \in \mathbb{R} \} \) with

\[
x_1 = \overline{X}_1|_{t=0}(\alpha), \quad X_2|_{t=0}(\overline{X}_1|_{t=0}^{-1}(x_1)) = \phi_1(x_1), \quad U_1|_{t=0}(\overline{X}_1|_{t=0}^{-1}(x_1)) = \phi_2(x_1),
\]

such that when the initial conditions are reexpressed in the new parameterization one has

\[
\varepsilon^{7/2} \left\| \begin{pmatrix} R_{Z_1}|_{t=0} \\ R_{X_2}|_{t=0} \\ R_{U_1}|_{t=0} \end{pmatrix} \right\|_{\mathcal{H}^*} = \left\| \begin{pmatrix} Z_1|_{t=0} - \varepsilon^2 \psi Z_1|_{t=0} \\ X_2|_{t=0} - \varepsilon^2 \psi X_2|_{t=0} \\ U_1|_{t=0} - \varepsilon^2 \psi U_1|_{t=0} \end{pmatrix} \right\|_{\mathcal{H}^*} \leq C \varepsilon^{7/2},
\]

with \( \psi \) defined in Theorem 2.4.

**Proof.** Recall that \( \psi_{Z_1} = \psi_{X_2} \), and that \( Z_1 = \mathcal{K}_0 X_1 \approx -\partial_\alpha X_1 \). Thus we define the reparameterization by \( \partial_\alpha X_1 = -\phi_1(\alpha) \). Since \( \Phi_{1,2} \in H^{s+6}(2) \), one has \( \sup_{\alpha \in \mathbb{R}} |X_1|_{t=0}(\alpha)| \leq C \varepsilon \) for some constant \( C \) independent of \( \varepsilon \). With this definition we first note that

\[
\|X_2|_{t=0}(\cdot) - \varepsilon^2 \phi_1(\cdot)\|_{H^*} \leq \|\varepsilon^2 \Phi_1(\varepsilon X_1(\cdot)) - \varepsilon^2 \Phi(\varepsilon \cdot)\|_{H^*}
\]

\[
\leq C\|\varepsilon^3 \Phi_1(\varepsilon \cdot)\|_{H^*} \|\varepsilon X_1(\cdot)\|_{C^5} \leq C \varepsilon^{7/2}
\]

In like fashion, one finds \( \|U_1|_{t=0} - \varepsilon^2 \Phi_2(\varepsilon \cdot)\|_{H^*} \leq C \varepsilon^{7/2} \). Choosing the initial conditions for the KdV equations as specified in the hypotheses of the lemma, we see immediately from the definition of \( \psi \) that \( \|X_2|_{t=0} - \varepsilon^2 \psi X_2\|_{H^*} \leq C \varepsilon^{7/2} \), and \( \|U_1|_{t=0} - \varepsilon^2 \psi U_1\|_{H^*} \leq C \varepsilon^{7/2} \).

Finally, we observe that

\[
\|Z_1|_{t=0} - \varepsilon^2 \psi Z_1\|_{H^*} = \|\mathcal{K}_0 X_1|_{t=0} - \varepsilon^2 \psi X_2\|_{H^*}
\]

\[
= \|(\mathcal{K}_0 + \partial_\alpha) X_1|_{t=0} - (\partial_\alpha X_1|_{t=0} + \phi_1) + (\phi_1 - \varepsilon^2 \psi X_2)\|_{H^*} \leq C \varepsilon^{7/2},
\]
where the last estimate used Lemma 2.3 to estimate \((K_0 + \partial_\alpha)X_1\), (20) to estimate \((\phi_1 - \varepsilon^2\psi_{X_2})\), and the fact that \(\partial_\alpha X_1|_{t=0} = -\phi_1\) by definition.

Thus, in the long wave limit the initial conditions \(\phi_1\) and \(\phi_2\) can be separated in a simple way into \(A_1|_{t=0}\) and \(A_2|_{t=0}\) with an error \(\varepsilon^{7/2}R_{t=0}\) such that Theorem 2.4 can be applied. So the linearized water wave problem (16) in the long wave limit can be described by the linearized KdV–equations (18).

**Remark 2.7** Since Lemma 2.6 concerns only the initial conditions, it applies equally well to both the linear and the nonlinear problems.

It remains to control the Euler variables. Given Theorem 2.4 and Lemma 2.5, or Theorem 6.1 and Theorem 1.3, the estimate relating the Lagrangian variables to the Eulerian ones is the same in either the linear or nonlinear cases so we make no distinction between them.

**Proof of Corollary 1.2 and Corollary 1.5.**

We obtain

\[
X_1(\alpha) = X_1|_{\alpha=0} + \int_0^\alpha Z_1(\alpha')d\alpha'
\]

\[
= X_1|_{\alpha=0} + \int_0^\alpha (\varepsilon^2\psi_{Z_1} + \varepsilon^{7/2}R_{Z_1})d\alpha'
\]

a) Lemma 2.5 yields

\[
\sup_{t \in [0,T_1/\varepsilon^2]} |X_1|_{\alpha=0} \leq C\varepsilon^{1/2}.
\]

b) Exactly as in the proof of Lemma 2.5 it follows that

\[
\sup_{\alpha \in \mathbb{R}} \int_0^\alpha \varepsilon^2\psi_{Z_1}d\alpha' \leq C\varepsilon.
\]

c) Next we have

\[
\sup_{t \in [0,T_0/\varepsilon^3]} \sup_{\alpha \in [-2T_0/\varepsilon^3,2T_0/\varepsilon^3]} \int_0^\alpha \varepsilon^{7/2}R_{Z_1}d\alpha' \leq C\varepsilon^{1/2}.
\]

Summarizing the estimates a) – c) shows

\[
\sup_{t \in [0,T_0/\varepsilon^3]} \sup_{\alpha \in [-2T_0/\varepsilon^3,2T_0/\varepsilon^3]} \max_{j=0,...,s} |\partial_\alpha^j (\check{X}_1(\alpha) - \alpha)| \leq C\varepsilon^{1/2}.
\]  

(21)

The Eulerian variables \(u_1\) and \(\eta\) can now be estimated as follows

\[
\sup_{t \in [0,T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3,2T_0/\varepsilon^3]} |u_1(x_1) - \varepsilon^2\psi_{u_1}(x_1)|
\]

19
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |U_1(\bar{X}_1^{-1}(x_1)) - \varepsilon^2 \psi_{U_1}(x_1)|
\]
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |\varepsilon^2 \psi_{U_1}(\bar{X}_1^{-1}(x_1)) - \varepsilon^2 \psi_{U_1}(x_1) + \varepsilon^{7/2} R_{U_1}(\bar{X}_1^{-1}(x_1))|
\]
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} \left( |\varepsilon^2 \psi_{U_1}(\bar{X}_1^{-1}(x_1)) - \varepsilon^2 \psi_{U_1}(x_1)| + |\varepsilon^{7/2} R_{U_1}(\bar{X}_1^{-1}(x_1))| \right)
\]
\[
\leq C \varepsilon^{5/2} + C \varepsilon^{7/2}
\]
and
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |u_1(x_1) - \varepsilon^2 \psi_{U_1}(x_1)|
\]
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |U_1(\bar{X}_1^{-1}(x_1)) - \varepsilon^2 \psi_{U_1}(x_1)|
\]
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |\varepsilon^2 \psi_{U_1}(\bar{X}_1^{-1}(x_1)) - \varepsilon^2 \psi_{U_1}(x_1) + \varepsilon^{7/2} R_{U_1}(\bar{X}_1^{-1}(x_1))|
\]
\[
\leq \sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} \left( |\varepsilon^2 \psi_{U_1}(\bar{X}_1^{-1}(x_1))| + |\varepsilon^2 \psi_{U_1}(x_1)| + |\varepsilon^{7/2} R_{U_1}(\bar{X}_1^{-1}(x_1))| \right)
\]
\[
\leq C \varepsilon^6 + C \varepsilon^{7/2},
\]

where the last estimate holds due to the spatial concentration of the \(A_j \in H^{s+6}(2)\), i.e. \(\sup_{t \in [0, T_0/\varepsilon^3]} \sup_{x_1 \in [-2T_0/\varepsilon^3, 2T_0/\varepsilon^3]} |1 + (\varepsilon(x - t)^2)_{-1}^{-1}| = O(\varepsilon^5)\) and Sobolev’s embedding theorem used here as \(\|S_\varepsilon \cdot \|_{C_0^{s-1}} \leq \| \cdot \|_{C_0^{s-1}} \leq \| \cdot \|_{H^s}\). From which we see that there is no loss in \(\varepsilon\) due to the scaling. The variable \(\eta\) and all derivatives of \(u_1\) and \(\eta\) can be estimated in the same way.

Thus in order to conclude the validity of Theorem 1.1 and Theorem 1.3 it remains to transfer Theorem 2.4 from the linear to the nonlinear case, namely to prove Theorem 6.1.

The idea to show the approximation property in the nonlinear case is as follows. Suppose we have a system of the form

\[
\partial_t \mathcal{W} = \Lambda \mathcal{W} + B(\mathcal{W}, \mathcal{W}),
\]

(22)

where \(\Lambda\) is a linear operator which leads to a uniformly bounded semigroup \(e^{\Lambda t}\), and \(B(\mathcal{W}, \mathcal{W})\) is a symmetric bilinear mapping which represents the nonlinear terms. Furthermore, suppose that we have constructed an approximation \(\varepsilon^2 \Psi\) (with \(\Psi \approx O(1)\)), for the essential variables \(Z_1, X_2\) and \(U_1\) of the water wave problem. The error \(\varepsilon^\beta R\) made by the approximation \(\varepsilon^2 \Psi\) satisfies the differential equation

\[
\partial_t R = \Lambda R + 2\varepsilon \gamma B(\Psi, R) + \varepsilon^\beta B(R, R) + \varepsilon^{-\beta} \text{Res}(\varepsilon^2 \Psi),
\]

where \(\text{Res}(\mathcal{W}) = -\partial_t \mathcal{W} + \Lambda \mathcal{W} + B(\mathcal{W}, \mathcal{W})\) and \(\gamma = 2\). As already explained, we have to show that \(R\) is \(O(1)\)--bounded on an \(O(1/\varepsilon^3)\)--time scale for some \(\beta > 2\). In [KSM92] it
was pointed out that this can be proved by a simple application of Gronwall’s inequality provided \( \gamma = 3 \) and \( \varepsilon^{-\beta} \text{Res}(\varepsilon^2 \Psi) = \mathcal{O}(\varepsilon^3) \) for a \( \beta > 3 \).

In what follows, we address the following points.

a) A priori we only have \( \gamma = 2 \). So, the main difficulty consists in the long time existence of the solutions. Since we have essentially solutions of amplitude \( \mathcal{O}(\varepsilon^2) \) a simple application of Gronwall’s inequality only would give existence on a time interval of length \( \mathcal{O}(1/\varepsilon^2) \), and not on \( \mathcal{O}(1/\varepsilon^3) \). To circumvent this difficulty we take advantage of the fact that \( \Psi \) can be chosen to be more regular than \( R \) and a function of \( \alpha = \varepsilon \alpha \).

We then can use the derivative like structure in the nonlinearity in the long wave limit to estimate \( \varepsilon^2 B(\Psi, R) \) to be of order \( \mathcal{O}(\varepsilon^3) \) in a suitable chosen norm. The construction of this norm is the main part of Section 6. The difficulty for the water wave problem comes from the fact that in addition to this requirement the norm has to be chosen in such a way that the quasilinear terms cancel for local existence and uniqueness reasons. (See Remark 4.3.)

b) In \( \varepsilon^\beta B(R, R) \) we cannot use the derivative like structure in the nonlinearity in the long wave limit since this would correspond to a loss of regularity which cannot be allowed due to the fact that the linear semigroup \( e^{\lambda t} \) is not smoothing. Thus, we need \( \beta \) sufficiently big a priori, and there is no advantage gained from starting with (8) in the long wave limit scaling.

c) The restriction on \( \beta \) comes from the requirement \( \text{Res}(\varepsilon^2 \Psi) = \mathcal{O}(\varepsilon^{3+\beta}) \). In general this can be achieved by constructing higher order approximations to the solution. This does not affect the approximation to \( \mathcal{O}(\varepsilon^2) \). The precise form of these additional terms is given in (52).

\section{The operator \( \mathcal{K}(X) \)}

Before we start to construct a suitable quasilinear system out of (8), and before we consider the long wave limit in the nonlinear case we need more information about the operator \( \mathcal{K}(X) \). We prove that this operator can be represented by \( \mathcal{K}(X) = \mathcal{Q}(X) \circ \mathcal{K}_0 \circ \mathcal{Q}^{-1}(X) \), where \( \mathcal{Q} \) is a Riemann–mapping from the domain \( \Omega(t) \) to a flat domain and where \( \mathcal{K}_0 \) is a multiplication operator in Fourier space. After this we expand \( \mathcal{K}(X) \) with respect to \( X \). Finally, we prove that \( \mathcal{K}(X) = \mathcal{K}_0 + S_1(X) \), where the operator \( S_1(X) \) has certain smoothing properties. In particular, it maps \( H^s \) into \( H^{s+1/2} \). The results in this section are not new. They can already be found in the literature (cf. [Yo83], [Cr85]). However, our focus is somewhat different than these previous works.
3.1 The Riemann mapping

Lemma 3.1
There exists a conformal mapping

$$\Phi = (\Phi_1, \Phi_2) : \Omega(t) \to P^- = \{(x_1, x_2) \mid x_1 \in \mathbb{R}, -1 \leq x_2 \leq 0\}$$

with $\Phi_2((\alpha, -1)) = -1$ and $\Phi_2(\bar{X}_1(\alpha), \bar{X}_2(\alpha)) = 0$ for all $\alpha \in \mathbb{R}$. It is unique up to adding constants to $\Phi_1$.

Proof. This follows directly from Riemann’s mapping theorem. See the proof of Lemma 3.2 for a special case.

Given $\Phi$ we define a mapping

$$h(\alpha, t) = \Phi_1(\bar{X}(\alpha, t)): \mathbb{R}^1 \to \mathbb{R}^1$$

and a mapping $Q(X)$ as $Q(X)f = f \circ h$. The function $h$ can be expressed explicitly in terms of $X_1$ and $X_2$ and it is only unique up to a constant.

For our purposes it is sufficient to assume that the top surface $\Gamma$ can be parametrized by $\{(x_1, \eta(x_1)) \mid x_1 \in \mathbb{R}\}$. Then we have the following lemma.

Lemma 3.2
If the free surface is parameterized as $(x_1, \eta(x_1))$, the function $h^{-1}$ is given by

$$h^{-1}(x_1) = x_1 + N(x_1) - F^{-1} \left( \frac{k \cosh k - \sinh k}{i k \sinh k} \eta(k) \right)(x_1),$$

where $N(x) = \int_{-\infty}^{x} \eta(\xi) d\xi$. i.e. $h^{-1}$ is unique up to a constant.

Remark 3.3 For the moment we are interested in the dependence of $h$ on the configuration $(X_1, X_2)$ and so we suppress its dependence on $t$ to avoid cluttering the notation.

Proof. With the notation from above we construct the inverse Riemann–mapping $\Phi^{-1} : P^- \to \Omega(t)$ as follows. We solve the Cauchy–Riemann differential equations

$$\partial_{x_1} u + \partial_{x_2} v = 0, \quad \partial_{x_1} v - \partial_{x_2} u = 0$$

with the boundary conditions $v(x_1, 0) = \eta(x_1)$ and $v(x_1, -1) = -1$. Then the function $h^{-1}$ is defined (up to a constant) by

$$h^{-1}(x_1) = u(x_1, 0).$$
The function \( \psi \) satisfies \( \Delta \psi = 0 \) in \( P^- \) with the boundary conditions \( \psi(x_1,0) = \eta(x_1) \) and \( \psi(x_1,-1) = -1 \). We make the ansatz \( \psi(x_1,x_2) = -x_2 + \psi_1(x_1,x_2) \) and obtain that the Fourier transform of \( \psi_1 \) has to satisfy

\[
-k^2 \hat{\psi}_1 + \partial_{x_2}^2 \hat{\psi}_1 = 0, \quad \hat{\psi}_1(k,-1) = 0, \quad \hat{\psi}_1(k,0) = \hat{\eta}(k).
\]

The last system possesses the solution

\[
\hat{\psi}_1(k,x_2) = \hat{c}_k \sinh(k(x_2 + 1)) \quad \text{with} \quad \hat{c}_k = \frac{\hat{\eta}(k)}{\sinh(k)}.
\]

For the conjugate function \( u \) this yields

\[
u(x_1,x_2) = x_1 + \mathcal{N}(x_1) + u_1(x_1,x_2) \quad \text{with} \quad \hat{u}_1(k,x_2) = -\frac{k \cosh(k(x_2 + 1)) - \sinh(k)}{ik \sinh(k)} \hat{\eta}(k).
\]

Finally the function \( h^{-1} \) is defined by

\[
h^{-1}(x_1) = u(x_1,0) = x_1 + \mathcal{N}(x_1) - \mathcal{F}^{-1}\left(\frac{k \cosh(k) - \sinh(k)}{ik \sinh(k)} \hat{\eta}(k)\right).
\]

\[\blacksquare\]

**Corollary 3.4**

Fix \( s \geq 1 \). Then there exist \( C_1, C_2 > 0 \) such that for \( \eta \in H^s \) with \( \|\eta\|_{H^s} \leq C_1 \) the function \( h^{-1} \) is invertible and satisfies the estimates

\[
\|\partial_{x_1} h^{-1} - 1\|_{H^{s-1}} \leq C_2 \|\eta\|_{H^s} \quad \text{and} \quad \|\partial_{x_1} h^{-1} - 1\|_{H^{s-1}} \leq C_2 \|\eta\|_{H^s}.
\]

**Proof.** This follows immediately from

\[
\partial_{x_1} h^{-1}(x_1) = 1 - \mathcal{F}\left(\frac{k \cosh k}{\sinh k} \hat{\eta}(k)\right)
\]

\[\blacksquare\]

Now assume that the top surface is parametrized by \( \{(X_1(\alpha), X_2(\alpha)) | \alpha \in \mathbb{R}\} \). A comparison with the above parametrization yields

\[
x_1 = \tilde{X}_1(\alpha) \quad \text{and} \quad \eta(x_1) = \tilde{X}_2(\alpha),
\]

i.e. \( \eta(x_1) = \tilde{X}_2(\tilde{X}_1^{-1}(x_1)) \).

