

The Rigorous Approximation of Long-Wavelength Capillary-Gravity Waves

GUIDO SCHNEIDER & C. EUGENE WAYNE

Communicated by G. FRIESECKE

Abstract

In a previous paper we proved that long-wavelength solutions of the water-wave problem in the case of zero surface tension 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 (KdV) equation. In this paper we examine the effect of surface tension on this scenario. We find that we obtain three different physical regimes depending on the strength of the surface tension. For weak surface tension, the propagation of the wave packets is very similar to that in the zero surface tension case. For strong surface tension, the evolution is again governed by a pair of KdV equations, but the coefficients in these equations have changed in such a way that the KdV soliton now represents a wave of *depression* on the fluid surface. Finally, at a special, intermediate value of the surface tension (where the Bond number equals $\frac{1}{3}$) the KdV description breaks down and it is necessary to introduce a new approximating equation, the Kawahara equation, which is a fifth order, nonlinear partial differential equation. In each of these regimes we give rigorous estimates of the difference between the solution of the appropriate modulation equation and the solution of the true water-wave problem.

1. Introduction

We consider the irrotational flow of an incompressible, inviscid fluid in an infinitely long canal of fixed depth with impermeable bottom under the influence of gravity and surface tension. We prove that the evolution of long-wavelength initial data can be accurately approximated for very long times by the motion of a pair of wave packets, one moving to the right and one moving to the left, and depending on the values of the various physical parameters in the problem, each of these wave packets evolves according to either the Korteweg-de Vries or Kawahara

equations. What is more, to the order of approximation typically considered in the formal derivation of these approximate equations, there is no interaction between the counterpropagating waves.

To state our results more precisely, we introduce coordinates with $x_1 \in \mathbb{R}$ and x_2 in the bounded direction. The fluid fills the domain $\Omega(t)$ between the bottom $\{(\alpha, -1) | \alpha \in \mathbb{R}\}$ and the free surface

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

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

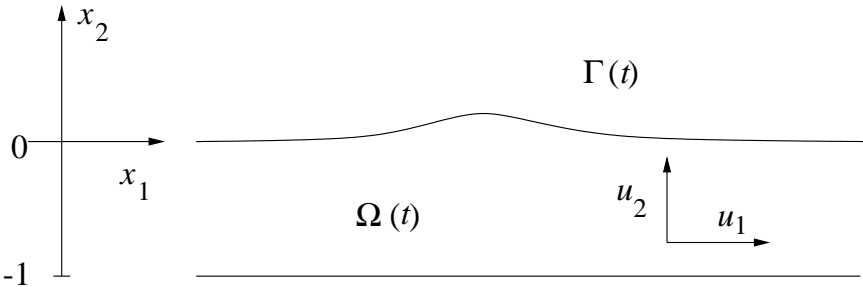


Fig. 1. The water-wave problem.

It turns out that under these assumptions [Yo83, equations (1.10)–(1.12)] the dynamics of the problem is completely determined by the evolution of the free surface $\Gamma(t)$ which is governed by the pair of partial differential equations,

$$\partial_t^2 X_1(1 + \partial_\alpha X_1) + \partial_\alpha X_2(1 + \partial_t^2 X_2) = \mu R(\partial_\alpha X, \partial_\alpha^2 X) + \mu S(\partial_\alpha X, \partial_\alpha^3 X), \quad (1)$$

$$\partial_t X_2 = \mathcal{K}(X)\partial_t X_1, \quad (2)$$

where

$$R(\partial_\alpha X, \partial_\alpha^2 X) = -3Q(\partial_\alpha X)^{-5}((1 + \partial_\alpha X_1)\partial_\alpha^2 X_1 + \partial_\alpha X_2\partial_\alpha^2 X_2) \\ \times (-\partial_\alpha X_2\partial_\alpha^2 X_1 + (1 + \partial_\alpha X_1)\partial_\alpha^2 X_2),$$

$$S(\partial_\alpha X, \partial_\alpha^3 X) = Q(\partial_\alpha X)^{-3}(-\partial_\alpha X_2\partial_\alpha^3 X_1 + (1 + \partial_\alpha X_1)\partial_\alpha^3 X_2),$$

$$Q(\partial_\alpha X) = ((1 + \partial_\alpha X_1)^2 + (\partial_\alpha X_2)^2)^{1/2}.$$

The parameter μ is called the Bond number and is proportional to the strength of the surface tension, and the bounded operator \mathcal{K} is related to the Dirichlet-Neumann operator. It depends linearly on $\partial_t X_1$, but nonlinearly on X . Properties of $\mathcal{K}(X)$ are explained at the beginning of Section 2. Also, we have “nondimensionalized” the equations so that the unperturbed fluid depth and gravitational acceleration are both equal to one.

For relatively flat surfaces we could assume without loss of generality an initial parametrization $X_1|_{t=0} = 0$, and so it turns out that we have essentially two independent initial conditions $X_2|_{t=0}$ and $\partial_t X_1|_{t=0}$. However, the freedom to choose

a nonzero initial parametrization for X_1 is essential for our approximation results and we use that freedom in both Theorem 1 and Theorem 2.

Notation. 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 $\|u\|_{H^s} = (\int |\hat{u}(k)|^2 (1 + |k|^2)^s dk)^{1/2}$. Moreover, let $\|u\|_{H^m(n)} = \|\rho^n u\|_{H^m}$, where $\rho(x) = (1 + x^2)^{1/2}$, and let $\|u\|_{C_b^n} = \sum_{j=0}^n \|\partial_x^j u\|_{C_b^0}$, where $\|u\|_{C_b^0} = \sup_{x \in \mathbb{R}} |u(x)|$.

Our goal is to approximate the evolution of solutions of (1) and (2) with general long-wavelength initial conditions. In order to make such an approximation, it is necessary to consider two different approximating equations since, as we explain below, the value of $\mu = \frac{1}{3}$ leads to a degeneracy in the equations. For $\mu \neq \frac{1}{3}$, the evolution of (long-wavelength) solutions of (1) and (2) can be split into a left and right moving wave packet, each of which is described by the Korteweg–de Vries (KdV) equation and there is *no* interaction between the left and right moving waves to the degree of approximation considered here. Nonetheless, as we shall see in the discussion following Theorem 1, the nature of the wave trains described by the two KdV equations is quite different depending on whether or not μ is greater or less than $\frac{1}{3}$. We now state our first approximation result.

Theorem 1. (The nondegenerate case: The KdV limit). *Let $\mu \neq \frac{1}{3}$. 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: Fix $s \geq 6$; then, 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

$$\|(\Phi_1, \Phi_2)\|_{H^{s+6}(2) \cap H^{s+11}(0)} \leq C_1.$$

Define amplitudes $A_1 = A_1(\underline{\alpha}, T) \in \mathbb{R}$ and $A_2 = A_2(\underline{\alpha}, T) \in \mathbb{R}$ satisfying the KdV equations

$$\begin{aligned} 2\partial_T A_1 &= (\mu - \frac{1}{3})\partial_{\underline{\alpha}}^3 A_1 - \frac{3}{2}A_1 \partial_{\underline{\alpha}} A_1 \quad \text{and} \\ 2\partial_T A_2 &= (\frac{1}{3} - \mu)\partial_{\underline{\alpha}}^3 A_2 + \frac{3}{2}A_2 \partial_{\underline{\alpha}} A_2, \end{aligned} \tag{3}$$

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

Then there exists a reparametrization of the fluid surface (denoted with the same symbols) such that the solution of the water wave-problem (1) and (2) satisfies

$$\sup_{t \in [0, T_0/\varepsilon^3]} |X_1(0, t)| \leq C_2 \varepsilon^{1/2}$$

and

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

where

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

Proof. See Section 4. \square

Remark 1. It might seem surprising that the error estimate is $\mathcal{O}(\varepsilon^{7/2})$. On the basis of formal calculations it would be expected to be $\mathcal{O}(\varepsilon^4)$. The loss of a half a power of ε results from the way the Sobolev norms scale — i.e., if the error term is of the form $\varepsilon^4 \mathcal{E}(\varepsilon\alpha)$, for some $\mathcal{E} \in L^2$, then $\|\varepsilon^4 \mathcal{E}(\varepsilon\cdot)\|_{L^2} = \varepsilon^{7/2} \|\mathcal{E}\|_{L^2}$.

While most rigorous studies of the water-wave problem have focused on the Lagrangian description used in Theorem 1, for comparison with experiments it is necessary to re-express these results in terms of Eulerian coordinates. Define $u_1(\tilde{X}_1(\alpha, t), t) = \partial_t X_1(\alpha, t)$ and $\eta(\tilde{X}_1(\alpha, t), t) = X_2(\alpha, t)$. In terms of these variables, we have the following approximation result.