### 3.2 The representation

The mapping \( \mathcal{Q}(X) \) can now be used to find a useful representation of the operator \( \mathcal{K}(X) \).
Lemma 3.5
The operator $\mathcal{K}(X)$ can be expressed as $\mathcal{K}(X) = \mathcal{Q}(X) \circ \mathcal{K}_0 \circ \mathcal{Q}(X)^{-1}$, where $\mathcal{K}_0$ is a multiplication operator in Fourier space defined by $(\mathcal{K}_0 u)(k) = -i(\tanh k) \hat{u}(k)$.

Proof. a) Define $f(x_1, x_2) = u_1(x_1, x_2) + i u_2(x_1, x_2)$ for $(x_1, x_2) \in \Omega(t)$. Then $f$ is an analytic function of $x = x_1 + ix_2$, because of incompressibility (2) and the irrotational nature (3) of the flow. Let $\Phi$ be the conformal transformation of Lemma 3.1. Then $g(x') = f(\Phi^{-1}(x'))$, with $x' = x_1' + ix_2'$, is an analytic function on $P^-$. Let $\bar{u}_1(x') = \Re g(x')$ and $\bar{u}_2(x') = \Im g(x')$. Then, by definition we have

$$\bar{u}_2(x', 0) = \mathcal{K}_0 \bar{u}_1(x', 0),$$

where $\mathcal{K}_0$ is the operator for the flat region $P^-$, i.e. $u_2 = \mathcal{Q}(X) \circ \mathcal{K}_0 \circ \mathcal{Q}(X)^{-1} u_1$.

b) The operator $\mathcal{K}_0$ is defined by $u_2|_{x_2=0} = \mathcal{K}_0 u_1|_{x_2=0}$, where $u_1$ and $u_2$ are solutions of

$$\partial_{x_1} u_1 + \partial_{x_2} u_2 = 0, \quad \partial_{x_2} u_1 - \partial_{x_1} u_2 = 0$$

for $(x_1, x_2) \in P^-$, with $u_2|_{x_2=-1} = 0$ and $u_1|_{x_2=0}$ known. We introduce a potential $\varphi$ with $u = (u_1, u_2) = \nabla \varphi$. The Fourier transform $\hat{\varphi}$ satisfies

$$-k^2 \hat{\varphi}(k, x_2) + \partial_{x_2}^2 \hat{\varphi}(k, x_2) = 0$$

with the boundary conditions $\partial_{x_2} \hat{\varphi}|_{x_2=-1} = 0$. We obtain solutions

$$\hat{\varphi}(k, x_2) = c_k \cosh (k(1 + x_2)),$$

and so $\bar{u}_1|_{x_2=0}(k) = i k c_k \cosh (k)$ and $\bar{u}_2|_{x_2=0}(k) = k c_k \sinh (k)$. Therefore, we have

$$\bar{u}_2|_{x_2=0}(k) = -i(\tanh k) \bar{u}_1|_{x_2=0}(k),$$

and so $\mathcal{K}_0$ is a multiplication operator in Fourier space with multiplier $\hat{\mathcal{K}}_0(k) = -i \tanh (k)$.

We next note that the nonuniqueness in the definition of $h$ does not affect the operator $\mathcal{K}(X)$.

Corollary 3.6
The operator $\mathcal{K}(X)$ is invariant under transformations $h(x) \mapsto h(x) + a$ for constants $a \in \mathbb{R}$.

Proof. We have

$$(\mathcal{K}(X) f)(\alpha) = \int \mathcal{K}_0(h(\alpha) - \alpha') f(h^{-1}(\alpha')) d\alpha' = \int \mathcal{K}_0(h(\alpha) - h(\alpha')) f(\alpha') h'(\alpha') d\alpha'$$

from which invariance under the above transformation follows immediately.

Remark 3.7 In $x$-space we have

$$(\mathcal{K}_0 u)(x) = -\frac{1}{2\pi} \int \frac{u(y)}{\sinh (x - y)} dy.$$
3.3 The formal expansion

Lemma 3.8

The operator $\mathcal{K}(X)$ possesses the following expansion

\[ \mathcal{K}(X)U = \mathcal{K}_0 U + B_1(X)U + S_2(X)U. \]

with

\[ B_1(X)U = [X_1, \mathcal{K}_0] \partial_a U - (X_2 + \mathcal{K}_0(X_2 \mathcal{K}_0)) \partial_a U \]
\[ S_2(X)U = \mathcal{O}(X^2)U. \]

Proof. This formal expansion can be found in [Cr85, Lemma 3.7: page 827].

In the following we show that commutators of the form $[a, \mathcal{K}_0]$, as they appear in this expansion with $a$ a smooth function has certain smoothing properties.

Remark 3.9 We have

\[ (X_2 + \mathcal{K}_0(X_2 \mathcal{K}_0)) \partial_a U = (1 + \mathcal{K}_0^2)(X_2 \partial_a U_1) + (1 - \mathcal{K}_0^2(X_2) + \mathcal{K}_0(X_2 \mathcal{K}_0)) \partial_a U \]
\[ = (1 + \mathcal{K}_0^2)(X_2 \partial_a U) + \mathcal{K}_0(\partial_a U), \]

i.e also the second term in $B_1(X)$ can be written as a commutator plus some term which is infinitely smoothing since $\mathcal{F}(1 + \mathcal{K}_0^2)(k) = \mathcal{O}(\exp(-|k|))$. This is also possible in all terms which occur in higher order expansions.

As explained in Section 2 the variable $X_1$ will be replaced by the variable $Z_1 = \mathcal{K}_0 X_1$. This is possible since the commutators can be written as smooth functions of $Z_1$ and so we have useful estimates in terms of $Z_1$ instead of $X_1$. See also Lemma 3.14. In detail, we have

\[ \mathcal{F}[X_1, \mathcal{K}_0]V_1 = \int \dot{X}_1(k - m) \dot{\mathcal{K}}_0(m) \dot{V}_1(m) - \dot{\mathcal{K}}_0(k) \dot{X}_1(k - m) \dot{V}_1(m) dm \]
\[ = \int (\dot{\mathcal{K}}_0(m) - \dot{\mathcal{K}}_0(k)) \dot{\dot{X}}_1(k - m) \dot{V}_1(m) dm \]
\[ = \int \frac{\dot{\mathcal{K}}_0(m) - \dot{\mathcal{K}}_0(k)}{\mathcal{K}_0(k-m)} \mathcal{K}_0(k-m) \dot{X}_1(k - m) \dot{V}_1(m) dm \]
\[ = \int \tilde{m}_1(k, k - m, m) \dot{Z}_1(k - m) \dot{V}_1(m) dm, \]

where $\tilde{m}_1(k, k - m, m)$ is smooth due to the Lipschitz-continuity of $\mathcal{K}_0$. In case of $\dot{X}_1$ a distribution the last formula can be written as

\[ \dot{X}_1[\dot{\mathcal{K}}_0(k - \cdot) \dot{V}_1(k - \cdot)] - \dot{\mathcal{K}}_0 \dot{X}_1[\dot{V}_1(k - \cdot)] \]
\[ = \dot{X}_1[\dot{\mathcal{K}}_0(\cdot) \left( \frac{\dot{\mathcal{K}}_0(k - \cdot) - \dot{\mathcal{K}}_0(k)}{\mathcal{K}_0(\cdot)} \right) V_m(k - \cdot)] \]
Definition 3.10
The operator $\mathcal{M}_1(Z_1, \cdot)$ is defined by
\[ \mathcal{M}_1(Z_1, \cdot) = [X_1, \mathcal{K}_0], \]
where $Z_1 = \mathcal{K}_0 X_1$

As a consequence of $\sup_{k, m \in \mathbb{R}} |\hat{m}(k, k - m, m)| < \infty$, $\mathcal{M}_1(Z_1, \cdot)$ is a linear operator which maps $H^s$ into $H^s$ as long as $Z_1 \in H^s$ for $s > 1/2$. This estimate will be improved in the next subsection. Therefore, we suppress the details for the moment.

In order to express the term $\partial_\alpha X_1$ in terms of $Z_1$ we define additionally the operator
\[ \mathcal{M}_2 = -\partial_\alpha (\mathcal{K}_0)^{-1}. \]

By looking at its symbol it is easy to see that this operator is well defined as a map from $H^{s+1}$ to $H^s$. We abuse notation and write $B_1(\mathcal{W})$ and $S_1(\mathcal{W})$ instead of $B_1(X)$ and $S_1(X)$. With these preparations we reformulate Lemma 3.8.

Lemma 3.11
The operator $\mathcal{K}(\mathcal{W})$ possesses the following expansion
\[ \mathcal{K}(\mathcal{W})U_1 = \mathcal{K}_0 U_1 + B_1(\mathcal{W}) U_1 + S_2(\mathcal{W}) U_1. \]

with
\[ B_1(\mathcal{W}) U_1 = \mathcal{M}_1(Z_1, \partial_\alpha U_1) - (X_2 + \mathcal{K}_0 (X_2 \mathcal{K}_0)) \partial_\alpha U_1 \]
\[ S_2(\mathcal{W}) U_1 = O(\|\mathcal{W}\|^2) U_1. \]

3.4 Analytic properties of the expansion

Beside these formal calculations we need estimates for the terms appearing in this expansion. First we recall a well known estimate for commutators.

Lemma 3.12
Let $r \geq 0$, $q > 1/2$ and $0 \leq p \leq q$. Then there exists a $C > 0$ such that
\[ \| [a, \mathcal{K}_0] u \|_{H^r} \leq C \| a \|_{H^{r+p}} \| u \|_{H^{q-p}}. \]

Proof. In Fourier space the term $[a, \mathcal{K}_0] u$ can be written as
\[ \int (1 + k^2)^r \left( \int (\hat{a}(k-m) \hat{\mathcal{K}}_0(m) - \hat{\mathcal{K}}_0(k) \hat{a}(k-m)) \hat{u}(m) dm \right)^2 dk \]
\[ \leq \left( \sup_{k, m \in \mathbb{R}} \frac{(1 + k^2)^{r/2} (\hat{\mathcal{K}}_0(m) - \hat{\mathcal{K}}_0(k)) (1 + m^2)^{(p-q)/2+s}}{(1 + (k-m)^2)^{(r+p)/2}} \right)^2 \|a\|_{H^{r+p}}^2 \|u\|^2_{H^{q-p}}, \]
for any $s > 1/4$. From the exponential convergence of $\hat{K}(k)$ towards $\pm 1$ for $k \to \pm \infty$ the boundedness of the supremum follows. In more detail, consider first the case in which $k$ and $m$ have the same sign. Then the supremum can be bounded by

$$
\sup_{k, m \in \mathbb{R}^+} \left| \frac{1 + k^2)^{r/2}(\hat{K}_0(m) - \hat{K}_0(k))(1 + m^2)(p-q)/(r+s)}{(1 + (k - m)^2)^{(r+p)/2}} \right| \leq C(r, p),
$$

where we choose $s < q/2$. If $k$ and $m$ have different signs we bound the factor $|\hat{K}_0(m) - \hat{K}_0(k)|$ by 2, and the supremum is bounded by

$$
2 \sup_{k \in \mathbb{R}^+, m \in \mathbb{R}^-} \left| \frac{(1 + k^2)^{r/2}(1 + m^2)(p-q)/(r+s)}{(1 + (k - m)^2)^{(r+p)/2}} \right| \leq 2 \sup_{k, m \in \mathbb{R}^+} \left| \frac{(1 + k^2)^{r/2}(1 + l^2)(p-q)/(r+s)}{(1 + (k + l)^2)^{(r+p)/2}} \right| \leq C(r, p),
$$

where again we use $1/4 < s < q/2$. □

In a very similar fashion, we obtain:

**Corollary 3.13**

Let $r \geq 0$, $q > 1/2$ and $0 \leq p \leq q$. Then there exists a $C > 0$ such that

$$
\|\mathcal{M}_1(a, u)\|_{H^r} \leq C \|a\|_{H^{r+p}} \|u\|_{H^{q-p}}.
$$

**Proof.** Using the expression derived above for the Fourier transform of $[X_1, \mathcal{K}_0]u$, we see that the square of the $H^r$ norm of $\mathcal{M}_1(a, u)$ is bounded by

$$
\int (1 + k^2)^r \left( \int \frac{\hat{K}_0(m) - \hat{K}_0(k)}{\hat{K}_0(k - m)} \hat{a}(k - m) \hat{\mu}(m) dm \right)^2 dk
\leq \left( \sup_{k, m \in \mathbb{R}} \left| \frac{(1 + k^2)^{r/2}(\hat{K}_0(m) - \hat{K}_0(k))(1 + m^2)(p-q)/(r+s)}{\hat{K}_0(k - m)(1 + (k - m)^2)^{(r+p)/2}} \right| \right)^2 \|a\|^2_{H^{r+p}} \|u\|^2_{H^{q-p}}
$$

for $s > 1/4$. The boundedness of the supremum then follows just as in Lemma (3.12). □

As a direct consequence of this corollary it follows that the first order expansion $B_1(X)$ of the operator $\mathcal{K}(X)$ is smoothing. We prove now that the same is true for $\mathcal{S}_1(W) = \mathcal{K}(W) - \mathcal{K}_0$. In order to do so we recall that we defined the space $H^s = H^s \times H^s \times H^{s-1/2}$. The following Lemma improves results given for instance in [BHL93].
Lemma 3.14
Fix \( s \geq 7/2 \). Write \( h(x) = x + g(x) \). Then there exist \( C_1 \in (0, \frac{1}{2}) \) and \( C_s > 0 \) such that for \( \|\partial_\alpha g\|_{H^1} \leq C_1 \) and \( \|\partial_\alpha g\|_{H^{s-1}} \leq C_s \) the operator \( S_1(W) = K(W) - K_0 \) satisfies
\[
\|S_1(W)U\|_{H^s} \leq C\|W\|_{H^s}\|U\|_{H^s}.
\]

Proof. We use Lemma 3.5 and represent the difference as
\[
K(X) - K_0 = Q \circ K_0 \circ Q^{-1} - K_0 \circ Q \circ Q^{-1} = (Q \circ K_0 - K_0 \circ Q) \circ Q^{-1}.
\]
Since \( (Q^{-1}U)(\alpha) = U(h^{-1}(\alpha)) = U(\alpha) + \partial_\alpha U \cdot (h^{-1}(\alpha))' + O(h^2) \) the assertion obviously follows from
\[
\|(K_0 \circ Q - Q \circ K_0)U\|_{H^s} \leq C\|W\|_{H^s}\|U\|_{H^s}.
\]
In Fourier space we use the representation
\[
\mathcal{F}[(K_0 \circ Q - Q \circ K_0)f] = \int \int \hat{K}_0(k)e^{ikx}e^{-ith(x)}\hat{f}(l)dl \, dx - \int \int \hat{e}^{ikx}e^{-ith(x)}\hat{K}_0(l)\hat{f}(l)dl \, dx
\]
\[
= \int \int (\hat{K}_0(k) - \hat{K}_0(l))e^{ikx}e^{-ith(x)}\hat{f}(l)dl \, dx
\]
\[
= \int \int (\hat{K}_0(k) - \hat{K}_0(l))e^{i(k-l)x}e^{-it\sigma(x)}\hat{f}(l)dl \, dx
\]
\[
= \int \int \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k-l)}e^{i(k-l)x}e^{-it\sigma(x)}\hat{f}(l)dl \, dx
\]
\[
= \int \int \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k-l)}e^{i(k-l)x}ilg'(x)e^{-it\sigma(x)}\hat{f}(l)dl \, dx
\]
\[
= \int \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k-l)}H(l,k-l)il\hat{f}(l)dl.
\]
In order to estimate the last line we need three preparatory steps.

a) We improve the estimate used in Lemma 3.12 slightly
\[
\left|\frac{\hat{K}_0(k) - \hat{K}_0(l)}{i(k-l)}\right| \leq 2\min(1,(k-l)^{-1})\begin{cases} 1 & \text{for } k, l \text{ different sign} \\ \max(e^{-|k|}, e^{-|l|}) & \text{for } k, l \text{ same sign} \end{cases}
\]
b) Next we consider
\[
H(l, p) = \int e^{ipx}g'(x)e^{-it\sigma(x)}dx.
\]
By Parseval’s identity we have
\[
\int |H(l, p)|^2dp = \int |g'(x)|^2dx.
\]
c) Moreover, we use
\[
\int \left( \int a(l, p) f(l) dl \right)^2 dp = \int \left( \int a(l, p) f(l) dl \right) \left( \int a(m, p) f(m) dm \right) dp
\]
\[
= \int \int (\int a(l, p) a(m, p) dp) f(l) f(m) dm dl
\]
\[
= \int \int (\int a(l, p)^2 dp)^{1/2} (\int a(m, p)^2 dp)^{1/2} f(l) f(m) dm dl
\]
\[
\leq \sup_l (\int a(l, p)^2 dp)^{1/2} \sup_m (\int a(m, p)^2 dp)^{1/2}
\times \int |f(l)| dl \int |f(m)| dm
\]\n\[
\leq \sup_l (\int a(l, p)^2 dp) (\int |f(l)| dl)^2
\]

Then we distinguish three cases.