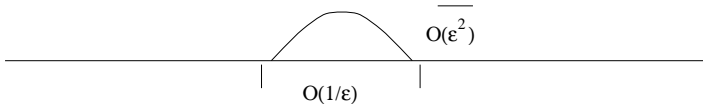
Corollary 1. *Assume that Theorem 1 holds. Then the following estimate holds:*

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

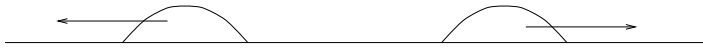
Proof. This follows from Theorem 1 exactly as [SW00a, Corollary 1.5]. \square

The long-time behavior for the water-wave problem in the long-wave limit is sketched in the Figs. 2 and 3. On a time scale $\mathcal{O}(1/\varepsilon)$ the solutions split up into two wave packets one moving to the right and one to the left, regardless of whether μ is greater or less than $\frac{1}{3}$. Each of these wave packets evolve independently as solutions of the KdV equations whose long-time behavior can be computed explicitly with the help of the inverse scattering transform. However, the behavior of these wave packets is dramatically different depending on the value of μ , in spite of the fact that they are all governed by the KdV equation. For $\mu < \frac{1}{3}$ (“weak” surface tension) the KdV solitons correspond to solitary waves of elevation on the fluid surface. The largest solitary waves propagate fastest, followed by dispersive ripples. This is very similar to the case of zero surface tension studied in [SW00a]. However, if $\mu > \frac{1}{3}$, the KdV solitons represent solitary waves of depression on the fluid surface and the largest waves move most slowly, preceded by the dispersive waves. Experimentally, it is difficult to observe these solitary waves of depression on the surface of water at room temperature since values of $\mu > \frac{1}{3}$ correspond to very shallow water depths where dissipation is no longer negligible, and also because the dispersive ripples propagating in front of the solitary wave tend to obscure it. However Benjamin [Be82] reports tentative experimental observation of such solutions. See also the comment before Corollary 2.

$t = 0$



$t = O(1/ε)$



$t = O(1/ε³)$

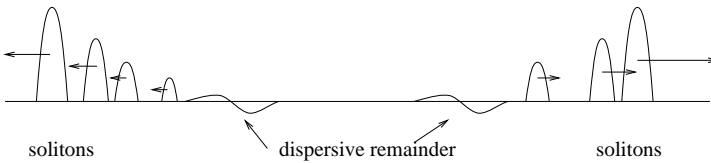


Fig. 2. The long-time behavior for the water-wave problem in the long-wave limit for $\mu < \frac{1}{3}$.

Remark 2. In what follows we prove Theorem 1 for all $0 < \mu, \mu \neq \frac{1}{3}$. However, the constants C_2 and ε_0 in the theorem depend on μ and in particular the estimates are not uniform in μ as ε tends to zero. Thus, while the statement of the theorem is true for $\mu = 0$ due to the results of [SW00a], we cannot recover our previous results from those proved below. The technical reason for this difficulty is that the surface tension enters as a singular perturbation of the $\mu = 0$ problem, and this requires very significant modifications of the former proof. Thus, while the general approach here parallels that of [SW00a], we must develop a completely new existence theory for solutions of (1) and (2) in Section 2, since that of [SW00a] no longer applies.

If $\mu = \frac{1}{3}$, the coefficient of the third order dispersive terms in (3) vanish. These terms in the KdV equation represent the leading order dispersive effects in the water-waves problem and their disappearance means that in this parameter regime water waves are almost dispersionless. This fact has been recognized for a long time by the fluid-mechanics community and has been experimentally exploited in an interesting way. Sound waves in air are also dispersionless, but in contrast to water waves, are hard to observe visually. By studying surface-tension-driven water waves, in water with $\mu = \frac{1}{3}$ (which corresponds to a water depth of about 5 mm at 20° C) it has been possible to exhibit visually many of the theoretically predicted properties of two-dimensional sound waves [Li78, Sections 1.8 and 3.4].

Furthermore, the point $\mu = \frac{1}{3}$ is a codimension-two bifurcation point for the stationary water-wave problem and this fact has been exploited in a series of recent works ([BGT96, BG99]) to construct a host of interesting traveling water waves. The waves constructed in those works require very special choices of initial conditions

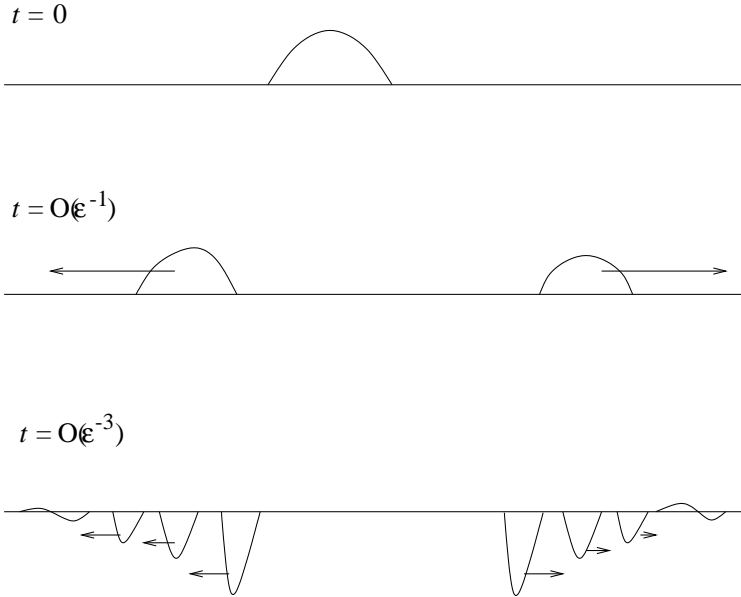


Fig. 3. The long-time behavior for the water-wave problem in the long-wave limit for $\mu > \frac{1}{3}$.

and our goal here is to derive equations that describe the evolution of general long-wavelength initial conditions. If we try to approximate such solutions using Theorem 1, the coefficients of the third derivative terms in (3) are equal to zero, and the resulting equations are just Burgers equation whose solutions typically form shocks in finite time. Thus, if we wish to model interesting behavior in the water-wave problem for this parameter regime, we must modify the scaling and the approximating equations (3). The next theorem explains how this can be done.

Theorem 2 (The degenerate case: The Kawahara limit). *Let $\mu = \frac{1}{3} + \nu\epsilon^2$. 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 Kawahara equation.*

More precisely: Fix $s \geq 6$; then, for all $C_1, T_0 > 0$, there exist $C_2, \epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$ the following is true.

Let $X_1|_{t=0} = 0, X_2|_{t=0}(\alpha) = \epsilon^4 \Phi_1(\epsilon\alpha)$ and $\partial_t X_1|_{t=0}(\alpha) = \epsilon^4 \Phi_2(\epsilon\alpha)$ with

$$\|(\Phi_1, \Phi_2)\|_{H^{s+8}(2) \cap H^{s+17}(0)} \leq C_1.$$

Define amplitudes $A_1 = A_1(\underline{\alpha}, T)$ and $A_2 = A_2(\underline{\alpha}, T)$ satisfying the Kawahara equations

$$\begin{aligned} 2\partial_T A_1 &= \nu \partial_{\underline{\alpha}}^3 A_1 - \frac{2}{90} \partial_{\underline{\alpha}}^5 A_1 - \frac{3}{2} A_1 \partial_{\underline{\alpha}} A_1 \quad \text{and} \\ 2\partial_T A_2 &= -\nu \partial_{\underline{\alpha}}^3 A_2 + \frac{2}{90} \partial_{\underline{\alpha}}^5 A_2 + \frac{3}{2} A_2 \partial_{\underline{\alpha}} A_2, \end{aligned} \tag{4}$$

with $2A_1|_{T=0} = \Phi_2 + \Phi_1$ and $2A_2|_{T=0} = \Phi_2 - \Phi_1$. Denote by $[0, T_1]$ the existence interval of the solutions A_1 and A_2 in $H^{s+8}(2)$ and set $T_2 = \min(T_0, T_1)$.

Then there exists a reparametrization of the fluid surface (denoted with the same symbols) such that the solution of the water wave-problem (1) and (2) satisfies

$$\sup_{t \in [0, T_2/\varepsilon^5]} |X_1(0, t)| \leq C_2 \varepsilon^{1/2}$$

and

$$\sup_{t \in [0, T_2/\varepsilon^5]} \left\| \begin{pmatrix} \partial_\alpha X_1 \\ X_2 \\ \partial_t X_1 \end{pmatrix} - \psi_{\text{LWL}}(\varepsilon, \cdot, t) \right\|_{H^{s-1} \times H^s \times H^{s-3/2}} \leq C_2 \varepsilon^{11/2},$$

where

$$\psi_{\text{LWL}}(\varepsilon, \alpha, t) = \varepsilon^4 A_1(\varepsilon(\alpha - t), \varepsilon^5 t) \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} + \varepsilon^4 A_2(\varepsilon(\alpha + t), \varepsilon^5 t) \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Proof. See Section 4. \square

Remark 3. The local existence and uniqueness of solutions of the Kawahara equations in $H^{8+s}(2)$ is guaranteed for initial conditions in $H^{8+s}(2) \cap H^{16+s}(0)$. However, we are not aware of any studies of the global existence of solutions to (4), in contrast to the case of the KdV equation. The work of Il'ichev [Il90] shows that the Kawahara equations (4) possess stable solitary waves at least when $\nu = 1$. This gives global existence for solutions with initial conditions close to the solitary waves.