**Case I:** \( k, l \) same sign, \(|k| \leq |l| \). With b) and c) we obtain
\[
\left( \int |k|^s \int \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k - l)} H(l, k - l) i l \hat{f}(l) dl^2 dk \right)^{1/2}
\leq 2 \sup_k |k^s e^{-|k|}| (\int (\int |H(l, k - l) i l \hat{f}(l)| dl)^2 dk)^{1/2}
\leq C (\int (\int |H(l, p)|^2 dp)^{1/2} (\int |il \hat{f}(l)| dl)
\leq C \|g\|_{L^2} \|f\|_{H^2}.
\]

**Case II:** \( k, l \) same sign, \(|k| \geq |l| \). We have
\[
\left( \int |k|^s \int \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k - l)} H(l, k - l) i l \hat{f}(l) dl^2 dk \right)^{1/2}
\leq C (\int (\int |k - l|^s + |l|^s) \frac{(\hat{K}_0(k) - \hat{K}_0(l))}{i(k - l)} H(l, k - l) i l \hat{f}(l) dl^2 dk)^{1/2}
\leq C (\int (\int |k - l|^s e^{-|l|} |H(l, k - l) i l \hat{f}(l)| dl)^2 dk)^{1/2}
\]
\[
+ C (\int (\int |l|^s e^{-|l|} |H(l, k - l) i l \hat{f}(l)| dl)^2 dk)^{1/2}
\leq C \|g\|_{H^{-1}} \|f\|_{H^1},
\]

where we used
i)
\[
(\int (\int |k - l|^s e^{-|l|} |H(l, k - l) i l \hat{f}(l)| dl)^2 dk)^{1/2}
\]
\[
\begin{align*}
&\leq (\int (\int |p|^{s-1}e^{-|l|})(\int e^{ipx}g'(x)e^{i\sigma(x)}dx)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq (\int (\int e^{-|l|})(\int e^{ipx}D_{x}^{s-1}(g'(x)e^{i\sigma(x)}dx)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq (\int (\int e^{ipx}e^{-|l|}|g^{(s)}(x)+\ldots+(il)^{s-1}(g'(x))^s)\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq C\|g'\|_{H^{s-1}}\|f\|_{H^1},
\end{align*}
\]
by \(\sup_{l}|e^{-|l|}|^{s} \leq \infty\) and applying b) to \(g^{(s)}(x),\ldots,(il)^{s-1}(g'(x))^s\).

ii) With b) we obtained
\[
\begin{align*}
&\left(\int (\int |l|^{s}e^{-|l|}|H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq \sup_{l}|e^{-|l|}|^{s+1}\left(\int (\int |H(l,p)|^{2}dp)^{1/2} \int |i\hat{f}(l)|dl \\
&\leq C\|g'\|_{L^2}\|f\|_{H^1}.
\end{align*}
\]

**Case III:** \(k, l\) different sign and \(|k-l| \geq |l|\). Then
\[
\begin{align*}
&\left(\int |k|^{s}\int \frac{(\hat{\mathcal{K}}(k) - \hat{\mathcal{K}}(l))}{i(k-l)}H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq C\left(\int (\int |H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq C\sup_{l \in \mathbb{R}}(\int |H(l,p)|^{2}dp)^{1/2} \int \|i\hat{f}(l)|dl \\
&\leq C\|g'\|_{L^2}\|f\|_{L^2},
\end{align*}
\]
where the first inequality used the fact that if \(k\) and \(l\) have different signs and \(|k-l| < 1\) then both \(|k| < 1\) and \(|l| < 1\), while the second inequality used c).

**Case IV:** \(k, l\) different sign, i.e \(|p| \geq |l|\) and \(|k-l| > 1\). Then
\[
\begin{align*}
&\left(\int |k|^{s}\int \frac{(\hat{\mathcal{K}}(k) - \hat{\mathcal{K}}(l))}{i(k-l)}H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq \left(\int (\int |k|^{s}\frac{|k|}{|k-l|}H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq C\left(\int (\int |k|^{s}\frac{(k-l)^{s-1}}{|k-l|^{s}}H(l,k-l)i\hat{f}(l)|dl|^{2}dk)^{1/2} \\
&\leq C\sup_{l}(\int \frac{P^{s-1}}{(1+l^{2})^{1/2}}H(l,p)|^{2}dp)^{1/2} \int |i\hat{f}(l)|^{1/2}\hat{f}(l)|dl \\
&\leq C\sup_{l}(\int \frac{P^{s-1}}{(1+l^{2})^{1/2}}H(l,p)|^{2}dp)^{1/2}\|f\|_{H^3}
\end{align*}
\]

30
We now consider the term

\[
p^{s-1}H(l,p) = \int p^{s-1}e^{ip\lambda}g'(x)e^{-il\lambda(x)}dx
\]

\[
= \int p^{s-2}\left(\frac{-p}{ip - il\lambda'(x)}g'(x)e^{ip\lambda(x)}\right)dx
\]

\[
= \int p^{s-2}\left(\frac{-p}{ip - il\lambda'(x)}g'(x)e^{ip\lambda(x)}\right)dx
\]

\[
- (\frac{p}{ip - il\lambda'(x)})g''(x)e^{ip\lambda(x)}dx
\]

\[
= \ldots
\]

\[
= \int p^{s-3}\left(\frac{-p}{ip - il\lambda'(x)}\right)g'(x)
\]

\[
+ \ldots + (\frac{p}{ip - il\lambda'(x)})^{s-1}g^{(s)}(x)e^{ip\lambda(x)}dx.
\]

We must show in this expansion has bounded \(L^2\)-norm (when considered as a function of \(p\)). We consider explicitly the first and the last terms which have been written out in full. The intermediate terms are handled in an identical fashion.

We first observe that \(\|g'\|_{L^\infty} \leq 2C_1 < 1\) and that \(\left|\frac{p}{ip - il\lambda'(x)}\right|_{L^\infty} \leq C\), where \(C\) is independent of \(p\) and \(l\). (Recall \(p = k - l\) and that \(k\) and \(l\) have different signs.) Similarly, \(\|g^{(r)}\|_{L^\infty} \leq 2C_s\) for \(1 \leq r \leq s - 1\).

Thus, we can bound

\[
|((\partial_x^{s-1})\left(\frac{-1}{ip - il\lambda'(x)}\right))| \leq C \sum_{r=1}^{s} \frac{|g^{(r)}(x)|}{|p - l\lambda'(x)|} \leq \tilde{C} \sum_{r=1}^{s} \frac{|g^{(r)}(x)|}{(1 + p^2)^{1/2}},
\]

where the constants \(C\) and \(\tilde{C}\) depend on \(s\) and \(\|\partial g\|_{H^{s-1}}\), but not on \(p\) or \(l\) and the last inequality used the fact that \(k\) and \(l\) have different signs and \(|k - l| > 1\), i.e. \(|p - l\lambda'(x)| \geq \tilde{C}(1 + p^2)^{\frac{1}{2}}\). But with (25) we see that

\[
\sup_l\left(\int \frac{1}{1 + l^2}\left|\int ((\partial_x^{s-1})\left(\frac{-1}{ip - il\lambda'(x)}\right))g'(x)|dx|^2dp\right)^{\frac{1}{2}}\right)
\]

\[
\leq \tilde{C}(\int \frac{1}{1 + l^2}\left(\int \sum_{r=1}^{s} \frac{|g^{(r)}(x)|}{(1 + p^2)^{1/2}}|g'(x)|dx|^2dp\right)^{\frac{1}{2}}
\]

\[
\leq \tilde{C}(\int \frac{1}{1 + p^2}\sum_{r=1}^{s}(\|g^{(r)}\|_{L^2}\|g\|_{L^2})dp)^{\frac{1}{2}}
\]

\[
\leq C\|\partial g\|_{H^{s-1}}.
\]
Finally, we consider
\[
\int \left( \frac{p}{ip - ig'(x)} \right)^{s-1} g(s)(x)e^{ipx - itg(x)}dx.
\]
Begin by writing
\[
\left( \frac{p}{ip - ig'(x)} \right)^{s-1} = 1 + \left( \sum_{j=0}^{s-2} \left( \frac{p}{p - lg'(x)} \right)^j \left( \frac{lg'(x)}{p - lg'(x)} \right) \right)\]
Thus,
\[
\int \left( \frac{p}{ip - ig'(x)} \right)^{s-1} g(s)(x)e^{ipx - itg(x)}dx
\]
\[= (-i)^{s-1} \left( \int g(s)(x)e^{ipx - itg(x)}dx \right)
\]
\[+ \int \sum_{j=0}^{s-2} \left( \frac{p}{p - lg'(x)} \right)^j \left( \frac{lg'(x)}{p - lg'(x)} \right) g(s)(x)e^{ipx - itg(x)}dx. \]
Noting that \( \| \sum_{j=0}^{s-2} \left( \frac{p}{p - lg'(x)} \right)^j \|_{L^\infty} \leq C \), we have
\[
\sup_l \left( \int \frac{1}{1 + l^2} \left( \int \sum_{j=0}^{s-2} \left( \frac{p}{p - lg'(x)} \right)^j \left( \frac{lg'(x)}{p - lg'(x)} \right) g(s)(x)e^{ipx - itg(x)}dx \right)^2 dp \right)^{1/2}
\]
\[\leq C \int \left( \int \frac{1}{(1 + p^2)^{1/2}} |g'(x)| |g(s)(x)| dx \right)^2 dp \right)^{1/2}
\]
\[\leq C \| g' \|_{H^{s-1}}. \]
Finally, the term \( \int g(s)(x)e^{ipx - itg(x)}dx \) has \( L^2 \)-norm (with respect to \( p \)) bounded by \( \| g(s) \|_{L^2} \) by Parseval’s inequality.
Thus all terms in \( \frac{p^{s-1}H(p,l)}{(1 + p^2)^{1/2}} \) have \( L^2 \)-norm (w.r.t \( p \)) bounded by \( C \| g' \|_{H^{s-1}} \) independent of \( l \).
The final step in the proof of the lemma is to show that we can bound \( \| g' \|_{H^{s-1}} \) by \( C \| \mathcal{W} \|_{H^s} \). Note that (23) implies that
\[
\| g' \|_{H^{s-1}} \leq C \| \eta \|_{H^s}.
\]
But \( \eta(x) = X_2 \circ \bar{X}_1^{-1}(x) \) and so
\[
\| \eta \|_{L^2}^2 = \int (X_2(\bar{X}_1^{-1}(x))) dx = \int (X_2(y))^2 |1 + \partial_y X_1(y)| dy
\]
\[= \int (X_2(y))^2 |1 + M_2Z_1(y)| dy \leq C \| X_2 \|_{L^2}^2 (1 + \| M_2Z_1 \|_{L^\infty})
\]
\[\leq C \| X_2 \|_{L^2}^2 (1 + \| M_2Z_1 \|_{H^{1}}) \leq C \| X_2 \|_{L^2}^2 (1 + \| Z_1 \|_{H^2}). \]
In like fashion, we find that for all \(2 \leq r \leq s\), we can bound \(\|\eta\|_{H^r} \leq C(\|X_2\|_{H^r} + \|Z_1\|_{H^r}) \leq C\|\mathcal{W}\|_{H^r}^r\).

As already explained in section 2 the local existence and uniqueness theory for quasilinear systems does not apply directly to system (8). Therefore, in Section 4 this system is extended and a suitable quasilinear system is constructed for the original variables and another variable \(V_1 = \partial_t U_1\). We collected these variables in the extended vector \(\mathcal{W}_e = (Z_1, X_2, U_1, V_1)\) and we look for solutions of the extended system in

\[
\mathcal{H}_e^s = \mathcal{H}_s \times H^{s-1} = H^s \times H^s \times H^{s-1/2} \times H^{s-1}.
\]

In the extended system the following terms appear.

**Lemma 3.15**

Assume the situation of Lemma 3.14. Then for all \(s \geq 4\) we have

\[
\|\partial_\alpha(S_1(\mathcal{W})U_1)\|_{H^{s-1}} \leq C\|\mathcal{W}\|_{H^s}(\|U_1\|_{H^s} + \|V_1\|_{H^4}),
\]

\[
\|\partial_\alpha^2(S_1(\mathcal{W})U_1)\|_{H^{s-1}} \leq C\|\mathcal{W}_e\|_{H^s}(\|U_1\|_{H^s} + \|V_1\|_{H^4}),
\]

\[
\|\partial_\alpha(S_1(\mathcal{W})U_1)\|_{H^{s-1}} \leq C\|\mathcal{W}\|_{H^s}\|U_1\|_{H^4}.
\]

**Proof.** The first and the third line follow immediately from Lemma 3.14 and the representation of \(h\) in terms of \(X_2\) and \(X_1\).

For the second estimate we use the representation of the commutator from [Wu97]. Wu works with fluids of infinite depth but this turns out to cause only minor changes. In particular, the key lemma, Lemma 2.1 of [Wu97] is almost unchanged. Formula (2.5) is exactly the same in our case, if we merely substitute \(K_0\) for the Hilbert transform \(H\), while (2.6) of that reference is replaced by

\[
[\partial_\delta h \circ h^{-1}, K_0]f = -\text{Im}(\frac{(Q^{-1}U_1 - iQ^{-1}U_2)}{\partial_\delta(Q^{-1}X_1 - iQ^{-1}X_2)}(f - iK_0f) - K_0\text{Re}(\frac{(Q^{-1}U_1 - iQ^{-1}U_2)}{\partial_\delta(Q^{-1}X_1 - iQ^{-1}X_2)}(f - iK_0f))\] (6)

\[
-\text{Im}(\frac{(Q^{-1}U_1 - iQ^{-1}U_2)}{\partial_\delta(Q^{-1}X_1 + iQ^{-1}X_2)}f).
\]

Aside from the change of \(K_0\) for \(H\), and a different sign convention, the only significant difference with [Wu97] is the term \((K_0^2 + 1)\text{Im}(\frac{(Q^{-1}U_1 - iQ^{-1}U_2)}{\partial_\delta(Q^{-1}X_1 + iQ^{-1}X_2)}f)\), which results from the finite depth. However, this term is infinitely smoothing since the Fourier multiplier of \(1 + K_0^2\) is \(O(e^{-|k|})\). Following the calculations leading to (4.5) of [Wu97], we find that in analogy with the formula just prior to (4.5), we have:

\[
(Q^{-1}V_1, K_0)\partial_\delta Q^{-1}U_1 + (Q^{-1}V_2, K_0)\partial_\delta Q^{-1}V_2 + 4[Q^{-1}U_1, K_0]\partial_\delta Q^{-1}V_1
\]

33
\[ +4 \mathcal{Q}^{-1} U_1, \mathcal{K}_0] \delta_\beta (\text{Re} \frac{(\mathcal{Q}^{-1} U_1 - i \mathcal{Q}^{-1} U_2) \partial_\beta (\mathcal{Q}^{-1} U_1 + i \mathcal{Q}^{-1} U_2)}{\partial_\beta (\mathcal{Q}^{-1} X_1 - i \mathcal{Q}^{-1} X_2)} ) (27) \]

\[ +2 ((\mathcal{Q}^{-1} U_1)^2 - (\mathcal{K}_0 (\mathcal{Q}^{-1} U_1))^2, \mathcal{K}_0] \delta_\beta (\text{Re} \frac{\partial_\beta (\mathcal{Q}^{-1} U_1 + i \mathcal{Q}^{-1} U_2)}{\partial_\beta (\mathcal{Q}^{-1} X_1 - i \mathcal{Q}^{-1} X_2)} ) \]

\[ + \Sigma (\mathcal{W})(U_1 + V_1), \]

where \( \beta = h(\alpha) \) and the term \( \Sigma (\mathcal{W}) \) is the collection of all the contributions coming from the infinitely smoothing part of (26). Due to the smoothing properties of the commutators (Lemma 3.12), the terms on the right hand side of this expression are all bounded by \( C \| \mathcal{W}_c \|_{\mathcal{H}^s} (\| U_1 \|_{\mathcal{H}^3} + \| V_1 \|_{\mathcal{H}^3}) \). Next note that the estimates we have proved above imply that \( \partial_\beta \mathcal{Q}^{-1} X_1 \) is close to 1 and \( \partial_\beta \mathcal{Q}^{-1} X_2 \) is small. Thus, regarding the left hand side of (27) as a linear operator acting on \( [\partial_\beta^2, \mathcal{S}_1 (\mathcal{W})] U_1 \), we see that this operator is of the form identity plus something small. Inverting this operator by a Neumann series then completes the proof of this estimate.

Combining the above results about the smoothing properties of the low order expansions we obtain the following corollary.

**Corollary 3.16**

*Assume the situation of Lemma 3.14. Then for all \( s \geq 4 \) we have*

\[ \| \mathcal{S}_2 (\mathcal{W}) U \|_{\mathcal{H}^s} \leq C \| \mathcal{W} \|^2_{\mathcal{H}^s} \| U \|_{\mathcal{H}^s} \]

\[ \| \partial_1 (\mathcal{S}_2 (\mathcal{W}) U_1) \|_{\mathcal{H}^{s-1}} \leq C \| \mathcal{W} \|^2_{\mathcal{H}^s} (\| U_1 \|_{\mathcal{H}^s} + \| V_1 \|_{\mathcal{H}^s}) \]

\[ \| [\partial_\beta^2, \mathcal{S}_2 (\mathcal{W})] U_1 \|_{\mathcal{H}^{s-1}} \leq C \| \mathcal{W}_c \|^2_{\mathcal{H}^s} (\| U_1 \|_{\mathcal{H}^s} + \| V_1 \|_{\mathcal{H}^s}) \]

\[ \| \partial_\alpha (\mathcal{S}_2 (\mathcal{W}) U_1) \|_{\mathcal{H}^{s-1}} \leq C \| \mathcal{W} \|^2_{\mathcal{H}^s} \| U_1 \|_{\mathcal{H}^s}. \]

## 4 Formulation as evolutionary problem

In this section we prove the local existence and uniqueness for the solutions of (8). As already explained this cannot be achieved directly. We obtain the result by embedding (8) into a bigger system which is quasilinear in the sense of [Kat75], such that standard existence and uniqueness theory can be applied.

This bigger system for \( \mathcal{V} = (X_1, X_2, U_1, V_1) \) is first order in time and consists mainly of a nonlinear wave equation. As a general rule the coefficients \( a, b \) in a nonlinear wave equation

\[ \partial_t^2 u = a(u)L(\partial_u)u + b(u), \quad (a > 0) \]

with \( L \) an elliptic operator, can contain half the derivatives of \( L \) and terms with \( \partial_t u \). In the water wave problem (8) the operator \( L \) is given by the multiplier \( L(k) = -k \tan h \)
in Fourier space. As a consequence \(|\hat{L}(k)| \sim |k|\) for \(|k| \to \infty\). Thus, if we look for solutions \(u\) in \(H^{s-1/2}\) we can allow for coefficients \(a, b \in H^{s-1}\), i.e. the coefficients \(a\) and \(b\) can loose half a derivative. Similar to the calculations in Section 2 we then have \(\partial_t u \in H^{s-1}\).