As remarked above, for water at room temperature, the value $\mu = \frac{1}{3}$ corresponds to a depth of about 5 mm. Because dissipation is no longer negligible in such thin layers, it is difficult to observe Kawahara dynamics in such shallow fluids. However, as noted by Marchenko [Ma88], a thin layer of ice floating on the top of the water can lead to effects similar to those that would arise from a greatly enhanced surface tension. Therefore, the evolution of waves on fluids of moderate depth, under a thin layer of ice are also expected to be described by the Kawahara equations at wavelengths where dissipation should be negligible.

Corollary 2. Assume that Theorem 2 holds. Then, for the Eulerian variables $u_1 = u_1(x, t)$ and $\eta = \eta(x, t)$ defined by $u_1(\tilde{X}_1(\alpha, t), t) = \partial_t X_1(\alpha, t)$ and $\eta(\tilde{X}_1(\alpha, t), t) = X_2(\alpha, t)$, the following estimate holds:

$$\sup_{t \in [0, T_0/\varepsilon^5]} \left\| \begin{pmatrix} \eta \\ u_1 \end{pmatrix} - \varepsilon^4 A_1(\varepsilon(\cdot - t), \varepsilon^5 t) \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \varepsilon^4 A_2(\varepsilon(\cdot + t), \varepsilon^5 t) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|_{C_b^{s-3/2}} \leq C_2 \varepsilon^{9/2}.$$

Proof. This follows exactly as [SW00a, Corollary 1.5]. \square

The plan of the paper is as follows. In Section 2 we discuss the local existence and uniqueness theory for (1) and (2). In Section 3 we consider the formal long-wave limit and derive the KdV and Kawahara equations. Finally, in Section 4 we prove the error estimates given in Theorems 1 and 2.

Before we start let us make some remarks.

Remark 4. The main difficulty in proving these results stems from the long time scales, $\mathcal{O}(1/\varepsilon^3)$ in the KdV case and $\mathcal{O}(1/\varepsilon^5)$ in the Kawahara case, which have to be considered. Since the amplitudes of the approximations are of order $\mathcal{O}(\varepsilon^2)$ in the KdV case and $\mathcal{O}(\varepsilon^4)$ in the Kawahara case a simple application of Gronwall’s inequality would only give estimates on a “short” time scale $\mathcal{O}(1/\varepsilon^2)$ in the KdV case and $\mathcal{O}(1/\varepsilon^4)$ in the Kawahara case. To close the gap we must exploit some additional structure of the water-wave problem. With the introduction of a suitable energy norm and some normal form transformations we are able to take advantage of the gradient-like structure of the nonlinearity, allowing us to extend the estimates to the desired time scale.

Remark 5. For a system

$$\partial_t \mathcal{W} = \Lambda \mathcal{W} + B(\mathcal{W}, \mathcal{W}) \quad (\Lambda \text{ linear, } B \text{ bilinear})$$

and an approximation $\varepsilon^2 \psi$, the error $\varepsilon^\beta R = W - \varepsilon^2 \psi$ for a $\beta > 3$ satisfies

$$\partial_t R = \Lambda R + 2\varepsilon^2 B(\psi, R) + \varepsilon^\beta B(R, R) + \varepsilon^{-\beta} \text{Res}(\varepsilon^2 \psi),$$

where the formal error $\text{Res}(\varepsilon^2 \psi)$ is small by construction. In order to estimate R an energy functional $E_s(\cdot, \cdot)$ has to be constructed which simultaneously yields the local existence and uniqueness of the solutions of the quasilinear system and allows us to follow the solution for long enough to obtain the desired error estimates, i.e.,

$$\begin{aligned} \frac{1}{2} \partial_t E_s(R, R) &= E_s(R, \Lambda R) + 2\varepsilon^2 E_s(R, B(\psi, R)) + \text{h.o.t} \\ &\leq 0 + C\varepsilon^3 E_s(R, R) + \text{h.o.t} \end{aligned}$$

Then by Gronwall’s inequality the $\mathcal{O}(1)$ boundedness of $E_s(R, R)$ and so also of R on an $\mathcal{O}(1/\varepsilon^3)$ time scale follows.

Remark 6. The cases of zero and nonzero surface tension must be distinguished since the local existence and uniqueness theory for the solutions is different and the energy $E_s(\cdot, \cdot)$ described in Remark 5 is (somewhat surprisingly) quite different in the two cases.

Remark 7. The theory of modulation equations is a weakly nonlinear theory, where the dispersion relation

$$\lambda^2 = -(k + \mu k^3) \tanh k$$

for the eigenvalues λ of the linearization of (1) and (2) plays a crucial role. In the long-wave limit the Fourier modes are concentrated in an $\mathcal{O}(\varepsilon)$ neighborhood of the Fourier wave number $k = 0$. Blowing up this neighborhood yields

$$\lambda_{\pm}(\varepsilon K) = \pm \varepsilon K \pm \varepsilon^3 \frac{3\mu - 1}{6} K^3 \pm \varepsilon^5 \frac{2 - 5\mu}{30} K^5 + \mathcal{O}(\varepsilon^7 K^7).$$

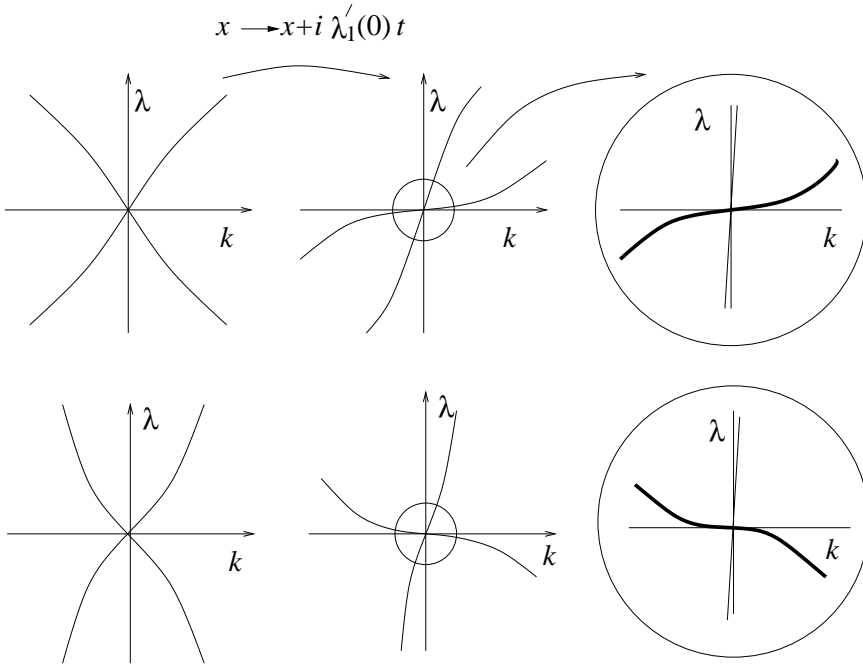


Fig. 4. Expansion of the eigenvalues for $\mu < \frac{1}{3}$ and $\mu > \frac{1}{3}$.

The $\mathcal{O}(\varepsilon)$ terms correspond to transport to the right and to the left, whereas the terms of order $\mathcal{O}(\varepsilon^3)$ contribute to the linear part of the KdV equations. This expansion shows why the case $\mu = \frac{1}{3}$ is exceptional and higher order terms have to be included.

Remark 8. In the local existence and uniqueness theory we use in an essential way the fact that $R + S$ is a gradient, in detail

$$R + S = \partial_\alpha(Q(\partial_\alpha X)^{-3}(-(\partial_\alpha X_2)\partial_\alpha^2 X_1 + (1 + \partial_\alpha X_1)\partial_\alpha^2 X_2)).$$

Remark 9. We refer to [SW00a] for a discussion of the mathematical literature about the approximation by the KdV equation [Cr85, KN86, Kal89]. Other models for the long wave limit are described in [Bo1877, Lo64, BS71, BBM72, Me79, Ma91]. A mathematical description of the long-wave limit by the KdV or the Kawahara equation in various other problems can be found in [Sch98, SW00b, SW00c]. In contrast to other models the KdV or Kawahara description has the advantage of being simple, exactly solvable in the KdV case, and independent of ε .