In Section 2 we explained that we first choose \(X_{1,2} \in H^s, U_1 = \partial_t X_1 \in H^{s-1/2},\) and \(V_1 = \partial_t^2 X_1 \in H^{s-1}\). We construct the nonlinear wave equation (28) for the variable \(U_1\). And so all terms with \(\partial_a X_{1,2} \in H^{s-1}\) have the right regularity. Terms with \(\partial_a U \in H^{s-3/2}\) create difficulties. They have to be separated into a part which contributes to the quasilinear part \(a(u)L(\partial_a)u\) and a part which contributes to \(b\).

With the notation \(U_2 = \partial_t X_2\) the equations (6) and (7) become

\[
(1 + \partial_a X_1)\partial_t U_1 + \partial_a X_2(1 + \partial_t U_2) = 0
\]

\[
U_2 = K_0 U_1 + S_1(X) U_1.
\]

Differentiation of (6) with respect to time yields

\[
(1 + \partial_a X_1)\partial_t^2 U_1 + (\partial_t U_1)\partial_a U_1 + (1 + \partial_t U_2)\partial_a U_2 + (\partial_a X_2)\partial_t^2 U_2 = 0.
\]

(29)

Differentiation of (7) with respect to time and space yields

\[
\partial_a U_2 = K_0 \partial_a U_1 + \partial_a (S(X) U_1)
\]

\[
\partial_t U_2 = K_0 \partial_t U_1 + \partial_t (S_1(X) U_1)
\]

\[
\partial_t^2 U_2 = K_0 \partial_t^2 U_1 + S_1(X) \partial_t^2 U_1 + [\partial_t^2, S_1(X)] U_1
\]

Replacing \(\partial_a U_2\) and \(\partial_t^2 U_2\) in (29) gives

\[
(1 + \partial_a X_1)\partial_t^2 U_1 + \partial_t U_1 \partial_a U_1 + (1 + \partial_t U_2) (K_0 \partial_a U_1 + \partial_a (S_1(X) U_1))
\]

\[
+(\partial_a X_2) (K_0 \partial_t^2 U_1 + S_1(X) \partial_t^2 U_1 + [\partial_t^2, S_1(X)] U_1) = 0
\]

and so

\[
(f_1 + f_2(K_0 + S_1(X))) \partial_t^2 U_1 + (g_1 + g_2 K_0) K_0 \partial_a U_1 + G_1 = 0,
\]

(30)

with

\[
f_1 = (1 + \partial_a X_1)
\]

\[
f_2 = \partial_a X_2
\]

\[
g_1 = (1 + \partial_t U_2)
\]

\[
g_2 = -\partial_t U_1
\]

\[
G_1 = \partial_t U_1 (1 + K_0^2) \partial_a U_1 + (1 + \partial_t U_2) \partial_a (S_1(X) U_1)
\]

\[
+(\partial_a X_2) ([\partial_t^2, S_1(X)] U_1).
\]
We multiply (30) with \((f_1 - f_2 K_0)\) and find that the first term in the equation becomes

\[
(f_1 - f_2 K_0)(f_1 + f_2(K_0 + S_1(X)))\partial_t^2 U_1
= (f_1)^2 \partial_t^2 U_1 - f_2 K_0 (f_1 \partial_t^2 U_1) + (f_1 - f_2 K_0) f_2 S_1(X) \partial_t^2 U_1
+ f_1 f_2 K_0 \partial_t^2 U_1 - f_2 K_0 (f_2 \partial_t^2 U_1)
= (f_1)^2 \partial_t^2 U_1 - f_2 K_0 \partial_t^2 U_1 + (f_1 - f_2 K_0) f_2 S_1(X) \partial_t^2 U_1
+ f_2 K_0 [f_0, f_2] \partial_t^2 U_1 - f_2 (1 + K_0^2) (f_2 \partial_t^2 U_1) + f_2 \partial_t^2 U_1.
\]

The second term can be rewritten as:

\[
(f_1 - f_2 K_0) (g_1 + g_2 K_0) K_0 \partial_\alpha U_1
= f_1 g_1 K_0 \partial_\alpha U_1 + f_1 g_2 K_0^2 \partial_\alpha U_1
- f_2 K_0 (g_1 K_0 \partial_\alpha U_1) - f_2 K_0 (g_2 K_0^2 \partial_\alpha U_1)
= f_1 g_1 K_0 \partial_\alpha U_1 - f_2 [K_0, g_1] K_0 \partial_\alpha U_1
- f_2 [K_0, g_1] K_0 \partial_\alpha U_1 - f_2 g_2 (1 + K_0^2) K_0 \partial_\alpha U_1
+ f_2 g_2 K_0 \partial_\alpha U_1 + (f_1 g_2 - f_2 g_1) K_0^2 \partial_\alpha U_1.
\]

Therefore, (31) transforms into

\[
(f_3 + H_1) \partial_t^2 U_1 + f_4 K_0 \partial_\alpha U_1 + f_5 K_0^2 \partial_\alpha U_1 + G_2 = 0
\]

(31)

with

\[
\begin{align*}
  f_3 &= f_1^2 + f_2^2 \\
  f_4 &= f_1 g_1 + f_2 g_2 \\
  f_5 &= f_1 g_2 - f_2 g_1 \\
  H_1 &= -f_2 [K_0, f_1] \cdot f_2 K_0 [K_0, f_2] - f_2 (1 + K_0^2) (f_2 \cdot f_2) - (f_1 - f_2 K_0) f_2 S_1(X) \cdot \\
  G_2 &= (f_1 - f_2 K_0) G_1 - f_2 [K_0, g_2] K_0^2 \partial_\alpha U_1 \\
  &\quad - f_2 [K_0, g_1] K_0 \partial_\alpha U_1 - f_2 g_2 (1 + K_0^2) K_0 \partial_\alpha U_1.
\end{align*}
\]

We multiply this with \((f_3 + H_1)^{-1}\) and obtain

\[
\partial_t^2 U_1 + f_6 K_0 \partial_\alpha U_1 + (f_3 + H_1)^{-1} (f_5 K_0^2 \partial_\alpha U_1) + G_3 = 0
\]

(32)

with

\[
\begin{align*}
  f_6 &= f_4 / f_3 \\
  G_3 &= (f_3 + H_1)^{-1} G_2 + \sum_{n=1}^{\infty} (-f_3^{-1} H_1)^n (f_3^{-1} f_4 K_0 \partial_\alpha U_1).
\end{align*}
\]

36
This is more or less the standard form (28) of a nonlinear wave equation.

By Lemma 3.12 and Remark 3.9 the commutators and \((1 + \mathcal{K}_0^2)\) are smoothing from 
\(H^{1/2}\) into \(H^{s}\) for all \(s \geq 0\). As a consequence \(G_3 \in H^{s-1}\) can be handled as semilinear.

The quasilinear terms
\[ f_6 \mathcal{K}_0 \partial_x U_1 + (f_3 + H_1)^{-1}(f_5 \mathcal{K}_0^2 \partial_x U_1) \]

have coefficients \(f_j \in H^{s-1}\). Obviously the equation
\[ \partial_t^2 U_1 = f_6 \mathcal{K}_0 \partial_x U_1 \]
defines a uniformly bounded \(C^0\)-semigroup if \(f_6 > 0\), which we can insure by taking \(\varepsilon\)
sufficiently small. The term \((f_3 + H_1)^{-1}(f_5 \mathcal{K}_0^2 \partial_x U_1)\) is a quasilinear perturbation which can lead to a loss of regularity. See remark 4.1. But since \(f_5 = 0\) is another notation for (6), this term vanishes identically and we finally end up with the system.

\[
\begin{align*}
\partial_t X_1 &= U_1 \\
\partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{S}_1(X)U_1 \\
\partial_t U_1 &= V_1 \\
\partial_t V_1 &= -f_6 \mathcal{K}_0 \partial_x U_1 - G_3,
\end{align*}
\]

where

\[
\begin{align*}
f_1 &= (1 + \partial_x X_1) \\
f_2 &= \partial_x X_2 \\
f_3 &= f_1^2 + f_2^2 \\
f_4 &= f_1 g_1 + f_2 g_2 \\
f_6 &= f_4 / f_3 \\
g_1 &= (1 + V_2) \\
g_2 &= -V_1 \\
G_1 &= V_1(1 + \mathcal{K}_0^2) \partial_x U_1 + (1 + V_2) \partial_x (\mathcal{S}_1(X) U_1) \\
&\quad + (\partial_x X_2)[\partial_t^2, \mathcal{S}_1(X)] U_1 \\
G_2 &= (f_1 - f_2 \mathcal{K}_0) G_1 - f_2 [\mathcal{K}_0, g_2] \partial_x \mathcal{K}_0^2 U_1 \\
&\quad - f_2[\mathcal{K}_0, g_1] \mathcal{K}_0 \partial_x U_1 - f_2 g_2 (1 + \mathcal{K}_0^2) \mathcal{K}_0 \partial_x U_1 \\
G_3 &= (f_3 + H_1)^{-1} G_2 + \sum_{n=1}^{\infty} (-f_3^{-1} H_1)^n (f_3^{-1} f_4 \mathcal{K}_0 \partial_x U_1) \\
H_1 &= -f_2[\mathcal{K}_0, f_1] + f_2 \mathcal{K}_0 [\mathcal{K}_0, f_2] - f_2 (1 + \mathcal{K}_0^2) (f_2') + (f_1 - f_2 \mathcal{K}_0) f_2 \mathcal{S}_1(X) \cdot \\
V_2 &= \mathcal{K}_0 V_1 + \partial_x (\mathcal{S}_1(X) U_1)
\end{align*}
\]
Remark 4.1 The semilinear perturbation $G_3$ can lead to a growth of the solutions, but it does not lead to a loss of regularity in contrast to the quasilinear term $(f_3 + H_1)^{-1}(f_5 K(X)^2 \partial_\alpha U_1)$. As an example consider $\partial_t^2 U = (1 + i \delta) \partial_\alpha^2 U + \partial_\alpha U$. If $\delta = 0$ a $C^0$-semigroup is generated.

The history of the local existence and uniqueness theory for (6) and (7) starts with a canal with infinite depth, i.e. when in Fourier space the operator $\hat{K}_0$ has to be replaced by the Heaviside function $\hat{H}$, where $\hat{H}(k) = 1$ for $k \geq 0$ and $\hat{H}(k) = -1$ for $k < 0$. For this problem results are available in [Na74], [Sh76], and [Wu97]. In case of finite depth results are available in [Yo82] without and in [Yo83] with surface tension. For our system we obtain in a similar way.

Theorem 4.2
Let $\mathcal{H}_s^c = H^s \times H^s \times H^{s-1/2} \times H^{s-1}$. For all $s \geq 4$ there exists a $C_1 > 0$ such that for all $C_2 \in (0, C_1]$ we have a $T_0 > 0$ such that the following is true. For each initial condition $\mathcal{V}_0 \in \mathcal{H}_s^c$ with $\|\mathcal{V}_0\|_{\mathcal{H}_s^c} \leq C_2$ there exists a unique solution $\mathcal{V} \in C([0, T_0], \mathcal{H}_s^c)$ of (33) with $\mathcal{V}|_{t=0} = \mathcal{V}_0$.

Remark 4.3 The existence proofs are based on the choice of a suitable energy

$$E_s(\mathcal{V}, \mathcal{V}) = (X_1, X_1)_{H^s} + (X_2, X_2)_{H^s} + (U_1, U_1)_{H^{s-1}} + (V_1, V_1)_{H^{s-1}}$$

$$+ \int_{\mathbb{R}} (\lambda \partial_x^{s-1} U_1)^2 \, dx \, dx$$

which is equivalent to the usual $\mathcal{H}_s^c$-scalar product. Here $(\cdot, \cdot)_{H^s}$ stands for the usual $H^s$-scalar product and $\lambda^2 = -\partial_\alpha K_0$. The energy $E_s(\cdot, \cdot)$ is chosen in such a way that in the computation of $\frac{d}{dt} E_s(\mathcal{V}, \mathcal{V})$ all quasilinear terms from (33) cancel and we can estimate $\frac{d}{dt} E_s(\mathcal{V}, \mathcal{V})$ can in terms of $E_s(\mathcal{V}, \mathcal{V})$. Then Gronwall’s inequality can be applied and a priori estimates on the solution can be obtained. We provide additional details in the proof of Theorem 4.7, which is the form of the existence theorem that we actually use.

Theorem 4.2 is not in itself sufficient for our purposes, since the $L^2$-norm of $X_1$ does not stay $O(1)$-bounded on a time scale $O(1/\varepsilon^3)$ as we have seen in Section 2. It will be refined below.

Remark 4.4 Previous local existence and uniqueness theorems differ in the way they construct a suitable quasilinear system which is constructed out of (6) and (7). In [Na74], [Yo82], [Cr85] an eight-dimensional system for $(X_1, X_2, Z_1, Z_2, U_1, U_2, Y_1, V_1)$ with $Z_j = \partial_\alpha X_j, Y_1 = \partial_\alpha^2 X_1$ and $V_1 = \partial_\alpha^3 X_1$ is considered. The main idea in showing local existence and uniqueness of solutions is to choose $X_j \in H^s$ and $Z_j \in H^{s-1/2}$.
ignoring the relation $Z_j = \partial_a X_j$. In [Wu97] a two-dimensional system for our $U_1$ and $V_1$ alone is obtained by using the representation $\mathcal{K}(X) = \mathcal{Q}(X) \circ \mathcal{K}_0 \circ \mathcal{Q}(X)^{-1}$. We note that recently Wu has also proved the local existence and uniqueness of solutions of the water wave problem in three space dimensions [Wu99].

Not all initial conditions $\mathcal{V}_0$ of (33) lead to solutions of the water wave problem (6) and (7), only those which have been computed from $X_1|_{t=0}$, $X_2|_{t=0}$, and $U_1|_{t=0}$. Therefore, we introduce the space $\mathcal{C}_{p,X}$ of functions which satisfy the compatibility conditions.

**Definition 4.5**

We define

$$\mathcal{C}_{p,X} = \{ \mathcal{V} = (\phi_0, \phi_1, \phi_2, \phi_3) |$$

$$\begin{align*}
a) & \quad (1 + \partial_a \phi_0) \phi_3 + (\partial_a \phi_1)(1 + \phi_4) = 0 \\
b) & \quad \phi_4 = \mathcal{K}(\phi_0, \phi_1) \phi_3 + [\partial_t, \mathcal{K}(\phi_0, \phi_1)] \phi_2 \}
\end{align*}$$

**Remark 4.6** For finite depth we do not expect that such a system exists in the variables $(U_1, V_1)$ alone independent of $X_2$. Moreover, as explained in Section 2 a $(U_1, V_1)$-system leads to some secular growth for $U_1$ if the compatibility conditions are not satisfied. As explained above it is difficult to make a direct use of these conditions in proving the error estimates.

We have that $[\partial_t, \mathcal{K}(\phi_0, \phi_1)]$ is a function of $\phi_0$ and $\phi_1$ since $\partial_t \phi_0 = \phi_2$ and $\partial_t \phi_1 = \mathcal{K}(\phi_0, \phi_1) \phi_2$. From $\mathcal{V}|_{t=0} \in \mathcal{C}_{p,X}$ it follows $\mathcal{V}(t) \in \mathcal{C}_{p,X}$ for all $t > 0$ due to the construction of (33).

As already explained in Sections 1 and 2 for our approximation problem we must avoid trying to estimate the variable $X_1$ in some Sobolev space. Due to this fact we replaced $X_1$ by $Z_1 = \mathcal{K}_0 X_1$. This is possible due to the considerations of Section 3. Thus, we finally consider

$$\begin{align*}
\partial_t Z_1 &= \mathcal{K}_0 U_1 \\
\partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{S}_1(\mathcal{V}) U_1 \\
\partial_t U_1 &= V_1 \\
\partial_t V_1 &= -f_6 \mathcal{K}_0 \partial_a U_1 - G_3
\end{align*}$$

for $\mathcal{V}_e = (Z_1, X_2, U_1, V_1)$. Again we have the local existence and uniqueness of solutions.

**Theorem 4.7**

Let $\mathcal{H}_e = H^s \times H^s \times H^{s-1/2} \times H^{s-1}$. For all $s \geq 4$ there exists a $C_1 > 0$ such that for all
$C_2 \in (0, C_1]$ we have a $T_0 > 0$ such that the following is true. For each initial condition $\mathcal{W}_{e,0} \in \mathcal{H}_e$ with $\|\mathcal{W}_{e,0}\|_{\mathcal{H}_e} \leq C_2$ there exists a unique solution $\mathcal{W}_e \in C([0,T_0], \mathcal{H}_e)$ of (34) with $\mathcal{W}_{e|t=0} = \mathcal{W}_{e,0}$.

**Proof.** As in the proof of Theorem 4.2, the proof is based on estimating the time derivative of the energy function

$$E_s(\mathcal{W}, \mathcal{W}) = (Z_1, Z_1)_{H^s} + (X_2, X_2)_{H^s} + (U_1, U_1)_{H^{s-1}} + (V_1, V_1)_{H^{s-1}}$$

$$+ \int_{\mathbb{R}} (\lambda \partial_x^{s-1} U_1)^2 f_6 d\alpha$$

Lemmas 3.12 and 3.15 imply that there exists $C_1, C_2 > 0$ such that if $\|\mathcal{W}\|_{\mathcal{H}_e} \leq C_1$, then

$$\max(\|f_1 - 1\|_{H^{s-1}}, \|f_2\|_{H^{s-1}}, \|f_3 - 1\|_{H^{s-1}}, \|f_4 - 1\|_{H^{s-1}}, \|f_6 - 1\|_{H^{s-1}}, \|g_1 - 1\|_{H^{s-1}}, \|g_2\|_{H^{s-1}}) \leq C_2 \|\mathcal{W}\|_{\mathcal{H}_e}.$$  

In particular, if we make $C_1 > 0$ small enough, we can be sure that $\|f_6 - 1\|_{H^{s-1}} < \frac{1}{4}$. The two lemmas also imply that (again, possibly by choosing $C_1$ somewhat smaller) $\|G_3\|_{H^{s-1}} \leq C_2 \|\mathcal{W}\|_{\mathcal{H}_e}$.