Notation. 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 $[L, M] = LM - ML$.

2. Formulation as evolutionary problem

In this section we prove local existence and uniqueness for the solutions of (1) and (2). Due to Remark 5 our local existence and uniqueness approach differs from the existing literature. We obtain the result by embedding (1) and (2) into a bigger system which is quasilinear in the sense of [Kat75], such that standard existence and uniqueness theory can be applied. In order to do so we construct out of (1) and (2) a nonlinear wave equation.

Notation. We introduce $U_1 = \partial_t X_1$, and $V_1 = \partial_t U_1$. All variables are collected in the vector $\mathcal{V} = (X_1, X_2, U_1, V_1)$. Later on we need the additional variable $Z_1 = \mathcal{K}_0 X_1$, where \mathcal{K}_0 is the linear part of the operator $\mathcal{K}(X)$ and has Fourier symbol $\hat{\mathcal{K}}_0(k) = -i \tanh(k)$, and the collections of variables $\mathcal{W} = (Z_1, X_2, U_1)$ and $\mathcal{W}_e = (Z_1, X_2, U_1, V_1)$. The vectors \mathcal{W} and \mathcal{W}_e will be in the spaces $\mathcal{H}^s = H^s \times H^s \times H^{s-3/2}$ and $\mathcal{H}_e^s = H^s \times H^s \times H^{s-3/2} \times H^{s-3}$ respectively. We abuse notation and do not distinguish between operators which depend on \mathcal{V} or \mathcal{W} ; for instance, we write for $\mathcal{K}(X)$ also $\mathcal{K}(\mathcal{V})$ and $\mathcal{K}(\mathcal{W})$.

Preliminaries. Before we start to construct the quasilinear system out of (1) and (2) we recall some facts. For more details we refer to [SW00a] and the literature cited there. The operator $\mathcal{K}(X)$ is of the form $\mathcal{K}(X) = \mathcal{K}_0 + \mathcal{S}_1(X)$, where the operator $\mathcal{S}_1(X)$ has certain smoothing properties. As a rule, a term with $\mathcal{S}_1(X)$ in front will have the regularity of X . More precisely, in terms of the variables \mathcal{W} used later on:

Lemma 1. *Fix $s \geq 11/2$. Then there exist $C > 0$ and $C_s > 0$ such that for $\|(Z_1, X_2)\|_{H^s \times H^s} \leq C_s$ the operator $\mathcal{S}_1(\mathcal{W}) = \mathcal{K}(\mathcal{W}) - \mathcal{K}_0$ satisfies*

$$\|\mathcal{S}_1(\mathcal{W})U\|_{H^s} \leq C \|\mathcal{W}\|_{\mathcal{H}^s} \|U\|_{H^3}.$$

Proof. See [SW00a, Lemma 3.14]. \square

Lemma 2. *Assume the situation of Lemma 1. Then, for all $s \geq 6$,*

$$\begin{aligned} \|\partial_t(\mathcal{S}_1(\mathcal{W})U_1)\|_{H^{s-3}} &\leq C \|\mathcal{W}\|_{\mathcal{H}^s} (\|U_1\|_{H^3} + \|V_1\|_{H^3}), \\ \|[\partial_t^2, \mathcal{S}_1(\mathcal{W})]U_1\|_{H^{s-3}} &\leq C \|\mathcal{W}_e\|_{\mathcal{H}_e^s} (\|U_1\|_{H^3} + \|V_1\|_{H^3}), \\ \|\partial_\alpha(\mathcal{S}_1(\mathcal{W})U_1)\|_{H^{s-1}} &\leq C \|\mathcal{W}\|_{\mathcal{H}^s} \|U_1\|_{H^3}. \end{aligned}$$

Proof. See [SW00a, Lemma 3.15]. \square

In the construction of the quasilinear system commutators $[a, \mathcal{K}_0]$ play a big role. As a rule, $[a, \mathcal{K}_0]u$ smoothes u and has the regularity of the function a .

Lemma 3. *Let $r \geq 0$, $q > \frac{1}{2}$ 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. See [SW00a, Lemma 3.12]. \square

As mentioned in the beginning of this section, it will be convenient to introduce a new variable $Z_1 = \mathcal{K}_0 X_1$. We associate with Z_1 the operator

$$\mathcal{M}_1(Z_1, \cdot) = [X_1, \mathcal{K}_0] \cdot$$

which satisfies

Lemma 4. *Let $r \geq 0, q > \frac{1}{2}$ 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. See [SW00a, Corollary 3.13]. \square

Remark 10. The operator \mathcal{M}_1 is well defined, even though \mathcal{K}_0 is not invertible in general, due to the commutator in its definition.

Remark 11. Examining the expression for $\|[a, \mathcal{K}_0]u\|_{H^r}$ on [SW00a, p. 1499] (which is expressed in terms of the Fourier transforms of a and u), we also see that it can be bounded by

$$\|[a, \mathcal{K}_0]u\|_{H^r} \leq C \min(\|\hat{a}\|_{L^1(r+p)} \|u\|_{H^{q-p}}, \|a\|_{H^{r+p}} \|\hat{u}\|_{L^1(q-p)}),$$

where $\|\hat{f}\|_{L^1(n)} = \int (1 + |k|^2)^{(n/2)} |\hat{f}(k)| dk$. There is then a similar bound for $\|\mathcal{M}_1(a, u)\|_{H^r}$, namely

$$\|\mathcal{M}_1(a, u)\|_{H^r} \leq C \min(\|\hat{a}\|_{L^1(r+p)} \|u\|_{H^{q-p}}, \|a\|_{H^{r+p}} \|\hat{u}\|_{L^1(q-p)}).$$

Estimates of this kind were also used in [SW00a]: see, for example, (6.8). These alternate forms of Lemmas 3 and 4 will sometimes be used in Section 4 since they do not suffer from the loss of half a power of ε described in Remark 1, i.e., if $f(x) = \varepsilon^2 F(\varepsilon x)$, then $\|\hat{f}\|_{L^1(n)} \leq \varepsilon^2 \|\hat{F}\|_{L(n)}$ whereas $\|f\|_{H^n} \leq \varepsilon^{3/2} \|F\|_{H^n}$.

In order to express the term $\partial_\alpha X_1$ in terms of Z_1 we define additionally the operator

$$\mathcal{M}_{2\cdot} = -\partial_\alpha (\mathcal{K}_0)^{-1}.$$

which is a map from H^{s+1} to H^s .

Finally, the operator $(1 + \mathcal{K}_0^2) \cdot$ is infinitely smoothing due to the fact that in Fourier space its symbol $(1 + \hat{\mathcal{K}}_0(k)^2)$ vanishes with some exponential rate for $|k| \rightarrow \infty$.

Remark 12. As a general rule, it is possible to obtain existence and uniqueness of solutions of the nonlinear wave equation

$$\partial_t^2 u = a(u)L(\partial_\alpha)u + b(u), \quad (a > 0) \tag{5}$$

with L an elliptic operator, provided the coefficients $a(u)$ and $b(u)$ contain not more than half the number of derivatives of L and terms with $\partial_t u$. In the water-wave problem (1) and (2) the operator L is essentially given by the multiplier $\hat{L}(k) = -(k + \mu k^3) \tanh k$ in Fourier space. As a consequence $|\hat{L}(k)| \sim |k|^3$ for $|k| \rightarrow \infty$. Thus, if we look for solutions u in $H^{s-3/2}$ we can allow for coefficients $a, b \in H^{s-3}$, i.e., the coefficients a and b can lose $\frac{3}{2}$ derivatives. Moreover, we then have $\partial_t u \in H^{s-3}$.

The nonlinear wave equation will be constructed for the variable U_1 . We have to choose $X_{1,2} \in H^s$, $U_1 = \partial_t X_1 \in H^{s-3/2}$, and $V_1 = \partial_t^2 X_1 \in H^{s-3}$. So all terms up to $\partial_\alpha^3 X_{1,2} \in H^{s-3}$ and $\partial_\alpha U \in H^{s-5/2}$ have the right regularity to be included in the functions a and b . Terms containing $\partial_\alpha^2 U \in H^{s-7/2}$ create difficulties and must be included in the operator L .