With these estimates, one sees immediately that for any $r \geq 1$, if we differentiate along solutions of (34), there exists a constant $C_E$ such that

$$\frac{d}{dt} E_r(\mathcal{W}, \mathcal{W}) \leq C_E ((Z_1, Z_1)_{H^r} + (X_2, X_2)_{H^r} + (U_1, U_1)_{H^r} + (V_1, V_1)_{H^r}) .$$  

(36)

We use (36) to control derivatives of the solutions of (34) with $r \leq s - 1$. It is not suitable for controlling the highest derivatives because it bounds $\|U_1\|_{H^{r-\frac{1}{2}}}$ and $\|V_1\|_{H^{r-1}}$ in terms of $\|U_1\|_{H^r}$ and $\|V_1\|_{H^r}$ respectively. This problem arises because of the derivatives of the terms $(Z_1, Z_1)_{H^r}$ and $(X_2, X_2)_{H^r}$ in the energy – the remaining terms were chosen precisely so that the quasi-linear terms would cancel. To bound the highest derivative, we note that because of the form of the right hand side of the equations in (34), we can control the norms of $\|\partial^s Z_1\|_{L^2}$ and $\|\partial^s U_1\|_{L^2}$ in terms of $\|V_1\|_{H^{r-1}}$. More precisely, consider

$$\frac{1}{2} \partial_\alpha \int [\partial^s_z Z_1 + \frac{1}{f_6}(\partial^s_r V_1)]^2 d\alpha = - \int [(\partial^s_z Z_1 + \frac{1}{f_6}(\partial^s_r V_1))$$

$$\times(\partial^s_\alpha K_0 U_1 - \frac{1}{f_6} \partial^s_r f_6 K_0 U_1 \frac{f_6}{f^2_6} \partial^s_r V_1 - \frac{1}{f_6} \partial^s_r G_3)] d\alpha$$

$$- \int [(\partial^s_z Z_1 + \frac{1}{f_6}(\partial^s_r V_1))$$

$$\times(\frac{1}{f_6}[f_6, \partial^s_r G_3 K_0 U_1 + \frac{f_6}{f^2_6} \partial^s_r V_1 - \frac{1}{f_6} \partial^s_r G_3)] d\alpha$$

40
Applying the Cauchy-Schwartz inequality we see that there exists a constant $C_4 > 0$ such that the right hand side of this expression is bounded in magnitude by

$$C_4 (\| (\partial^s_\alpha Z_1) + \frac{1}{f_0}(\partial^s_\alpha V_1) \|^2_{L^2} + \| W \|_{H^s}) .$$

In this estimate we used the fact that if $\| W \|_{H^s} < C_1$, $\| f_0 \|_{L^\infty} \geq \frac{1}{2}$, and $\| f_0 \|_{L^\infty} \leq \| f_0 \|_{H^1} \leq C_3$, for some constant $C_3$. In addition, terms in the commutator $[f_0, \partial^s_\alpha]K_0 U_1$ of the form $(\partial^r_\alpha f_0)(\partial^{s-1-r}_\alpha K_0 U_1)$ with $r < s - 1$ are estimated by bounding the $L^\infty$ of $\partial^r_\alpha f_0$ by a constant with the aid of the Sobolev embedding theorem and the estimate above on $\| f_0 - 1 \|_{H^{s-1}}$ and bounding the $L^2$ norm of $\partial^{s-1-r}_\alpha K_0 U_1$ by $C \| W \|_{H^s}$. The term in the commutator of the form $(\partial^s_\alpha f_0)(K_0 U_1)$ is controlled by bounding the $L^\infty$ norm of $K_0 U_1$ by the Sobolev theorem and then controlling the $L^2$ norm of $\partial^s_\alpha f_0$ by the estimate on $\| f_0 - 1 \|_{H^{s-1}}$. In like fashion we find

$$\frac{1}{2} \partial_t \int [(\partial^s_\alpha X_2) + \frac{1}{f_0}(\partial^s_\alpha V_1)]^2 d\alpha \leq C_4 (\| (\partial^s_\alpha X_2) + \frac{1}{f_0}(\partial^s_\alpha V_1) \|^2_{L^2} + \| S_1(W)U_1 \|^2_{H^s} + \| G_3 \|^2_{H^{s-1}} + \| V_1 \|^2_{H^{s-1}}) . (38)$$

Define $Z_1^s = (\partial^s_\alpha Z_1 + \frac{1}{f_0}(\partial^s_\alpha V_1))$, and $X_2^s = (\partial^s_\alpha X_2 + \frac{1}{f_0}(\partial^s_\alpha V_1))$, and

$$E_s(W, W) = (Z_1^s, Z_1^s)_{L^2} + (X_2^s, X_2^s)_{L^2} + (U_1, U_1)_{H^{s-1}} + (V_1, V_1)_{H^{s-1}} + \int_{\mathbb{R}} (\lambda \partial^s_\alpha U_1)^2 f_0 d\alpha$$

Then by (36), (37), and (38), there exists a constant $C_5$, such that

$$\frac{d}{dt} (E_s(W, W) + E_{s-1}(W, W)) \leq C_5 (E_s(W, W) + E_{s-1}(W, W)) . (39)$$

However, if $\| W \|_{H^s} \leq C_1$, the norm defined by $E_s(W, W) + E_{s-1}(W, W)$ is equivalent to the $H^s$ norm.

With the help of this a priori estimate the existence theory locally in time is completed by applying the usual fixed point argument to the iteration scheme [Kat75]

$$\begin{align*}
\partial_t Z_{1, j} &= K_0 U_{1, j} \\
\partial_t X_{2, j} &= K_0 U_{1, j} + S_1(W_{j-1}) U_{1, j-1} \\
\partial_t U_{1, j} &= V_{1, j} \\
\partial_t V_{1, j} &= -f_{6, j-1} K_0 \partial_\alpha U_{1, j} - G_{3, j-1}
\end{align*} (40)$$

with $W_{c, j}|_{t=0} = W_{c, 0}$ for all $j \in \mathbb{N}$ in the space $C([0, T_0], H^s_c)$ for a $T_0 > 0$ sufficiently small, and hence, Theorem 4.7 follows immediately.

**Definition 4.8**

The compatibility conditions for (34) are denoted by $\mathcal{C}_p$. They are the obvious analogues of $C_{p, X}$.
5 Long wave analysis

In this section we derive the KdV–equations. We construct from the solutions of the KdV–equations an approximation for the solutions of (8) and we estimate the formal error, the so called residual.

5.1 Derivation of the KdV–equations

We are interested in the long wave limit, i.e. in the case when the initial conditions are functions of the large spatial variable $\alpha = \varepsilon \alpha$. As we have seen for the linearized system there will be no fast dynamics and so we also introduce the variable for the slow temporal scale $\xi = \varepsilon t$. The amplitude is scaled in such a way that the leading linear and leading nonlinear terms are of the same order. As explained in Section 2 we refine this approximation to make the residual sufficiently small. Therefore we make the ansatz

$$X_1(\varepsilon, \alpha, t) = \varepsilon X_1(\alpha, \xi)$$

$$X_2(\varepsilon, \alpha, t) = \varepsilon^2 X_2(\alpha, \xi)$$

Expansion of (6) and (7) to quadratic order (see Lemma 3.8) gives

$$0 = (1 + \partial_\alpha X_1)\partial_\xi^2 X_1 + \partial_\alpha X_2(1 + \partial_\xi^2 X_2)$$

$$\partial_\xi X_2 = \mathcal{K}_0 \partial_\xi X_1 + [X_1, \mathcal{K}_0] \partial_\alpha \partial_\xi X_1 - (X_2 + \mathcal{K}_0(X_2 \mathcal{K}_0)) \partial_\alpha \partial_\xi X_1$$

$$\quad + \text{h.o.t.}$$

(41)

In the long wave limit we obtain

$$\mathcal{K}_0(\partial_\alpha) = -\varepsilon \partial_\alpha - \frac{1}{3} \varepsilon^3 \partial_\alpha^3 - \mathcal{O}(\varepsilon^5)$$

and consequently

$$\partial_\xi^2 X_1 + \partial_\alpha^2 X_2 = -\varepsilon^2 ((\partial_\alpha^2 X_1)(\partial_\xi^2 X_1)) - \varepsilon^4 (\partial_\alpha^2 X_2)(\partial_\xi^2 X_2)$$

$$\partial_\alpha X_2 + \partial_\alpha \partial_\xi X_1 = \varepsilon^2 (-\frac{1}{3} \partial_\alpha^3 X_1 + (\partial_\alpha^2 X_1)(\partial_\alpha \partial_\xi X_1) - X_2 \partial_\alpha \partial_\xi X_1) + \mathcal{O}(\varepsilon^4).$$

The first order long wave approximation focuses on those terms of $\mathcal{O}(\varepsilon^0)$ and $\mathcal{O}(\varepsilon^2)$, i.e. on the system of equations

$$\partial_\xi^2 X_1 + \partial_\alpha^2 X_2 = -\varepsilon^2 ((\partial_\alpha^2 X_1)(\partial_\xi^2 X_1))$$

$$\partial_\alpha X_2 + \partial_\alpha \partial_\xi X_1 = \varepsilon^2 (-\frac{1}{3} \partial_\alpha^3 X_1 + (\partial_\alpha^2 X_1)(\partial_\alpha \partial_\xi X_1) - X_2 \partial_\alpha \partial_\xi X_1).$$

(42)

(43)
We solve these equations (up to terms of $O(\varepsilon^4)$) by choosing

$$X_2 + \partial_\alpha X_1 = \varepsilon^2\left(-\frac{1}{3}\partial_\alpha^3 X_1 + \left(\partial_\alpha^2 X_1\right)^2\right). \tag{44}$$

Differentiating this with respect to time gives

$$\partial_t X_2 + \partial_\alpha \partial_\alpha X_1 = -\frac{\varepsilon^2}{3}\partial_\alpha^3 \partial_\alpha X_1 + 2\varepsilon^2(\partial_\alpha \partial_\alpha X_1)(\partial_\alpha X_1) \tag{45}$$

$$= -\frac{\varepsilon^2}{3}\partial_\alpha^3 \partial_\alpha X_1 + \varepsilon^2(\partial_\alpha^2 \partial_\alpha X_1)(\partial_\alpha X_1) - \varepsilon^2 X_2(\partial_\alpha \partial_\alpha X_1) + \varepsilon^4 \epsilon_1^4,$$

where

$$\epsilon_1^4 = \left(-\frac{1}{3}\partial_\alpha^3 X_1 + (\partial_\alpha X_1)^2\right)(\partial_\alpha \partial_\alpha X_1) .$$

This means that by the choice (44) equation (43) is satisfied up to terms of $O(\varepsilon^4)$.

Inserting (44) into (42) gives

$$\partial_t^2 X_1 - \partial_\alpha^2 X_1 = \varepsilon^2\left(\frac{1}{3}\partial_\alpha^4 X_1 - \partial_\alpha \partial_\alpha X_1 \partial_\alpha^2 X_1 - \partial_\alpha^2((\partial_\alpha X_1)^2))\right) \tag{46}$$

and so

$$\partial_t^2 X_1 = (1 + \varepsilon^2 \partial_\alpha X_1)^{-1}(\partial_\alpha^2 X_1 + \varepsilon^2\frac{1}{3}\partial_\alpha^4 X_1 - \varepsilon^2 \partial_\alpha((\partial_\alpha X_1)^2)) \tag{47}$$

$$= \partial_\alpha^2 X_1 + \varepsilon^2\frac{1}{3}\partial_\alpha^4 X_1 - \varepsilon^2 \partial_\alpha((\partial_\alpha X_1)^2) - \varepsilon^2 \partial_\alpha \partial_\alpha X_1 \partial_\alpha^2 X_1 + \varepsilon^4 \epsilon_2^4$$

$$= \partial_\alpha^2 X_1 + \varepsilon^2\frac{1}{3}\partial_\alpha^4 X_1 - \varepsilon^2\frac{3}{2} \partial_\alpha((\partial_\alpha X_1)^2) + \varepsilon^4 \epsilon_2^4,$$

where

$$\epsilon_2^4 = \varepsilon^4\left(\frac{1}{3}\partial_\alpha^4 X_1 - \varepsilon^2 \partial_\alpha((\partial_\alpha X_1)^2)\right) \sum_{n=1}^{\infty} (-\varepsilon^2 \partial_\alpha X_1)^n + (\partial_\alpha^2 X_1) \sum_{n=2}^{\infty} (-\varepsilon^2 \partial_\alpha X_1)^n$$

Corresponding to our variable $Z_1 = \mathcal{K}_0 X_1$ we introduce in the long wave limit $Z_1 = -\partial_\alpha X_1$. From (46), we see that $Z_1 = -\partial_\alpha X_1$ satisfies

$$\partial_t^2 Z_1 - \partial_\alpha^2 Z_1 = \varepsilon^2\left(\frac{1}{3}\partial_\alpha^4 Z_1 + \frac{3}{2} \partial_\alpha^2((Z_1)^2)\right) + O(\varepsilon^4). \tag{47}$$

**Remark 5.1** Neglecting terms of order $O(\varepsilon^4)$ system (47) is ill-posed. Using $\partial_t^2 X_1 - \partial_\alpha^2 X_1 = O(\varepsilon^2)$ gives the well-posed system

$$\partial_t^2 Z_1 - \partial_\alpha^2 Z_1 = \varepsilon^2\left(\frac{1}{3}\partial_\alpha^2 \partial_\alpha^2 Z_1 + \frac{3}{2} \partial_\alpha^2((Z_1)^2)\right). \tag{48}$$

Equations (46), (47), and (48) have first been derived by Boussinesq (cf. [Bo1877]).

43
Our goal is to obtain approximation equations which are independent of $\varepsilon$. Therefore, we make the ansatz
\[ Z_1 = A_1(\alpha - t, \varepsilon^2 t) + A_2(\alpha + t, \varepsilon^2 t) \] (49)
for the solutions of (47). This yields with $T = \varepsilon^2 t$ to
\[
\varepsilon^2(-2\partial_\alpha \partial_t (A_1 - A_2)) + \varepsilon^4 \partial^2_t (A_1 + A_2)
\]
\[ = \varepsilon^2\left(\frac{1}{3} \partial^4_\alpha (A_1 + A_2) + \frac{3}{2} \partial^2_\alpha ((A_1)^2 + (A_2)^2) + r_1\right) + \mathcal{O}(\varepsilon^4)
\]
with $r_1 = 3\partial^2_\alpha (A_1 A_2)$. Since $A_1$ and $A_2$ are spatially localized and since the two wave packets only met on a time scale $\mathcal{O}(1/\varepsilon)$ which is relatively short compared to the over all time scale $\mathcal{O}(1/\varepsilon^3)$ we claim that the influence of the term $r_1$ on the dynamics of $A_{1,2}$ is of order $\mathcal{O}(\varepsilon^2)$ and as a consequence we claim that the water wave problem can completely be described by two decoupled KdV-equations
\[
2\partial_T A_1 = -\frac{1}{3} \partial^3_\alpha A_1 - \frac{3}{2} \partial_\alpha ((A_1)^2), \quad 2\partial_T A_2 = \frac{1}{3} \partial^3_\alpha A_2 + \frac{3}{2} \partial_\alpha ((A_2)^2).
\] (50)
This has already been shown in [Sch98] for (48) as a model problem.

**Remark 5.2** The general, first order long wave limit is represented by the equations (42) and (43). The freedom in the parametrization of the top surface will allow us to choose (44) and (50) as approximation equations.

### 5.2 The approximation and estimates for the residual

It is the aim of this section to construct an approximation $\varepsilon^2 \Psi_e$ out of (44) and (50) for the extended system (34). This has to be done in such a way that the extended residual
\[ \text{Res}_e(\mathcal{W}_e) = (\text{Res}_Z(\mathcal{W}_e), \text{Res}_X(\mathcal{W}_e), \text{Res}_U(\mathcal{W}_e), \text{Res}_V(\mathcal{W}_e)) \]
with
\[
\text{Res}_Z(\mathcal{W}_e) = -\partial_t Z_1 + \mathcal{K}_0 U_1
\]
\[
\text{Res}_X(\mathcal{W}_e) = -\partial_t X_2 + \mathcal{K}_0 U_1 + S_1(\mathcal{W}) U_1
\]
\[
\text{Res}_U(\mathcal{W}_e) = -\partial_t U_1 - V_1
\]
\[
\text{Res}_V(\mathcal{W}_e) = -\partial_t V_1 - f_3 \mathcal{K}_0 \partial_\alpha U_1 - G_3.
\]
is small. The residual contains all terms which do not drop out after inserting the extended approximation $\varepsilon^2 \Psi_e$ into system (34). We additionally define
\[ \text{Res}(\mathcal{W}) = (\text{Res}_Z(\mathcal{W}), \text{Res}_X(\mathcal{W}), \text{Res}_U(\mathcal{W})) , \]
where in computing \( \text{Res}_{U_1}(W) \), we use the fact that \( \partial_t U_1 \) can also be written as \( \partial_t U_1 = (1 + M_2 Z_1 + (\partial_\alpha X_2) K(W))^{-1}[(\partial_\alpha X_2)(1 + [\partial_t, K(W)] U_1)] \) and hence, \( \text{Res}_{U_1} = -\partial_t U_1 + (1 + M_2 Z_1 + (\partial_\alpha X_2) K(W))^{-1}[(\partial_\alpha X_2)(1 + [\partial_t, K(W)] U_1)] \) With these preparations we define the approximations

\[
\varepsilon^2 \Psi = (\varepsilon^2 \psi_{Z_1}, \varepsilon^2 \psi_{X_2}, \varepsilon^2 \psi_{U_1}) \quad \text{and} \quad \varepsilon^2 \Psi_\varepsilon = (\varepsilon^2 \psi_{Z_1}, \varepsilon^2 \psi_{X_2}, \varepsilon^2 \psi_{U_1}, \varepsilon^3 \psi_{V_1}),
\]

(51)

where

\[
\begin{align*}
\varepsilon^2 \psi_{Z_1}(\alpha, t) &= \varepsilon^2(A_1(\alpha - t, \varepsilon^2 t) + A_2(\alpha + t, \varepsilon^2 t)) \\
\varepsilon^2 \psi_{X_2}(\alpha, t) &= +\varepsilon^2 \psi_{Z_1} - \varepsilon^4 (\psi_{Z_1})^2 \\
\varepsilon^2 \psi_{U_1}(\alpha, t) &= \varepsilon^2 \partial_\alpha K_0^{-1} \psi_{Z_1}(\alpha, t), \\
\varepsilon^3 \psi_{V_1}(\alpha, t) &= \varepsilon^3 \partial_\alpha^3 K_0^{-1} \psi_{Z_1}(\alpha, t),
\end{align*}
\]

(52)

Note that the \( V_1 \)-component is scaled with \( \varepsilon^3 \).

**Remark 5.3** The motivation for the definitions of \( \psi_{U_1} \) and \( \psi_{V_1} \) are that \( \varepsilon^2 \psi_{U_1} = \varepsilon^2 \partial_\alpha \psi_{X_1} \) where \( \psi_{X_1} = K_0^{-1} \psi_{Z_1} \). However \( K_0^{-1} \psi_{Z_1} \) is undefined unless \( \int \psi_{Z_1}(\alpha, t) d\alpha = 0 \). However, since \( A_1 \) and \( A_2 \) satisfy the KdV equations, \( \int (\partial_t \psi_{Z_1}(\alpha, t)) d\alpha = 0 \) and hence \( \psi_{U_1} \) is well defined. A similar argument applies to \( \psi_{V_1} \).

The approximation \( \varepsilon^2 \Psi_\varepsilon \) has the following properties.

**Lemma 5.4**

Fix \( s \geq 1 \). For all \( C_A > 0 \) there exist \( C_\Psi, \varepsilon_0 > 0 \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) the following is true. Let \( A_1, A_2 \in C([0, T_0], H^{s+6}) \) be solutions of (50) with

\[
\sup_{T \in [0, T_0]} \|(A_1, A_2)(T)\|_{H^{s+6}} \leq C_A.
\]

Then we have

\[
\sup_{t \in [0, T_0 / \varepsilon^3]} \|\Psi_\varepsilon(t)\|_{H^s} \leq C_\Psi,
\]

where \( H^s = H^s \times H^s \times H^{s-1/2} \times H^{s-1} \).

**Proof.** Note that \( \partial_t \psi_{Z_1} = \varepsilon \partial_\alpha \psi_{Z_1} \) and so

\[
\begin{align*}
\partial_\alpha \psi_{Z_1}(\alpha, t) &= -\partial_\alpha A_1(\alpha - t, \varepsilon^2 t) + \partial_\alpha A_2(\alpha + t, \varepsilon^2 t) + \varepsilon^2 \partial_\alpha A_1(\alpha - t, \varepsilon^2 t) \\
&\quad + \varepsilon^2 \partial_\alpha A_2(\alpha + t, \varepsilon^2 t) \\
&= -\partial_\alpha A_1(\alpha - t, \varepsilon^2 t) + \partial_\alpha A_2(\alpha + t, \varepsilon^2 t) - \varepsilon^2 \left( \frac{1}{6} \partial_\alpha^3 A_1 + \frac{3}{4} \partial_\alpha (A_1^2) \right) \\
&\quad + \varepsilon^2 \left( \frac{1}{6} \partial_\alpha^3 A_2 + \frac{3}{4} \partial_\alpha (A_2^2) \right)
\end{align*}
\]

(53)
Taking Fourier transforms we see that
\[
(K_0^{-1} \partial_k \psi_{Z_1})(k, t) = \frac{1}{-i \tanh(k)} \left\{ -iKe^{itK} \hat{A}_1(K, \varepsilon^2 \frac{t}{2}) \right. \\
+ iKe^{-itK} \hat{A}_2(K, \varepsilon^2 t) + \frac{\varepsilon^2}{6} K^3 [e^{itK} \hat{A}_1 - e^{-itK} \hat{A}_2] \]
\[
- iK^2 \frac{3\varepsilon^2}{4} [e^{itK} (A_1)^2 - e^{-itK} (A_2)^2] \right\},
\]
where \( k = \varepsilon K \). Hence,
\[
\varepsilon^2 \|(K_0^{-1} \partial_k \psi_{Z_1})\|_{H^{s+\frac{5}{2}}}^2 = \varepsilon^2 \int (1 + k^2)^{s-\frac{1}{2}} |(K_0^{-1} \partial_k \psi_{Z_1})(k, t)|^2 dk \\
\leq C(\|A_1\|_{H^{s+\frac{5}{2}}}^2 + \|A_2\|_{H^{s+\frac{5}{2}}}^2 + \|A_1\|_{H^{s+\frac{5}{2}}}^4 + \|A_2\|_{H^{s+\frac{5}{2}}}^4) \tag{55}
\]
if \( \|A_{1, 2}\|_{H^{s+\frac{5}{2}}} \leq C_1 \), for all \( 0 \leq t \leq T_0/\varepsilon^2 \).

The following lemma ensures that the approximation \( \varepsilon^2 \Psi_\varepsilon \) defined in (51) is at least formally a good approximation. Let us define \( (\tau_t u)(\alpha) = u(\alpha + t) \) and \( (S_\varepsilon A)(\alpha) = A(\varepsilon \alpha) \).

**Lemma 5.5**

Fix \( s \geq 1 \). For all \( C_A > 0 \) there exist \( C_{Res}, \varepsilon_0 > 0 \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) the following is true. Let \( A_1, A_2 \in C([0, T_0], H^{s+6}(2)) \) be solutions of (50) with
\[
\sup_{T \in [0, T_0]} \|(A_1, A_2)(T)\|_{H^{s+6}(2)} \leq C_A
\]
or let \( A_1, A_2 \in C([0, T_0], H^{s+6}) \) be solutions of (50) with \( A_2 = 0 \) and
\[
\sup_{T \in [0, T_0]} \|(A_1, A_2)(T)\|_{H^{s+6}} \leq C_A.
\]

Then for
\[
Q_1 = \text{Res}_\varepsilon(\varepsilon^2 \Psi_\varepsilon(s)) \quad \text{and} \quad Q_2 = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \varepsilon^5 \partial_\alpha ((\tau_{-t} S_\varepsilon A_1)(\tau_{-t} S_\varepsilon A_2)) \\ -\frac{3}{2} \varepsilon^6 \partial_\alpha ((\tau_{-t} S_\varepsilon A_1)(\tau_{-t} S_\varepsilon A_2)) \end{pmatrix}
\]
we have the estimates
\[
\sup_{T \in [0, T_0/\varepsilon^3]} \|Q_1 - Q_2\|_{H^s} \leq C_{Res} \varepsilon^{13/2}
\]
and
\[
\varepsilon^{7/2} q(t) := \|Q_2(t)\|_{H^s} \leq C_{Res} \varepsilon^{9/2}/(1 + (\varepsilon t)^2).
\]
Proof. Recalling $Z_1 = \mathcal{K}_0 X_1$ and $Z_1 = -\partial_{\alpha} X_1$ we obtain

\[
\text{Res}_{X_2}(\mathcal{W}_e) = -\varepsilon^3 \partial_{\alpha} X_2 - \varepsilon^3 \partial_{\alpha} U_1 + \varepsilon^5 \left( -\frac{1}{3} \partial_{\alpha}^3 U_1 - Z_1 \partial_{\alpha} U_1 - X_2 \partial_{\alpha} U_1 \right) + O(\varepsilon^7)
\]

\[
\text{Res}_{U_1}(\mathcal{W}_e) = -\varepsilon^3 \partial_{\alpha} U_1 - \varepsilon^3 \partial_{\alpha} X_2 + \varepsilon^5 (Z_1 \partial_{\alpha} U_1) + O(\varepsilon^7)
\]

\[
\text{Res}_{V_1}(\mathcal{W}_e) = -\varepsilon^4 \partial_{\alpha} V_1 - \varepsilon^4 \partial_{\alpha} \partial_{\alpha} X_2 - \varepsilon^6 (\partial_{\alpha} Z_1)(\partial_{\alpha} X_2) - \varepsilon^6 Z_1 (\partial_{\alpha} \partial_{\alpha} X_2) + O(\varepsilon^8)
\]

\[
= -\varepsilon^4 \partial_{\alpha} V_1 + \varepsilon^4 \partial_{\alpha}^2 U_1 - \varepsilon^6 \partial_{\alpha} \left( -\frac{1}{3} \partial_{\alpha}^3 U_1 - Z_1 \partial_{\alpha} U_1 - X_2 \partial_{\alpha} U_1 \right)
\]

\[
+ \varepsilon^6 (\partial_{\alpha}^2 U_1)(\partial_{\alpha} X_2) + \varepsilon^6 Z_1 (\partial_{\alpha}^2 U_1) + O(\varepsilon^8)
\]

\[
= -\varepsilon^4 \partial_{\alpha} V_1 + \varepsilon^4 \partial_{\alpha}^2 U_1 + \varepsilon^6 \frac{1}{3} \partial_{\alpha}^3 U_1 + 2 \varepsilon^6 Z_1 \partial_{\alpha}^2 U_1 + 2 \varepsilon^6 (\partial_{\alpha} X_2)(\partial_{\alpha} U_1)
\]

\[
+ \varepsilon^6 (\partial_{\alpha} Z_1)(\partial_{\alpha} U_1) + \varepsilon^6 X_2 (\partial_{\alpha}^2 U_1) + O(\varepsilon^8).
\]

Inserting $\mathcal{W}_e = \varepsilon^2 \Psi_e$ gives by construction

\[
\text{Res}_{X_1}(\varepsilon^2 \Psi_e) = 0
\]

\[
\text{Res}_{X_2}(\varepsilon^2 \Psi_e) = \varepsilon^5 \partial_{\alpha} \left( (\tau_{-\mathcal{L}} S_\varepsilon A_1)(\tau_{-\mathcal{L}} S_\varepsilon A_2) \right) + O(\varepsilon^7)
\]

\[
\text{Res}_{U_1}(\varepsilon^2 \Psi_e) = -\frac{1}{2} \varepsilon^5 \partial_{\alpha} \left( (\tau_{-\mathcal{L}} S_\varepsilon A_1)(\tau_{-\mathcal{L}} S_\varepsilon A_2) \right) + O(\varepsilon^7)
\]

\[
\text{Res}_{V_1}(\varepsilon^2 \Psi_e) = -\frac{3}{2} \varepsilon^6 \partial_{\alpha} \left( (\tau_{-\mathcal{L}} S_\varepsilon A_1)(\tau_{-\mathcal{L}} S_\varepsilon A_2) \right) + O(\varepsilon^8)
\]

Using Lemma 2.3 establishes

\[
\sup_{t \in [0,T_0/\varepsilon^3]} \|Q_1 - Q_2\|_{H^1_e} \leq C_{\text{Res}} \varepsilon^{13/2}.
\]

From $A_j \in C([0,T_0], H^{s+6}(2))$ and Sobolev’s embedding theorem which we apply in the form

\[
\|S_\varepsilon(uv)\|_{H^s} \leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{\mathcal{C}_0} \|S_\varepsilon(u \rho_1)\|_{H^s} \|S_\varepsilon(v \rho_2)\|_{\mathcal{C}_0}
\]

\[
\leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{\mathcal{C}_0} \varepsilon^{-1/2} \|u \rho_1\|_{H^s} \|v \rho_2\|_{\mathcal{C}_0}
\]

\[
\leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{\mathcal{C}_0} \varepsilon^{-1/2} \|u \rho_1\|_{H^{s+1}} \|v \rho_2\|_{H^{s+1}}
\]

we have

\[
\|((\tau_{-\mathcal{L}} S_\varepsilon A_1)(\tau_{-\mathcal{L}} S_\varepsilon A_2))\|_{H^s} \leq C \sup_{\alpha} \left| ((1 + (\varepsilon(\alpha - t))^2)(1 + (\varepsilon(\alpha + t))^2))^{-1} \right|
\]

\[
\times \varepsilon^{-1/2} \left( \sup_{T \in [0,T_0]} \|A_1\|_{H^{s+2}(2)} \left( \sup_{T \in [0,T_0]} \|A_2\|_{H^{s+1}(2)} \right) \right)
\]

\[
\leq C_{\text{Res}} \varepsilon^{-1/2} / (1 + (\varepsilon t)^2)
\]

and so with $\partial_T A_j \in C([0,T_0], H^{s+3}(2))$ and $\partial_{\alpha} \partial_T A_j \in C([0,T_0], H^{s+2}(2))$ the estimate about $Q_2$ follows.
6 The error estimates

Now we are ready to formulate our main result. As already explained for (34) it is difficult to enforce the condition \( \mathcal{W}_c(t) \in C_p \) which is necessary to avoid the secular growth of the solutions. However, since system (8) is a subsystem of (34) we also have indirectly the local existence and uniqueness for the solutions of (8) which is written in the following as

\[
\begin{aligned}
\partial_t Z_1 &= K_0 U_1 \\
\partial_t X_2 &= K_0 U_1 + S_1(X) U_1 \\
\partial_t U_1 &= -(1 - \mathcal{M}_2 Z_1 + (\partial_a X_2) K_0 + (\partial_a X_2) S_1(X))^{-1} \left[ (\partial_a X_2)(1 + [\partial_1, S_1(X)] U_1) \right]
\end{aligned}
\]  

(56)

For system (56) we show that there exist solutions which behave in approximately the same way as predicted by the approximation \( \varepsilon^2 \Psi \) defined in (52) and constructed via the solutions of the two decoupled KdV–equations (50).

**Theorem 6.1**

Fix \( s \geq 4 \), let \( \beta = 7/2 \), and let \( \mathcal{H}^s = H^s \times H^s \times H^{s-1/2} \). For all \( C_A, C_0, T_0 > 0 \) there exist \( C_R, \varepsilon_0 > 0 \) such that for all \( \varepsilon \in (0, \varepsilon_0) \) the following is true. Let \( A = (A_1, A_2) \in C([0, T_0], (H^{s+6}(2))^2) \) be solutions of (50) with

\[
\sup_{T \in [0, T_0]} \| (A_1, A_2) \|_{(H^{s+6}(2))^2} \leq C_A
\]

or let \( A = (A_1, A_2) \in C([0, T_0], (H^{s+6})^2) \) be solutions of (50) with \( A_2 = 0 \) and

\[
\sup_{T \in [0, T_0]} \| (A_1, A_2) \|_{(H^{s+6})^2} \leq C_A
\]

and let \( \mathcal{W}|_{t=0} = \varepsilon^2 \Psi |_{t=0} + \varepsilon^\beta R |_{t=0} \in \mathcal{H}^s \) with \( \| R |_{t=0} \|_{\mathcal{H}^s} \leq C_0 \). Then we have a unique solution \( \mathcal{W} = \varepsilon^2 \Psi + \varepsilon^\beta R \in C([0, T_0/\varepsilon^3], \mathcal{H}^s) \) of (56) which satisfies

\[
\sup_{t \in [0, T_0/\varepsilon^3]} \| R(t) \|_{\mathcal{H}^s} \leq C_R.
\]

**Remark 6.2** Note that if Theorem 6.1 holds, it guarantees that Lemma 2.5 holds as well.

**Proof of Theorem 1.3:**

Due to Remark 6.2 we have the estimate

\[
\sup_{t \in [0, T_1/\varepsilon^3]} |X_1|_{t=0(t)} \leq C_2 \varepsilon^{1/2}.
\]

48
With Lemma 2.6 we split the initial conditions \((\phi_0, \phi_1, \phi_2)\) for (56) in the long wave limit into \(A_1|_{T=0}\) and \(A_2|_{T=0}\), the initial conditions of the two decoupled KdV-equations (50). The local existence and uniqueness for the solutions of the KdV-equation in \(H^{s+6}(2) \cap H^{s+10}(0)\) is known from [Kat83]. But since we want to describe the water wave problem for all \(t \in [T_0/\varepsilon^3]\) by the solutions of the KdV-equations for all fixed \(T_0 > 0\), we need the global existence of solutions in \(H^{s+6}(2) \cap H^{s+10}(0)\).

**Lemma 6.3**

Let \(s \geq -2\). Then for all \(C_1, T_0 > 0\) there exists a \(C_2 > 0\) such that the following is true. For the initial condition \(A_1|_{T=0} \in H^{s+6}(2) \cap H^{s+11}(0)\) there exists a unique solution \(A_1 \in C([0, \infty), H^{s+6}(2) \cap H^{s+10}(0))\) of the KdV-equation (50) which exists globally in time and satisfies \(\sup_{t \in [0, T_0]} \|A_1(T)\|_{H^s(0)} < \infty\).

**Proof.** From the global existence result in [Kat79] for solutions in \(H^2\) and from the conservation laws for the KdV-equation the global existence of solutions in each \(H^s(0)\) for \(s \geq 2\) is known ([KPV91]), i.e.

\[
\sup_{T \in [0, T_0]} \|A_1(T)\|_{H^s(0)} < \infty. \tag{57}
\]

Now fix \(T_0 > 0\). The global existence in \(H^{s+6}(2) \cap H^{s+10}(0)\) is proved with the help of the help of the variation of constant formula

\[
A_1(T) = G(T)A_1|_{T=0} + \frac{3}{4} \int_0^T G(T - s)\partial_\mu(A(s)^2)ds,
\]

where \(G(T)\) is defined by \((\mathcal{F}(G(T)A)(K) = e^{tK^3/6}A(K)\). Since

\[
\sup_{T \in [0, T_0]} \|G(T)\|_{H^{s+6}(2) \cap H^{s+10}(0)} \rightarrow H^{s+6}(2) \cap H^{s+10}(0)
\]

\[\leq \sup_{T \in [0, T_0]} \|K \rightarrow (1 + K^2)^{-2}e^{iK^3T/6}\|_{C^2_S} + \|K \rightarrow e^{iK^3T/6}\|_{C^2_S} < \infty
\]

and since

\[
\|A\partial_\mu A\|_{H^{s+6}(2) \cap H^{s+10}(0)} \leq C\|A\|_{H^{s+11}(0)}\|A\|_{H^{s+6}(2) \cap H^{s+10}(0)}
\]

we have with (57) that

\[
\|A(T)\|_{H^{s+6}(2) \cap H^{s+10}(0)} \leq C + C \int_0^T \|A(s)\|_{H^{s+6}(2) \cap H^{s+10}(0)}ds
\]

for all \(T \in [0, T_0]\). With Gronwall’s inequality the assertion follows. \(\blacksquare\)

Thus, Lemma 2.6 ensures that in the long wave limit the initial conditions \(\phi_1\) and \(\phi_2\) of (56) can be separated into \(A_1|_{T=0}\) and \(A_2|_{T=0}\) with an error \(\varepsilon^{7/2}R|_{t=0}\) such that Theorem 6.1 can be applied. Then by Lemma 6.3 we construct via (52) a formal approximation.
of the solutions of (56) for all $t \in [0, T_0/\varepsilon^3]$ with $T_0 > 0$ finite, but arbitrarily large. Then Theorem 6.1 ensures that the water wave problem (56) in the long wave limit really can be described approximately by two decoupled KdV-equations (50). So the only remaining step in the proof of Theorem 1.3 is to establish the validity of Theorem 6.1.