The quasilinearization. With this scheme in mind we begin with the construction of the quasilinear system. With the notation $U_2 = \partial_t X_2$ the equations (1) and (2) become

$$(1 + \partial_\alpha X_1) \partial_t U_1 + \partial_\alpha X_2 (1 + \partial_t U_2) = \mu R(\partial_\alpha X, \partial_\alpha^2 X) + \mu S(\partial_\alpha X, \partial_\alpha^3 X),$$

$$U_2 = \mathcal{K}_0 U_1 + \mathcal{S}_1(X) U_1.$$

Differentiation of (1) with respect to time yields

$$(1 + \partial_\alpha X_1) \partial_t^2 U_1 + (\partial_t U_1) \partial_\alpha U_1 + (1 + \partial_t U_2) \partial_\alpha U_2 + (\partial_\alpha X_2) \partial_t^2 U_2$$

$$= \mu \partial_1 R(\partial_\alpha X, \partial_\alpha^2 X) \partial_\alpha U + \mu \partial_2 R(\partial_\alpha X, \partial_\alpha^2 X) \partial_\alpha^2 U$$

$$+ \mu \partial_1 S(\partial_\alpha X, \partial_\alpha^3 X) \partial_\alpha U + \mu \partial_2 S(\partial_\alpha X, \partial_\alpha^3 X) \partial_\alpha^3 U. \quad (6)$$

Differentiation of (2) with respect to time and space yields

$$\partial_\alpha U_2 = \mathcal{K}_0 \partial_\alpha U_1 + \partial_\alpha (\mathcal{S}_1(X) U_1),$$

$$\partial_\alpha^2 U_2 = \mathcal{K}_0 \partial_\alpha^2 U_1 + \partial_\alpha^2 (\mathcal{S}_1(X) U_1),$$

$$\partial_\alpha^3 U_2 = \mathcal{K}_0 \partial_\alpha^3 U_1 + \partial_\alpha^3 (\mathcal{S}_1(X) U_1),$$

$$\partial_t U_2 = \mathcal{K}_0 \partial_t U_1 + \partial_t (\mathcal{S}_1(X) U_1),$$

$$\partial_t^2 U_2 = \mathcal{K}_0 \partial_t^2 U_1 + \mathcal{S}_1(X) \partial_t^2 U_1 + [\partial_t^2, \mathcal{S}_1(X)] U_1.$$

Replacing $\partial_\alpha U_2$, $\partial_\alpha^2 U_2$, $\partial_\alpha^3 U_2$, and $\partial_t^2 U_2$ in (6) gives

$$(1 + \partial_\alpha X_1) \partial_t^2 U_1 + \partial_t U_1 \partial_\alpha U_1 + (1 + \partial_t U_2) (\mathcal{K}_0 \partial_\alpha U_1 + \partial_\alpha (\mathcal{S}_1(X) U_1))$$

$$+ (\partial_\alpha X_2) (\mathcal{K}_0 \partial_t^2 U_1 + \mathcal{S}_1(X) \partial_t^2 U_1 + [\partial_t^2, \mathcal{S}_1(X)] U_1)$$

$$= \mu \partial_1 R(\partial_\alpha X, \partial_\alpha^2 X) \partial_\alpha U + \mu \partial_1 S(\partial_\alpha X, \partial_\alpha^3 X) \partial_\alpha U$$

$$- 3\mu Q(\partial_\alpha X)^{-5} ((1 + \partial_\alpha X_1) \partial_\alpha^2 U_1 + \partial_\alpha X_2 (\mathcal{K}_0 \partial_\alpha^2 U_1 + \partial_\alpha^2 (\mathcal{S}_1(X) U_1)))$$

$$\quad \times (-\partial_\alpha X_2 \partial_\alpha^2 X_1 + (1 + \partial_\alpha X_1) \partial_\alpha^2 X_2))$$

$$- 3\mu Q(\partial_\alpha X)^{-5} ((1 + \partial_\alpha X_1) \partial_\alpha^2 X_1 + \partial_\alpha X_2 \partial_\alpha^2 X_2)$$

$$\quad \times (-\partial_\alpha X_2 \partial_\alpha^2 U_1 + (1 + \partial_\alpha X_1) (\mathcal{K}_0 \partial_\alpha^2 U_1 + \partial_\alpha^2 (\mathcal{S}_1(X) U_1))))$$

$$+ \mu Q(\partial_\alpha X)^{-3} (-\partial_\alpha X_2 \partial_\alpha^3 U_1 + (1 + \partial_\alpha X_1) (\mathcal{K}_0 \partial_\alpha^3 U_1 + \partial_\alpha^3 (\mathcal{S}_1(X) U_1))),$$

and so

$$(f_1 + f_2 (\mathcal{K}_0 + \mathcal{S}_1(X))) \partial_t^2 U_1 + h_0 (-f_2 + f_1 \mathcal{K}_0) \partial_\alpha^3 U_1$$

$$+ 3\mu Q^{-5} ((-f_2 \partial_\alpha f_1 + f_1 \partial_\alpha f_2) (f_1 + f_2 \mathcal{K}_0) \partial_\alpha^2 U_1)$$

$$+ 3\mu Q^{-5} ((f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2) (-f_2 + f_1 \mathcal{K}_0) \partial_\alpha^2 U_1) + G_1 = 0, \quad (7)$$

with

$$\begin{aligned}
 f_1 &= (1 + \partial_\alpha X_1), \\
 f_2 &= \partial_\alpha X_2, \\
 h_0 &= -\mu Q(\partial_\alpha X)^{-3} = -\mu(f_1^2 + f_2^2)^{-3/2}, \\
 G_1 &= \partial_t U_1(1 + \mathcal{K}_0^2)\partial_\alpha U_1 + (1 + \partial_t U_2)\partial_\alpha(\mathcal{S}_1(X)U_1), \\
 &\quad + (\partial_\alpha X_2)([\partial_t^2, \mathcal{S}_1(X)]U_1) + ((1 + \partial_t U_2) - \partial_t U_1 \mathcal{K}_0)\mathcal{K}_0 \partial_\alpha U_1 \\
 &\quad - \mu \partial_1 R(\partial_\alpha X, \partial_\alpha^2 X)\partial_\alpha U - \mu \partial_1 S(\partial_\alpha X, \partial_\alpha^3 X)\partial_\alpha U \\
 &\quad - \mu Q(\partial_\alpha X)^{-3}(f_1 \partial_\alpha^3(\mathcal{S}(X)U_1)) \\
 &\quad + 3\mu Q^{-5}((-f_2 \partial_\alpha f_1 + f_1 \partial_\alpha f_2)f_2 \partial_\alpha^2(\mathcal{S}(X)U_1)) \\
 &\quad + 3\mu Q^{-5}((f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)f_1 \partial_\alpha^2(\mathcal{S}(X)U_1)).
 \end{aligned}$$

The terms are separated according to the following rules: G_1 contains the obviously semilinear terms, and the terms written explicitly are the quasilinear ones.

The quasilinear term $(-f_2 + f_1 \mathcal{K}_0)\partial_\alpha^3 U_1$ is rewritten as

$$(-f_2 + f_1 \mathcal{K}_0)\partial_\alpha^3 U_1 = (f_1 + f_2 \mathcal{K}_0)\mathcal{K}_0 \partial_\alpha^3 U_1 - f_2(1 + \mathcal{K}_0^2)\partial_\alpha^3 U_1.$$

Since $1 + \mathcal{K}_0^2$ is infinitely smoothing, the last term can be handled as a semilinear one. In a similar fashion we obtain

$$\begin{aligned}
 &(f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)(-f_2 + f_1 \mathcal{K}_0)\partial_\alpha^2 U_1 \\
 &= (f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)(f_1 + f_2 \mathcal{K}_0)\mathcal{K}_0 \partial_\alpha^2 U_1 \\
 &\quad - (f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)f_2(1 + \mathcal{K}_0^2)\partial_\alpha^2 U_1.
 \end{aligned}$$

Therefore, we arrive at

$$\begin{aligned}
 &(f_1 + f_2(\mathcal{K}_0 + \mathcal{S}_1(X)))\partial_t^2 U_1 + h_0(f_1 + f_2 \mathcal{K}_0)\mathcal{K}_0 \partial_\alpha^3 U_1 \\
 &\quad + 3\mu Q^{-5}((-f_2 \partial_\alpha f_1 + f_1 \partial_\alpha f_2)(f_1 + f_2 \mathcal{K}_0)\partial_\alpha^2 U_1) \\
 &\quad + 3\mu Q^{-5}((f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)(f_1 + f_2 \mathcal{K}_0)\mathcal{K}_0 \partial_\alpha^2 U_1) + G_2 = 0,
 \end{aligned} \tag{8}$$

with

$$\begin{aligned}
 G_2 &= G_1 - h_0 f_2(1 + \mathcal{K}_0^2)\partial_\alpha^3 U_1 \\
 &\quad + 3\mu Q^{-5}(f_1 \partial_\alpha f_1 + f_2 \partial_\alpha f_2)f_2(1 + \mathcal{K}_0^2)\partial_\alpha^2 U_1.
 \end{aligned}$$

We multiply (8) with $(f_1 - f_2 \mathcal{K}_0)$ and find that the first term in the equation becomes

$$\begin{aligned}
 &(f_1 - f_2 \mathcal{K}_0)(f_1 + f_2(\mathcal{K}_0 + \mathcal{S}_1(X)))\partial_t^2 U_1 \\
 &= (f_1)^2 \partial_t^2 U_1 - f_2 \mathcal{K}_0(f_1 \partial_t^2 U_1) + (f_1 - f_2 \mathcal{K}_0)f_2 \mathcal{S}_1(X)\partial_t^2 U_1 \\
 &\quad + f_1 f_2 \mathcal{K}_0 \partial_t^2 U_1 - f_2 \mathcal{K}_0(f_2 \mathcal{K}_0 \partial_t^2 U_1) \\
 &= (f_1)^2 \partial_t^2 U_1 - f_2[\mathcal{K}_0, f_1]\partial_t^2 U_1 + (f_1 - f_2 \mathcal{K}_0)f_2 \mathcal{S}_1(X)\partial_t^2 U_1 \\
 &\quad + f_2 \mathcal{K}_0[\mathcal{K}_0, f_2]\partial_t^2 U_1 - f_2(1 + \mathcal{K}_0^2)(f_2 \partial_t^2 U_1) + f_2^2 \partial_t^2 U_1.
 \end{aligned}$$

The second term can be rewritten as:

$$\begin{aligned}
& (f_1 - f_2\mathcal{K}_0)h_0(f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^3 U_1 \\
&= [(f_1 - f_2\mathcal{K}_0), h_0](f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^3 U_1 \\
&\quad + h_0(f_1 - f_2\mathcal{K}_0)(f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^3 U_1 \\
&= [(f_1 - f_2\mathcal{K}_0), h_0](f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^3 U_1 \\
&\quad + h_0(f_1)^2\mathcal{K}_0\partial_\alpha^3 U_1 - h_0f_2[\mathcal{K}_0, f_1]\mathcal{K}_0\partial_\alpha^3 U_1 \\
&\quad + h_0f_2\mathcal{K}_0[\mathcal{K}_0, f_2]\mathcal{K}_0\partial_\alpha^3 U_1 - h_0f_2(1 + \mathcal{K}_0^2)(f_2\mathcal{K}_0\partial_\alpha^3 U_1) \\
&\quad + h_0f_2^2\mathcal{K}_0\partial_\alpha^3 U_1.
\end{aligned}$$

For the third and fourth term, we obtain in the same fashion, with

$$\begin{aligned}
h_1 &= 3\mu Q^{-5}, \\
a_1 &= (-f_2\partial_\alpha f_1 + f_1\partial_\alpha f_2), \\
a_2 &= f_1\partial_\alpha f_1 + f_2\partial_\alpha f_2
\end{aligned}$$

the value

$$\begin{aligned}
& (f_1 - f_2\mathcal{K}_0)h_1(a_1((f_1 + f_2\mathcal{K}_0)\partial_\alpha^2 U_1 + a_2(f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^2 U_1)) \\
&= [(f_1 - f_2\mathcal{K}_0), h_1a_1](f_1 + f_2\mathcal{K}_0)\partial_\alpha^2 U_1 \\
&\quad + [(f_1 - f_2\mathcal{K}_0), h_1a_2](f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^2 U_1 \\
&\quad + h_1a_1(f_1^2 + f_2^2)\partial_\alpha^2 U_1 + h_1a_2(f_1^2 + f_2^2)\mathcal{K}_0\partial_\alpha^2 U_1 \\
&\quad - h_1a_1f_2[\mathcal{K}_0, f_1]\partial_\alpha^2 U_1 - h_1a_2f_2[\mathcal{K}_0, f_1]\mathcal{K}_0\partial_\alpha^2 U_1 \\
&\quad + h_1a_1f_2\mathcal{K}_0[\mathcal{K}_0, f_2]\partial_\alpha^2 U_1 + h_1a_2f_2\mathcal{K}_0[\mathcal{K}_0, f_2]\mathcal{K}_0\partial_\alpha^2 U_1 \\
&\quad - h_1a_1f_2(1 + \mathcal{K}_0^2)(f_2\partial_\alpha^2 U_1) - h_1a_2f_2(1 + \mathcal{K}_0^2)(f_2\mathcal{K}_0\partial_\alpha^2 U_1).
\end{aligned}$$

Therefore, (7) transforms into

$$\begin{aligned}
& (f_3 - H_1)\partial_t^2 U_1 + h_0f_3\mathcal{K}_0\partial_\alpha^3 U_1 \\
&\quad + h_1a_1f_3\partial_\alpha^2 U_1 + h_1a_2f_3\mathcal{K}_0\partial_\alpha^2 U_1 + G_3 = 0 \tag{9}
\end{aligned}$$

with

$$\begin{aligned}
f_3 &= f_1^2 + f_2^2, \\
H_1 &= f_2[\mathcal{K}_0, f_1] \cdot - f_2\mathcal{K}_0[\mathcal{K}_0, f_2] \cdot + f_2(1 + \mathcal{K}_0^2)(f_2 \cdot) \\
&\quad - (f_1 - f_2\mathcal{K}_0)f_2\mathcal{S}_1(X) \cdot, \\
H_2 &= f_2[\mathcal{K}_0, f_1] \cdot - f_2\mathcal{K}_0[\mathcal{K}_0, f_2] \cdot + f_2(1 + \mathcal{K}_0^2)(f_2 \cdot), \\
G_3 &= (f_1 - f_2\mathcal{K}_0)G_2 + [(f_1 - f_2\mathcal{K}_0), h_0](f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^3 U_1 \\
&\quad + [(f_1 - f_2\mathcal{K}_0), h_1a_1](f_1 + f_2\mathcal{K}_0)\partial_\alpha^2 U_1 \\
&\quad + [(f_1 - f_2\mathcal{K}_0), h_1a_2](f_1 + f_2\mathcal{K}_0)\mathcal{K}_0\partial_\alpha^2 U_1 \\
&\quad - h_0H_2\mathcal{K}_0\partial_\alpha^3 U_1 - h_1a_1H_2\partial_\alpha^2 U_1 - h_1a_2H_2\mathcal{K}_0\partial_\alpha^2 U_1.
\end{aligned}$$

We multiply this with

$$\begin{aligned} (f_3 - H_1)^{-1} &= (f_3(1 - f_3^{-1}H_1))^{-1} = (1 - f_3^{-1}H_1)^{-1} f_3^{-1} \\ &= \sum_{n=0}^{\infty} (f_3^{-1}H_1)^n f_3^{-1} \equiv f_3^{-1} + H_3 \end{aligned}$$

and obtain

$$\partial_t^2 U_1 + h_0 \mathcal{K}_0 \partial_\alpha^3 U_1 + h_1 a_1 \partial_\alpha^2 U_1 + h_1 a_2 \mathcal{K}_0 \partial_\alpha^2 U_1 + G_4 = 0 \tag{10}$$

with

$$G_4 = (f_3 - H_1)^{-1} G_3 + H_3 (h_0 f_3 \mathcal{K}_0 \partial_\alpha^3 U_1 + h_1 a_1 f_3 \partial_\alpha^2 U_1 + h_1 a_2 f_3 \mathcal{K}_0 \partial_\alpha^2 U_1).$$

By the above Lemmas $[\cdot, \mathcal{K}_0]$, $(1 + \mathcal{K}_0^2)$, $\mathcal{S}_1(X)$, and consequently H_1 , H_2 , and H_3 have smoothing properties. Hence, it is easy to see that all terms in G_4 , h_0 , h_1 , a_1 and a_2 are at least in H^{s-3} provided that $X_1, X_2 \in H^s$, $U_1 \in H^{s-3/2}$, and $V_1 \in H^{s-3}$ and have sufficiently small norm that the sum defining H_3 converges.