**Remark 6.4** In the following many estimates have to be made. In order to avoid each time restating all of the quantifiers explicitly, we use the following specification

(spec) For all $C_R > 0$ there exist $\varepsilon_0 > 0$ and constants $C > 0$ such that the following holds for all $\varepsilon \in (0, \varepsilon_0)$ and $t \geq 0$ as long as $\sup_{\tau \in (0, t)} \| R_\varepsilon \|_{L^2} \leq C_R$.

**Proof of Theorem 6.1.** We proceed as follows. We write a solution $\mathcal{W} = (Z_1, X_2, U_1)$ of (56) as a sum of the approximation $\varepsilon^2 \Psi$ and an error $\varepsilon^\beta R$ with

$$
\varepsilon^2 \Psi = (\varepsilon^2 \psi_{Z_1}, \varepsilon^2 \psi_{X_2}, \varepsilon^2 \psi_{U_1}) \quad \text{and} \quad \varepsilon^\beta R = (\varepsilon^\beta R_{Z_1}, \varepsilon^\beta R_{X_2}, \varepsilon^\beta R_{U_1}),
$$

and we write a solution $\mathcal{W}_e = (Z_1, X_2, U_1, V_1)$ of (34) as a sum of the approximation $\varepsilon^2 \Psi_e$ and an error $\varepsilon^\beta R_e$ with

$$
\varepsilon^2 \Psi_e = (\varepsilon^2 \psi_{Z_1}, \varepsilon^2 \psi_{X_2}, \varepsilon^2 \psi_{U_1}, \varepsilon^3 \psi_{V_1}) \quad \text{and} \quad \varepsilon^\beta R_e = (\varepsilon^\beta R_{Z_1}, \varepsilon^\beta R_{X_2}, \varepsilon^\beta R_{U_1}, \varepsilon^\beta R_{V_1}).
$$

As we have remarked previously, the main difficulty in proving the result stated in Theorem 6.1 comes from the fact that the method explained at the end of Section 2 for obtaining the error estimates and the local existence and uniqueness theory from Section 4 do not fit together. We solve this problem by using (56) for the evolution of $\mathcal{W}$ and (34) for the evolution of the derivatives of $\mathcal{W}$ and for $V_1$.

We expand (56) up to terms of quadratic order

$$
\begin{align*}
\partial_t Z_1 & = \mathcal{K}_U U_1 \\
\partial_t X_2 & = \mathcal{K}_U U_1 + \mathcal{M}_1(Z_1, \partial_a U_1) - (X_2 + \mathcal{K}_U (X_2 \mathcal{K}_U)) \partial_a U_1 + O(\| \mathcal{W} \|^3) \\
\partial_t U_1 & = -\partial_a X_2 - (\mathcal{M}_2 Z_1) \partial_a X_2 + (\partial_a X_2) \mathcal{K}_U \partial_a X_2 + O(\| \mathcal{W} \|^3).
\end{align*}
$$

**Remark 6.5** In order to make clear, where the smallness of several terms comes from we we consider (56) also in the long wave limit. We obtain

$$
\begin{align*}
\partial_t X_2 & = -\partial_a U_1 - \varepsilon^2 \frac{1}{3} \partial^3 U_1 - \varepsilon^2 Z_1 \partial_a U_1 - \varepsilon^2 X_2 \partial_a U_1 - \varepsilon^4 \partial_a (X_2 \partial_a) \partial_a U_1 + O(\varepsilon^4) \\
& = -\partial_a U_1 - \varepsilon^2 \frac{1}{3} \partial^3 U_1 - 2 \varepsilon^2 X_2 \partial_a U_1 + O(\varepsilon^4) \\
\partial_t U_1 & = -\partial_a X_2 - \varepsilon^2 Z_1 \partial_a X_2 + O(\varepsilon^4) \\
& = -\partial_a X_2 - \varepsilon^2 X_2 \partial_a X_2 + O(\varepsilon^4),
\end{align*}
$$

50
where we have used $Z_1 = X_2 + O(\varepsilon^2)$. With $\partial_{\alpha} U_1 = \partial_{\alpha} \partial_{\alpha} X_1 = -\partial_{\alpha} Z_1 = -\partial_{\alpha} X_2 + O(\varepsilon^2)$, we obtain

$$\partial_{\alpha} X_2 = -\partial_{\alpha} U_1 - \varepsilon^2 \frac{1}{3} \partial_{\alpha}^3 U_1 + 2 \varepsilon^2 X_2 \partial_{\alpha} X_2 + O(\varepsilon^4).$$

With the normal form transform $Y_2 = X_2 - \varepsilon^2 (X_2)^2$ we obtain

$$\partial_{\alpha} Y_2 = -\partial_{\alpha} U_1 - \varepsilon^2 \frac{1}{3} \partial_{\alpha}^3 U_1 + O(\varepsilon^4)$$

$$\partial_{\alpha} U_1 = -\partial_{\alpha} Y_2 - \varepsilon^2 \partial_{\alpha} (Y_2)^2 + O(\varepsilon^4)$$

Using $A_1 = U_1 + Y_2$ and $A_2 = U_1 - Y_2$ we obtain approximately system (5) from [Sch98] and thus we use the ideas of that paper to guide our search for an appropriate energy functional for (56).

To implement the preceding remark rigorously, insert the ansatz $\Psi = \varepsilon^2 \Psi + \varepsilon^\beta R$ into (58) yields

\begin{align*}
\partial_t R_{Z_1} &= \kappa_0 R_{U_1} \\
\partial_t R_{X_2} &= \kappa_0 R_{U_1} + \varepsilon^2 \lambda_1 (\psi_{Z_1}, \partial_{\alpha} R_{U_1}) - \varepsilon^2 (\psi_{X_2} + \kappa_0 (\psi_{X_2} \kappa_0)) \partial_{\alpha} R_{U_1} \\
&+ \varepsilon^3 \lambda_1 (R_{Z_1}, \partial_{\alpha} \psi_{U_1}) - \varepsilon^3 (R_{X_2} + \kappa_0 (R_{X_2} \kappa_0)) \partial_{\alpha} \psi_{U_1} + \mathcal{N}_1 \\
&= \kappa_0 R_{U_1} + \varepsilon^2 \lambda_1 (\psi_{Z_1}, \partial_{\alpha} R_{X_2}) - \varepsilon^2 (\psi_{X_2} + \kappa_0 (\psi_{X_2} \kappa_0)) \partial_{\alpha} R_{U_1} + \mathcal{N}_2 \\
&= \kappa_0 R_{U_1} + \varepsilon^2 s_1(\Psi, \partial_{\alpha} R_{U_1}) + \mathcal{N}_2 \\
\partial_t R_{U_1} &= -\partial_{\alpha} R_{X_2} - \varepsilon^2 (\lambda_2 \psi_{Z_1}) \partial_{\alpha} R_{X_2} + \varepsilon^3 (\partial_{\alpha} \psi_{X_2}) \kappa_0 \partial_{\alpha} R_{X_2} \\
&- \varepsilon^3 (\lambda_2 R_{Z_1}) \partial_{\alpha} \psi_{X_2} + \varepsilon^3 (\partial_{\alpha} \psi_{X_2}) \kappa_0 \partial_{\alpha} \psi_{X_2} + \mathcal{N}_3 \\
&= -\partial_{\alpha} R_{X_2} - \varepsilon^2 (\lambda_2 \psi_{Z_1}) \partial_{\alpha} R_{X_2} + \mathcal{N}_3 
\end{align*}

with

\begin{align*}
\|\mathcal{N}_1\|_{H^r} &\leq C (\varepsilon^4 \|R\|_{H^r} + \varepsilon^\beta \|R\|_{H^r}^2 + \varepsilon^3 C_{res} + q(t)) \\
\|\mathcal{N}_2\|_{H^r} &\leq C (\varepsilon^3 \|R\|_{H^r} + \varepsilon^\beta \|R\|_{H^r}^2 + \varepsilon^3 C_{res} + q(t)) \\
\|\mathcal{N}_3\|_{H^{r-1}} &\leq C (\varepsilon^4 \|R\|_{H^r} + \varepsilon^\beta \|R\|_{H^r}^2 + \varepsilon^3 C_{res} + q(t)) \\
\|\mathcal{N}_4\|_{H^{r-1}} &\leq C (\varepsilon^3 \|R\|_{H^r} + \varepsilon^\beta \|R\|_{H^r}^2 + \varepsilon^3 C_{res} + q(t)) \\
\varepsilon^2 s_1(\Psi, \partial_{\alpha} R_{U_1}) &= \varepsilon^2 \lambda_1 (\psi_{Z_1}, \partial_{\alpha} R_{U_1}) - \varepsilon^2 (\psi_{X_2} + \kappa_0 (\psi_{X_2} \kappa_0)) \partial_{\alpha} R_{U_1}
\end{align*}

under the specification (spec). The term $s_1$ is bilinear in its arguments with the regularity of $\psi$, i.e $\partial_{\alpha} R_{U_1}$ is smoothed by this operator according to Lemma 3.12 and Remark 3.9. Next we replace $\partial_{\alpha} R_{U_1}$ in $\varepsilon^2 s_1(\psi, \partial_{\alpha} R_{U_1})$ by

$$\partial_{\alpha} R_{U_1} = (\partial_{\alpha} \kappa_0^{-1}) (\partial_t R_{X_2} - \varepsilon^2 s_1(\Psi, \partial_{\alpha} R_{U_1}) + \mathcal{N}_2)$$
and so
\[
\partial_t R_{X_2} = K_0 R_{U_1} + \varepsilon^2 s_1(\Psi, (\partial_\alpha K_0^{-1}) R_{X_2}) + N_{2a},
\]
where \( N_{2a} \) obeys the same estimates as \( N_2 \). With \( \partial_t \Psi = O(\varepsilon) \) we obtain
\[
\partial_t (R_{X_2} - \varepsilon^2 s_1(\Psi, (\partial_\alpha K_0^{-1}) R_{X_2})) = K_0 R_{U_1} + N_{2b},
\]
where \( N_{2b} \) obeys the same estimates as \( N_2 \). Then we make the normalform transform
\[
R_{Y_2} = R_{X_2} - \varepsilon^2 s_1(\Psi, (\partial_\alpha K_0^{-1}) R_{X_2})
\]
(59)
or equivalently
\[
R_{X_2} = R_{Y_2} + \varepsilon^2 s_1(\Psi, (\partial_\alpha K_0^{-1}) R_{Y_2}) + O(\varepsilon^4).
\]
which is of the form identity plus something small and smooth in \( R_{Y_2} \). Thus, we obtain the system
\[
\begin{align*}
\partial_t R_{Z_1} &= K_0 R_{U_1} \\
\partial_t R_{Y_2} &= K_0 R_{U_1} + N_5, \\
\partial_t R_{U_1} &= -\partial_\alpha R_{Y_2} + \varepsilon^2 s_2(\Psi, \partial_\alpha R_{Y_2}) + N_{4a} 
\end{align*}
\]
with
\[
\begin{align*}
\|N_{4a}\|_{H^{-1}} &\leq C(\varepsilon^3 \|R\|_{H^0} + \varepsilon^3 \|R\|_{H^0}^2 + \varepsilon^3 C_{\varepsilon, \delta} + q(t)) \\
\|N_5\|_{H^0} &\leq C(\varepsilon^3 \|R\|_{H^0} + \varepsilon^3 \|R\|_{H^0}^2 + \varepsilon^3 C_{\varepsilon, \delta} + q(t)) \\
\varepsilon^2 s_2(\Psi, \partial_\alpha R_{Y_2}) &\leq -\varepsilon^2 \partial_\alpha (s_1(\Psi, \partial_\alpha K_0^{-1} R_{Y_2})) - \varepsilon^2 s_2(\Psi, \partial_\alpha R_{Y_2}) 
\end{align*}
\]
which holds under (spec). It is easy to see that the \( H^{s-1/2} \)-norm of \( R_{U_1} \) cannot be estimated by the term \( N_4 \) since we would loose regularity. Therefore, in the highest derivative we have to use the equation \( \partial_t U_1 = V_1 \) and extend (56) by the fourth equation of (34). Thus, we must supplement this system with the last two equations of (34). For clarity, we remark again that (34) describes in the subset \( \mathcal{C}_p \) of solutions of the water wave problem the same solutions as (56). Expanding the last two equations of (34) up to terms of quadratic order gives
\[
\begin{align*}
\partial_t U_1 &= V_1 \\
\partial_t V_1 &= -(1 + \mathcal{M}_2 Z_1 + K_0 V_1 + N_6(\mathcal{V}_c)) K_0 \partial_\alpha U_1 - V_1 (1 + K_0^2) \partial_\alpha U_1 \\
&\quad - \partial_\alpha \mathcal{M}_1 (Z_1, \partial_\alpha U_1) + \partial_\alpha (X_2 + K_0 (X_2 K_0)) \partial_\alpha U_1 + O(\|\mathcal{V}_c\|_{H^0}^3), \\
\end{align*}
\]
where
\[
\begin{align*}
\|N_6(\mathcal{V}_c)\|_{H^{-1}} &= \|f_6(\mathcal{V}_c) - 1 - \mathcal{M}_2 Z_1 - K_0 V_1\|_{H^{-1}} = O(\|\mathcal{V}_c\|_{H^0}^3)
\end{align*}
\]
52
The vector $R$ is changed to $\tilde{R} = (R_{Z_1}, R_{Y_2}, R_{U_1})$ and extended to $\tilde{R}_c = (R_{Z_1}, R_{Y_2}, R_{U_1}, R_{V_1})$. For $\varepsilon$ sufficiently small we have that $R$ can be estimated by $\tilde{R}$ and vice versa.

Writing $V_1 = \varepsilon^3 \psi_1 + \varepsilon^2 R_{V_1}$ yields

\[
\begin{align*}
\partial_t R_{U_1} &= R_{V_1} \\
\partial_t R_{V_1} &= -(1 + M_2 Z_1 + K_0 V_1 + N_6(W_e))K_0 \partial_\alpha R_{U_1} - \varepsilon^3 \psi_1 (1 + K_0^2) \partial_\alpha R_{U_1} \\
&\quad - \varepsilon^2 \partial_\alpha M_1(\psi Z_1, \partial_\alpha R_{U_1}) + \varepsilon^2 \partial_\alpha (\psi X_2 + K_0(\psi X_2 K_0)) \partial_\alpha R_{U_1} \\
&\quad + \varepsilon^3 (M_2 R_{Z_1} + K_0 R_{V_1} + \varepsilon^{-2} N_6(W_e) - \varepsilon^{-2} N_6(\varepsilon^2 \Psi_e))K_0 \partial_\alpha \psi_{U_1} \\
&\quad - \varepsilon^3 R_{V_1}(1 + K_0^2) \partial_\alpha \psi_{U_1} \\
&\quad - \varepsilon^3 \partial_\alpha M_1(R_{Z_1}, \partial_\alpha \psi_{U_1}) + \varepsilon^3 \partial_\alpha (R_{Y_2} + K_0(R_{Y_2} K_0)) \partial_\alpha \psi_{U_1} + N_7 \\
&\quad = -(1 + M_2 Z_1 + K_0 V_1 + N_6(W_e))K_0 \partial_\alpha R_{U_1} \\
&\quad - \varepsilon^2 \partial_\alpha M_1(\psi Z_1, \partial_\alpha R_{U_1}) + \varepsilon^2 \partial_\alpha (\psi X_2 + K_0(\psi X_2 K_0)) \partial_\alpha R_{U_1} + N_8 \\
&\quad = -(1 + M_2 Z_1 + K_0 V_1 + N_6(W_e))K_0 \partial_\alpha R_{U_1} + \varepsilon^2 s_3(\Psi, \partial_\alpha R_{U_1}) + N_8,
\end{align*}
\]

where we have

\[
\begin{align*}
\|N_7\|_{H^{-1}} &\leq C(\varepsilon^4 \|R_c\|_H^2 + \varepsilon^3 \|R_c\|_H^2 + \varepsilon^3 C_{res} + q(t)) \\
\|N_8\|_{H^{-1}} &\leq C(\varepsilon^3 \|R_c\|_H^2 + \varepsilon^2 \|R_c\|_H^2 + \varepsilon^3 C_{res} + q(t)) \\
s_3 &= -\partial_\alpha M_1(\psi Z_1, \partial_\alpha R_{U_1}) + \partial_\alpha (\psi X_2 + K_0(\psi X_2 K_0)) \partial_\alpha R_{U_1}
\end{align*}
\]

under (spec). Again we replace $\partial_\alpha R_{U_1}$ in $s_3$ by

\[
\partial_\alpha R_{U_1} = (\partial_\alpha K_0^{-1})(\partial_1 R_{Y_2} - N_5).
\]

By using $\partial_1 \Psi = \mathcal{O}(\varepsilon)$ we obtain

\[
\partial_t R_{V_1} = -(1 + M_2 Z_1 + K_0 V_1 + N_6(W_e))K_0 \partial_\alpha R_{U_1} + \varepsilon^2 \partial_1 s_3(\Psi, \partial_\alpha K_0^{-1} R_{Y_2}) + N_{8a},
\]

where $N_{8a}$ obeys the same estimates as $N_8$. With the normalform transform

\[
R_{W_1} = R_{V_1} - \varepsilon^2 s_3(\Psi, \partial_\alpha K_0^{-1} R_{Y_2})
\]

which is again of the form identity plus something smooth and small, we end up with

\[
\partial_t R_{W_1} = -(1 + M_2 Z_1 + K_0 W_1 + N_6(W_e))K_0 \partial_\alpha R_{U_1} + N_9
\]

\[
= -f_0 K_0 \partial_\alpha R_{U_1} + N_9,
\]

where $N_9$ again obeys the same estimates as $N_8$, i.e.

\[
\|N_9\|_{H^{-1}} \leq C(\varepsilon^3 \|R_c\|_H^2 + \varepsilon^2 \|R_c\|_H^2 + \varepsilon^3 C_{res} + q(t))
\]

53
under (spec). We collected all semilinear terms in $N_o$, but the quasilinear terms are still written explicitly. Now collect the new variables in $\tilde{R}_e = (R_{Z_1}, R_{Y_2}, R_{W_1})$. Again for small $\varepsilon$ the vector $\tilde{R}_e$ can be estimated by $R_e$ and vice versa.