Finally, we take out the linear term from G_4 and introduce the abbreviations

$$\begin{aligned} \mathcal{L}(\mathcal{V})U_1 &= -\partial_\alpha (h_0 \mathcal{K}_0 \partial_\alpha^2 U_1) - h_2 \partial_\alpha^2 U_1 - \mathcal{K}_0 \partial_\alpha U_1 \\ G_5 &= -G_4 + \mathcal{K}_0 \partial_\alpha U_1, \end{aligned}$$

where we set $h_2 = h_1 a_1$ and used

$$\begin{aligned} h_1 a_2 &= 3\mu Q^{-5} \partial_\alpha (f_1^2 + f_2^2)/2 = 3\mu Q^{-5} \partial_\alpha (Q^2)/2 \\ &= 3\mu Q^{-4} \partial_\alpha Q = -\mu \partial_\alpha (Q^{-3}) = \partial_\alpha h_0. \end{aligned}$$

Remark 13. Note that $h_0 = -\mu + \text{h.o.t.}$ (higher order terms) and that it is essential for the local existence and uniqueness theory that $h_1 a_2$ can be written as the gradient $\partial_\alpha h_0$. A linear system $\partial_t^2 U_1 = \mathcal{K}_0 \partial_\alpha^3 U_1 + \delta_1 \partial_\alpha^2 U_1 + \delta_2 \mathcal{K}_0 \partial_\alpha^2 U_1$ with $\delta_{1,2} \in \mathbb{R}$ has the eigenvalues λ satisfying $\lambda^2 = -(\tanh k)k^3 - \delta_1 k^2 + i\delta_2 (\tanh k)k^2$. The real term $\delta_1 k^2$ can be dominated by $-(\tanh k)k^3$ in the sense that the eigenvalues stay on the imaginary axis for $|k|$ sufficiently big. This is no longer true for the imaginary term $i\delta_2 (\tanh k)k^2$ which leads to eigenvalues leaving the imaginary axis by order $\mathcal{O}(k^{1/2})$ for $|k| \rightarrow \infty$. This would yield to an unacceptable loss of regularity of the solutions. Because the coefficient in front of this term is a gradient, $\partial_\alpha h_0$, it can be merged into the leading order term $-\partial_\alpha (h_0 \mathcal{K}_0 \partial_\alpha^2 U_1)$ and this loss of regularity can be avoided. (See the discussion following (15) for more details.)

Rewriting (10) and the associated equations for X_1 and X_2 as a first order system we obtain

$$\begin{aligned} \partial_t X_1 &= U_1, \\ \partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{S}_1(X)U_1, \\ \partial_t U_1 &= V_1, \\ \partial_t V_1 &= \mathcal{L}(\mathcal{V})U_1 + G_5 \end{aligned} \tag{11}$$

for which the local existence and uniqueness of solutions can be established.

Theorem 3. For all $s \geq 6$ there exists a $C_1 > 0$ such that, for all $C_2 \in (0, C_1]$, there is a $T_0 > 0$ such that the following is true. For each initial condition $\mathcal{V}_0 \in \mathcal{H}_e^s$ with $\|\mathcal{V}_0\|_{\mathcal{H}_e^s} \leq C_2$ there exists a unique solution $\mathcal{V} \in C([0, T_0], \mathcal{H}_e^s)$ of (11) with $\mathcal{V}|_{t=0} = \mathcal{V}_0$.

Remark 14. The existence proofs are based on the choice of a suitable energy $E_s(\mathcal{V}, \mathcal{V})$ which is equivalent to the usual \mathcal{H}_e^s scalar product. 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 (11) cancel and we can estimate $\frac{d}{dt} E_s(\mathcal{V}, \mathcal{V})$ 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, which is the form of the existence theorem that we actually use.

Remark 15. Various formulations of (1) and (2) as an evolutionary problem and associated local existence and uniqueness theorems have been obtained for the water-wave problem with and without surface tension for finite and infinite depth in two dimensions ([Na74, Sh76, Yo82, Yo83, Wu97]) and in three dimensions ([Wu99]). For a more complete historical discussion of the existence theory for the water-wave problem, see [SW00a, Remark 4.4].

Not all initial conditions \mathcal{V}_0 of (11) lead to solutions of the water-wave problem (1) and (2), 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 1. We define

$$\begin{aligned} \mathcal{C}_{p,X} = \{ \mathcal{V} = (\phi_0, \phi_1, \phi_2, \phi_3) \mid \\ \text{(a) } & (1 + \partial_\alpha \phi_0)\phi_3 + (\partial_\alpha \phi_1)(1 + \phi_4) \\ & = \mu R(\partial_\alpha(\phi_0, \phi_1), \partial_\alpha^2(\phi_0, \phi_1)) + \mu S(\partial_\alpha(\phi_0, \phi_1), \partial_\alpha^3(\phi_0, \phi_1)), \\ \text{(b) } & \phi_4 = \mathcal{K}(\phi_0, \phi_1)\phi_3 + [\partial_t, \mathcal{K}(\phi_0, \phi_1)]\phi_2 \}. \end{aligned}$$

We know that $[\partial_t, \mathcal{K}(\phi_0, \phi_1)]$ is a function of ϕ_0, ϕ_1 and ϕ_2 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 that $\mathcal{V}(t) \in \mathcal{C}_{p,X}$ for all $t > 0$ due to the construction of (11).

As explained in [SW00a] for our approximation problem we must avoid trying to estimate the variable X_1 in some Sobolev space. So as in [SW00a] the variable X_1 will now be replaced by the variable $Z_1 = \mathcal{K}_0 X_1$ which is possible due to [SW00a, Section 3, especially Lemma 3.14].

Thus, we finally consider

$$\begin{aligned} \partial_t Z_1 &= \mathcal{K}_0 U_1, \\ \partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{S}_1(\mathcal{W})U_1, \\ \partial_t U_1 &= V_1, \\ \partial_t V_1 &= \mathcal{L}(\mathcal{W})U_1 + G_5 \end{aligned} \tag{12}$$

for $\mathcal{W}_e = (Z_1, X_2, U_1, V_1)$.

For this system we now prove the local existence and uniqueness of solutions.

Theorem 4. *For all $s \geq 6$ there exists a $C_1 > 0$ such that, for all $C_2 \in (0, C_1]$, there is a $T_0 > 0$ such that the following is true. For each initial condition $\mathcal{W}_{e,0} \in \mathcal{H}_e^s$ with $\|\mathcal{W}_{e,0}\|_{\mathcal{H}_e^s} \leq C_2$ there exists a unique solution $\mathcal{W}_e \in C([0, T_0], \mathcal{H}_e^s)$ of (12) with $\mathcal{W}_e|_{t=0} = \mathcal{W}_{e,0}$.*

Proof. The proof of Theorem 4 is based on estimating the time derivative of the energy function

$$E_s(\mathcal{W}, \mathcal{W}) = (Z_1, Z_1)_{H^s} + (X_2, X_2)_{H^s} + (U_1, U_1)_{H^{s-3/2}} + (V_1, V_1)_{H^{s-3}}. \tag{13}$$

We see immediately that for any $r \geq 1$, if we differentiate along solutions of (12), there exists a constant C_E such that

$$\frac{d}{dt} E_r(\mathcal{W}, \mathcal{W}) \leq C_E((Z_1, Z_1)_{H^r} + (X_2, X_2)_{H^r} + (U_1, U_1)_{H^r} + (V_1, V_1)_{H^r}). \tag{14}$$

We use (14) to control derivatives of the solutions of (12) with $r \leq s - 3$. It is not suitable for controlling the highest derivatives because it bounds time derivatives of $\|U_1\|_{H^{r-3/2}}$ and $\|V_1\|_{H^{r-3}}$ on the left-hand side in terms of $\|U_1\|_{H^r}$ and $\|V_1\|_{H^r}$ on the right-hand side.

In order to estimate the highest order derivatives we first look at the quasilinear part of the last two equations of (12) which we write again as a nonlinear wave equation

$$\partial_t^2 U_1 = -\partial_\alpha(h_0 \mathcal{K}_0 \partial_\alpha^2 U_1) - h_2 \partial_\alpha^2 U_1 - \mathcal{K}_0 \partial_\alpha U_1. \tag{15}$$

We introduce the skew symmetric operators λ_1 and λ_2 by $\lambda_1^2 = -\mathcal{K}_0 \partial_\alpha$ and $\lambda_2^2 = -\mathcal{M}_2 = \partial_\alpha \mathcal{K}_0^{-1}$. Then $u = \lambda_1 U_1$ satisfies

$$\begin{aligned} \partial_t^2 u &= \partial_\alpha \lambda_1 (h_0 \partial_\alpha \lambda_1 u) - \lambda_1 (h_2 \mathcal{M}_2 \lambda_1 u) + \lambda_1^2 u \\ &= \partial_\alpha \lambda_1 (h_0 \partial_\alpha \lambda_1 u) + \lambda_1 \lambda_2 (h_2 \lambda_1 \lambda_2 u) + \lambda_1^2 u - \lambda_1 [\lambda_2, h_2] \lambda_1 \lambda_2 u, \end{aligned}$$

where we used $\lambda_1 \lambda_2 = \lambda_2 \lambda_1$. The last term is semilinear since λ_1 and λ_2 both lose half a derivative. Exploiting the smoothing property of the commutator, we find that $[\lambda_2, h_2]$ gains half a derivative and hence the combination $\lambda_1 [\lambda_2, h_2] \lambda_1 \lambda_2$ loses a total of one derivative. By Remark 12 we can still expect well-posedness of this equation provided this term loses less than 3/2 of a derivative so we ignore this part of the equation for the moment. (We intend here only to give a sketch of the argument – precise details are supplied below.) The operator

$$\Lambda \cdot = -\partial_\alpha \lambda_1 (h_0 \partial_\alpha \lambda_1) \cdot - \lambda_1 \lambda_2 (h_2 \lambda_1 \lambda_2) \cdot - \lambda_1^2$$

is self-adjoint in L^2 . For $\|h_2\|_{C_b^1}$ and $\|h_0 + \mu\|_{C_b^2}$ sufficiently small it is also positive in L^2 and hence according to [DS71] fractional powers of Λ can be defined. For $\|\mathcal{W}_e\|_{\mathcal{H}_e^s}$ sufficiently small the norm $\|\Lambda^{s/3} \cdot\|_{L^2}$ is equivalent to the usual H^s norm.