In the next step we construct a new scalar product $E_s(\cdot, \cdot)$ for the $(X_2, U_1, V_1)$-variables. It is equivalent to the usual $H^s = H^s \times H^{s-1/2} \times H^{s-1}$-scalar product. We define $R_i = (R_{X_2}, R_{U_1}, R_{V_1})$ the lower part of $R_e$. As explained in Remark 4.3 the main part of this scalar product must consist of $E_s(\cdot, \cdot)$ such that in the computation of $\frac{d}{dt}E_s(R_t, R_i)$ all quasilinear terms from (34) cancel and that $\frac{d}{dt}E_s(R_t, R_i)$ can be estimated in terms of $E_s(R_t, R_i)$ such that Gronwall’s inequality can be applied. On the other hand, in order to apply the ideas of the end of Section 2 we have to modify $E_s(\cdot, \cdot)$ so that the influence of the linear terms of order $O(\varepsilon^2\|R_t\|_{\mathcal{H}^s})$ measured with $E_s(\cdot, \cdot)$ is of order $O(\varepsilon^3)$. Finally, in the $(X_2, U_1)$-part the scalar product $E_s(\cdot, \cdot)$ has to be modified so that the linear terms of order $O(\varepsilon^3\|R_t\|_{\mathcal{H}^s})$ cancel.

Now we estimate $R_{Z_1}$ by $R_{Y_2}$. From the first two equations for the error it follows that

$$\partial_t R_{Y_2} - \partial_t R_{Z_1} = N_{10},$$

where

$$\|N_{10}\|_{H^s} \leq C(\varepsilon^3\|R\|_{\mathcal{H}^s} + \varepsilon^\beta\|R\|_{\mathcal{H}^s}^2 + \varepsilon^3C_{r_0} + q(t)).$$

Integration with respect to time and Gronwall’s inequality show the estimate

$$\forall C_A \exists C_4, C_5 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0) : \|R_{Z_1}(t)\|_{H^s} \leq C_4 + C_5\|R(t)\|_{\mathcal{H}^s}$$

as long as $\|R(t)\|_{\mathcal{H}^s} \leq C_R$.

For $\mathcal{W}_e \in C_p$ we can relate $\partial_x X_2$ to $V_1$. Thus by the implicit function theorem we obtain

$$\partial_x R_{X_2} = -R_{V_1} + \tilde{N}_{12}$$

or equivalently

$$\partial_x R_{Y_2} = -R_{W_1} + \tilde{N}_{12}$$

with

$$\|\tilde{N}_{12}\|_{H^{s-1}} + \|N_{12}\|_{H^{s-1}} \leq C(\varepsilon^2\|R\|_{\mathcal{H}^s} + \varepsilon^\beta\|R\|_{\mathcal{H}^s}^2 + \varepsilon^3C_{r_0} + q(t)).$$

In particular, we can estimate the highest derivatives $\partial_x^s R_{Y_2}$ appearing in the following by

$$\|\partial_x^s R_{Y_2}\|_{L^2} \leq \|\partial_x^{s-1} R_{W_1}\|_{L^2} + \|\partial_x^{s-1} N_{12}\|_{L^2},$$

so we can control $\|R_{Y_2}\|_{H^s}$ by $\|R_{Y_2}\|_{L^2}$, $\|R_{V_1}\|_{H^{s-1/2}}$, and $\|R_{W_1}\|_{H^{s-1}}$. 

54
Then we consider
\[
\partial_t \int (M_2^{1/2} R_{Y_2})^2 d\alpha/2 = \int R_{Y_2} (M_2 \partial_\alpha R_{Y_2}) d\alpha = \int R_{Y_2} (M_2 (\mathcal{K}_0 R_{Y_1} + \mathcal{N}_5)) d\alpha = \int (\partial_\alpha R_{Y_2}) R_{Y_1} d\alpha + \mathcal{N}_{\varepsilon_1},
\]
where
\[
|\mathcal{N}_{\varepsilon_1}| \leq C(\varepsilon^3 \|R_1\|_H^2 + \varepsilon^\beta \|R_1\|_H^3 + \varepsilon^3 C_{\text{res}} \|R_1\|_H + q(t) \|R_1\|_H^2)
\]
under (spec). Next we have
\[
\partial_t \int (R_{Y_1})^2 d\alpha/2 = \int R_{Y_1} (\partial_\alpha R_{Y_1}) d\alpha = \int R_{Y_1} (-\partial_\alpha R_{Y_2} + \varepsilon^2 s_2(\Psi, \partial_\alpha R_{Y_2}) + \mathcal{N}_{4a}) d\alpha
\]
\[
= -\varepsilon^2 \int R_{Y_1} (\partial_\alpha R_{Y_2}) d\alpha - \varepsilon^2 \int R_{Y_1} \psi_{Z_1} \partial_\alpha R_{Y_2} d\alpha - \varepsilon^2 \int R_{Y_1} \partial_\alpha s_1(\Psi, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) d\alpha + \int R_{Y_1} \mathcal{N}_{4a} d\alpha
\]
\[
= -\varepsilon^2 \int R_{Y_1} (\partial_\alpha R_{Y_2}) d\alpha - \varepsilon^2 \int R_{Y_1} \psi_{Z_1} \partial_\alpha R_{Y_2} d\alpha + \varepsilon^2 \int R_{Y_1} \partial_\alpha s_1(\Psi, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) d\alpha + \int R_{Y_1} \mathcal{N}_{4a} d\alpha
\]
\[
= -\varepsilon^2 \int R_{Y_1} (\partial_\alpha R_{Y_2}) d\alpha - \varepsilon^2 \int R_{Y_1} \psi_{Z_1} \partial_\alpha R_{Y_2} d\alpha + \varepsilon^2 \int (\partial_\alpha \mathcal{K}_{0}^{-1} \partial_\alpha R_{Y_2}) s_1(\Psi, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) d\alpha + \mathcal{N}_{\varepsilon_2},
\]
Where we used that \(\partial_\alpha R_{Y_1} = \partial_\alpha \mathcal{K}_{0}^{-1} \partial_\alpha R_{Y_2} + \mathcal{O}(\varepsilon^2)\) and \(\mathcal{N}_{\varepsilon_2}\) satisfies the estimates
\[
|\mathcal{N}_{\varepsilon_2}| \leq C(\varepsilon^3 \|R_1\|_H^2 + \varepsilon^\beta \|R_1\|_H^3 + (\varepsilon^3 C_{\text{res}} + q(t)) \|R_1\|_H^2)
\]
Using the fact that the time derivatives of \(\Psi\) are \(\mathcal{O}(\varepsilon)\), and the fact that
\[
\int (\partial_\alpha \mathcal{K}_{0}^{-1} \partial_\alpha R_{Y_2}) s_1(\Psi, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) d\alpha = \int (\partial_\alpha \mathcal{K}_{0}^{-1} \partial_\alpha R_{Y_2}) \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_1}) d\alpha
\]
\[
+ \int (\partial_\alpha \mathcal{K}_{0}^{-1} \partial_\alpha R_{Y_2}) (\psi_{X_2} + \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0)) \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_1} d\alpha
\]
with \(\mathcal{M}_1(\psi_{Z_1}, \cdot)\) and \((\psi_{X_2} + \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0))\) symmetric, smoothing operators, the terms of \(\mathcal{O}(\varepsilon^2)\) in this expression can be written as
\[
\frac{\varepsilon^2}{2} \partial_t \{ \int \psi_{Z_1} (R_{Y_1})^2 d\alpha + \int (\partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) s_1(\Psi, \partial_\alpha \mathcal{K}_{0}^{-1} R_{Y_2}) d\alpha \} + \mathcal{O}(\varepsilon^3)
\]
More precisely, we have
\[
\partial_t \left\{ \frac{1}{2} \int (R_{U_1})^2 d\alpha - \frac{\varepsilon^2}{2} \int \psi Z_1 (R_{U_1})^2 d\alpha - \frac{\varepsilon^2}{2} \int (\partial_a K_0^{-1} R_{Y_2}) s_1 (\Psi, \partial_a K_0^{-1} R_{Y_2}) d\alpha \right\} = N_{e4}
\]
With the estimate
\[
|N_{e4}| \leq C (\varepsilon^3 \| R_t \|_{H_1}^2 + \varepsilon^\beta \| R_t \|_{H_1}^3 + (\varepsilon^3 C_{\psi es} + q(t)) \| R_t \|_{H_1}^2)
\]
(64)

For the quasilinear terms we obtain
\[
\partial_t \int \sum_{j=0}^{s-1} (\partial_a^{i} R_{W_1})^2 d\alpha / 2 = \int \sum_{j=0}^{s-1} (\partial_a^{i} R_{W_1}) (\partial_a^{i} \partial_t R_{W_1}) d\alpha
\]
\[
= \int \sum_{j=0}^{s-1} (\partial_a^{i} R_{W_1}) (\partial_a^{i} (- f_6 K_0 \partial_a R_{U_1} + N_9)) d\alpha
\]
\[
= \int \sum_{j=0}^{s-1} (\partial_a^{i} R_{W_1}) (\partial_a^{i} (f_6 \lambda^2 R_{U_1} + N_9)) d\alpha
\]
\[
= \int \sum_{j=0}^{s-1} (\partial_a^{i} R_{W_1}) f_6 \partial_a^{i} \lambda^2 R_{U_1} d\alpha + N_{e5},
\]
where (64) also holds for $N_{e5}$.

**Remark 6.6** We have used that $f_6 = 1 + \varepsilon^2 M_2 \psi Z_1 + \varepsilon^3 K_0 \psi V_1 + \tilde{N}_6$, where $\tilde{N}_6$ satisfies the same estimates as $N_6$, and so
\[
\| \partial_a^{i} f_6 \|_{L^2} \leq C (\varepsilon^3 \| R_t \|_{H_1}^2 + \varepsilon^\beta \| R_t \|_{H_1}^3)
\]
for $j = 0, \ldots, s - 1$. We also recall that $\lambda(k)$ is the operator with Fourier multiplier $i \sqrt{k \tanh(k)}$ which implies that $\lambda$ is skew symmetric.

The counterpart for the cancellation of the quasilinear terms looks like
\[
\partial_t \int \sum_{j=0}^{s-1} (\partial_a^{i} \lambda R_{U_1})^2 f_6 d\alpha / 2 = \int \sum_{j=0}^{s-1} (\partial_a^{i} \lambda R_{U_1})^2 (\partial_t f_6) d\alpha / 2
\]
\[
+ \int \sum_{j=0}^{s-1} f_6 (\partial_a^{i} \lambda R_{U_1}) (\partial_a^{i} \lambda \partial_t R_{U_1}) d\alpha
\]
\[
= \int \sum_{j=0}^{s-1} f_6 (\partial_a^{i} \lambda R_{U_1}) (\partial_a^{i} \lambda R_{W_1}) d\alpha + N_{e6}
\]

The estimate (64) also holds for $N_{e6}$, where we used
\[
\| \partial_t f_6 \|_{H^1} \leq C (\varepsilon^3 \| R_t \|_{H_1}^2 + \varepsilon^\beta \| R_t \|_{H_1}^3 + \varepsilon^3 C_{\psi}).
\]

56
The term
\[ \int \sum_{j=0}^{s-1} f_6(\partial_{\alpha}^j \lambda R_{U_1})(\partial_{\alpha}^j \lambda R_{W_1}) d\alpha/2 \]
can be written in Fourier space as
\[ \int \int \hat{f}_6(k-l)(\hat{\lambda}(l)(il)^{s-1} \hat{R}_{U_1}(l))(\hat{\lambda}(k)(ik)^{s-1} \hat{R}_{V_1}(k)) \]
\[ = \int \int \hat{f}_6(k-l)(\hat{\lambda}(l) - \hat{\lambda}(k))\hat{\lambda}(l)(il)^{s-1} \hat{R}_{U_1}(l)((ik)^{s-1} \hat{R}_{V_1}(k)) \]
\[ + \int \int \hat{f}_6(k-l)(-\hat{\lambda}(l^2)(il)^{s-1} \hat{R}_{U_1}(l)((ik)^{s-1} \hat{R}_{V_1}(k)), \]
where we have used \( \hat{\lambda}(k) = -\hat{\lambda}(k) \). Since a) \( f_6 = 1 + \epsilon^2 M_{2\psi Z_1} + \epsilon^3 K_0\psi V_1 + \tilde{N}_6 \), b) \( |\lambda(k) - \lambda(l)| \leq C|k-l| \), and c) \( i k \hat{\Psi} = O(\epsilon) \) we finally have
\[ \partial_t \int \sum_{j=0}^{s-1} (\partial_{\alpha}^j \lambda R_{U_1})^2 f_6 d\alpha/2 = - \int \sum_{j=0}^{s-1} f_6(\partial_{\alpha}^j \lambda^2 R_{U_1})(\partial_{\alpha}^j R_{W_1}) d\alpha + N_{c7}, \]
where estimate (64) also holds for \( N_{c7} \).

If we define \( E_s(\cdot, \cdot, \cdot) \) in the following way we find that all terms of order \( O(\epsilon^2) \) in \( \frac{d}{dt} E_s(\mathcal{W}_t, \mathcal{W}_t) \) cancel. We define
\[ E_s(R_t, R_t) = \int (M_{1/2}^1 R_{Y_2})^2 + \int (R_{U_1})^2 \]
\[ - \frac{\epsilon^2}{2} \int \psi_{Z_1} (R_{U_1})^2 - \frac{\epsilon^2}{2} \int (\partial_{\alpha} K_{0}^{-1} R_{Y_2}) s_1(\Psi, \partial_{\alpha} K_{0}^{-1} R_{Y_2}) \]
\[ + \int \sum_{j=0}^{s-1} (\partial_{\alpha}^j R_{W_1})^2 - \int \sum_{j=0}^{s-1} (\lambda \partial_{\alpha}^j R_{U_1})^2 f_6. \]

We easily see by using (61) and (62) that the scalar product \( E_s(\cdot, \cdot, \cdot) \) is equivalent to the usual \( \mathcal{H}_0^2 \) scalar product, i.e. there exist positive constants \( c_1 \) and \( c_2 \) and an \( \epsilon_0 > 0 \) such that for all \( \epsilon \in (0, \epsilon_0) \) we have
\[ ||\mathcal{W}_t||_{\mathcal{H}_0^2}^2 \leq c_1 E_s(\mathcal{W}_t, \mathcal{W}_t) \leq c_2 ||\mathcal{W}_t||_{\mathcal{H}_0^2}^2. \] (65)

Therefore, we can sum up our above estimates and obtain with constants \( C_1 = C_1(C_{\psi}, C_{Res}, c_j), C_2 = C_2(C_{\psi}, C_{Res}, C_{R}, c_j), \) and \( C_3 = C_3(C_{\psi}, C_{Res}, c_j), \) that
\[ \frac{1}{2} \partial_t E_s(R_t, R_t) \leq \epsilon^3 C_1 E_s(R_t, R_t) + \epsilon^3 C_2 E_s(R_t, R_t)^{3/2} + (\epsilon^3 C_3 + q(t)) E_s(R_t, R_t)^{1/2} \]
\[ \leq \epsilon^3 C_1 E_s(R_t, R_t) + \epsilon^3 C_2 E_s(R_t, R_t)^{3/2} \]
\[ + (C_3 \epsilon^3 + q(t)) + (\epsilon^3 C_3 + q(t)) E_s(R_t, R_t), \]
57
i.e. for \( y(T) = E_s(R_l(t), R_r(t)) \) with \( T = \varepsilon^3 t \) we obtain the ordinary differential equation

\[
y = (C_1 + C_3 + \varepsilon^{-3}q(t))y + \varepsilon^0 C_2 y^{3/2} + C_3 + \varepsilon^{-3}q(t)
\leq (C_1 + C_3 + 1 + \varepsilon^{-3}q(t))y + C_3 + \varepsilon^{-3}q(t)
\]

Applying Gronwall’s inequality shows for all \( t \in [0, T_0/\varepsilon^3] \) that

\[
y_{\varepsilon}(t) = y_{\varepsilon}(0) + \int_0^T e^{-\int_0^t (C_1 + C_3 + 1 + \varepsilon^{-3}q(\varepsilon^{-3}s))ds} (C_3 + 1 + \varepsilon^{-3}q(\varepsilon^{-3}s))ds,
\]

where we have chosen \( \varepsilon_0 > 0 \) so small that \( \varepsilon^{\theta-3}C_2(C_R)(c_2C_R)^2 \leq 1 \), where \( C_R = 2 \limsup_{\varepsilon \to 0} y_{\varepsilon}(T_0) = O(1) \). This yields with (59), (60), (61), (62), and (65) that

\[
\sup_{\varepsilon \in [0,T_0/\varepsilon^3]} \| R_{\varepsilon} \|_{L^2}^2 \leq c_1 C_1.
\]

Therefore, we are done. \( \square \)

References


59


60