Again, in order to emphasize the essential outline of the argument, we assume for the moment that h_0 and h_2 are independent of time. Then

$$\frac{d}{dt} \left(\left(\Lambda^{\frac{s}{3}-\frac{7}{6}} \partial_t u, \Lambda^{\frac{s}{3}-\frac{7}{6}} \partial_t u \right)_{L^2} + \left(\Lambda^{\frac{s}{3}-\frac{2}{3}} u, \Lambda^{\frac{s}{3}-\frac{2}{3}} u \right)_{L^2} \right) = 0$$

and the local and global existence for solutions $u \in C([0, T_0], H^{s-2}) \cap C^1([0, T_0], H^{s-7/2})$ of

$$\partial_t^2 u = -\Lambda u$$

follows. Here $(\cdot, \cdot)_{L^2}$ stands for the L^2 scalar product.

We now fill in the details of the above sketch. Returning to the full system we introduce, in addition to u , the variables $z = \lambda_1 Z_1$, $x = \lambda_1 X_2$, $v = \lambda_1 V_1$, $g_0 = \lambda_1 S_1(X)U_1$ and $g_5 = \lambda_1 G_5 + \lambda_1[\lambda_2, h_2]\lambda_1\lambda_2 u$.

Then the variables $\omega = (z, x, u, v)$ satisfy

$$\begin{aligned} \partial_t z &= \mathcal{K}_0 u, \\ \partial_t x &= \mathcal{K}_0 u + g_0, \\ \partial_t u &= v, \\ \partial_t v &= -\Lambda u + g_5. \end{aligned}$$

We are looking now for solutions satisfying $(\Lambda^{r+1} z, \Lambda^{r+1} x, \Lambda^{r+1/2} u, \Lambda^r v) \in (L^2)^4$, where $r = \frac{s}{3} - \frac{7}{6}$. By construction we then have $\Lambda^{r+1} g_0 \in L^2$ and $\Lambda^r g_5 \in L^2$. Remembering $x = \lambda_1 X_2$ and $z = \lambda_1 Z_1$ shows $\Lambda^{r+5/6} h_0 \in L^2$ and $\Lambda^{r+1/2} h_2 \in L^2$. Looking again at the definition of h_0 and h_2 , we see that their time derivatives satisfy $\Lambda^{r+1/3} \partial_t h_0 \in L^2$ and $\Lambda^r \partial_t h_2 \in L^2$.

We begin by reconsidering the time derivative of $(\Lambda^r v, \Lambda^r v)_{L^2} + (\Lambda^{r+1/2} u, \Lambda^{r+1/2} u)_{L^2}$, but now include the fact that h_0 and h_2 depend on time.

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} & \left((\Lambda^r v, \Lambda^r v)_{L^2} + (\Lambda^{r+1/2} u, \Lambda^{r+1/2} u)_{L^2} \right) \\ &= (\Lambda^r v, \Lambda^r g_5)_{L^2} + (\Lambda^r v, \partial_t(\Lambda^r v))_{L^2} + (\Lambda^{r+1/2} u, \partial_t(\Lambda^{r+1/2} u))_{L^2}. \end{aligned}$$

By the Cauchy-Schwarz inequality, the first term can be immediately bounded by $C(\|\Lambda^r v\|_{L^2}^2 + \|\Lambda^r g_5\|_{L^2}^2)$. The term $(\Lambda^r v, \partial_t(\Lambda^r v))_{L^2}$ can be written as

$$\begin{aligned} & (\Lambda^r v, \partial_t(\Lambda^r v))_{L^2} \\ &= \sum_{j=0}^{r-1} (\Lambda^r v, \Lambda^j (\partial_\alpha \lambda_1 (\partial_t h_0) \partial_\alpha \lambda_1 + \lambda_1 \lambda_2 (\partial_t h_2) \lambda_1 \lambda_2) \Lambda^{r-j-1} v)_{L^2}. \end{aligned}$$

Since $\Lambda^{r+1/3} \partial_t h_0 \in L^2$ and $\Lambda^r \partial_t h_2 \in L^2$ as we observed above, we see that this expression is bounded by

$$C(\|\Lambda^r v\|_{L^2}^2 + E_s(\mathcal{W}_e, \mathcal{W}_e))$$

where $s = 3r + \frac{7}{2}$. The term $(\Lambda^{r+1/2}u, \partial_t(\Lambda^{r+1/2}u))_{L^2}$ is bounded in like fashion and we have

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} ((\Lambda^r v, \Lambda^r v)_{L^2} + (\Lambda^{r+1/2}u, \Lambda^{r+1/2}u)_{L^2}) \\ & \leq C((\Lambda^r v, \Lambda^r v)_{L^2} + (\Lambda^{r+1/2}u, \Lambda^{r+1/2}u)_{L^2}) + E_s(\mathcal{W}_e, \mathcal{W}_e). \end{aligned}$$

In order to estimate the the highest order derivatives of x and z we must take advantage of the special form of the equations. Define $\tilde{u} = \mathcal{K}_0 U_1$, and $\tilde{z} = \partial_\alpha Z_1$. Then

$$\begin{aligned} \partial_t V_1 &= -\partial_\alpha h_0 \partial_\alpha (\partial_\alpha \tilde{u}) - h_2 (\partial_\alpha \mathcal{K}_0^{-1}) (\partial_\alpha \tilde{u}) - (\partial_\alpha \tilde{u}) + G_5 \equiv N(\partial_\alpha \tilde{u}) + G_5 \\ \partial_t \tilde{z} &= \partial_\alpha \tilde{u} \end{aligned}$$

Recalling that $\|h_0 + \mu\|_{H^{s-1}} \leq C\|\mathcal{W}\|_{\mathcal{H}^s}$, we see that if $\|\mathcal{W}\|_{\mathcal{H}^s}$ is sufficiently small, N is invertible and maps $H^t \rightarrow H^{t+2}$. Furthermore, if $Z_1 \in H^s$ and $U_1 \in H^{s-\frac{3}{2}}$, then $\tilde{z} \in H^{s-1}$ and $\tilde{u} \in H^{s-\frac{3}{2}}$. With this in mind, we consider the quantity:

$$\|\partial_\alpha^{s-1} \tilde{z} - N^{-1}(\partial_\alpha^{s-1} V_1) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-2} \tilde{z}\|_{L^2}^2.$$

Note that if we know that $V_1 \in H^{s-3}$ and if $\|\mathcal{W}_e\|_{\mathcal{H}_e^s}$ is sufficiently small, this quantity controls the L^2 norm of $\partial_\alpha^{s-1} \tilde{z}$, (and hence the L^2 norm of $\partial_\alpha^s z$.)

Differentiating this with respect to time gives:

$$\begin{aligned} & ([\partial_\alpha^{s-1} \tilde{z} - N^{-1}(\partial_\alpha^{s-1} V_1) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-2} \tilde{z}], [\partial_\alpha^s \tilde{u} - N^{-1} \partial_\alpha^{s-1} (N(\partial_\alpha \tilde{u}) + G_5) \\ & \quad + (\partial_t N^{-1})(\partial_\alpha^{s-1} V_1) + (\partial_t [N^{-1}(\partial_\alpha N)]) \partial_\alpha^{s-2} \tilde{z} + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-1} \tilde{u}]_{L^2} \\ & = ([\partial_\alpha^{s-1} \tilde{z} - N^{-1}(\partial_\alpha^{s-1} V_1) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-2} \tilde{z}], \\ & \quad [-N^{-1}[\partial_\alpha^{s-1}, N] \partial_\alpha \tilde{u} - N^{-1} \partial_\alpha^{s-1} G_5 \\ & \quad + (\partial_t N^{-1})(\partial_\alpha^{s-1} V_1) + (\partial_t [N^{-1}(\partial_\alpha N)]) \partial_\alpha^{s-2} \tilde{z} + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-1} \tilde{u}]_{L^2}, \end{aligned}$$

where we have simplified the expression by commuting the factor of N through ∂_α^{s-1} . We now examine each of the remaining pieces in turn.

Note first that

$$N^{-1}[\partial_\alpha^{s-1}, N](\partial_\alpha \tilde{u}) = \sum_{j=2}^{s-1} c_j N^{-1}(\partial_\alpha^j N) \partial_\alpha^{s-1-j} (\partial_\alpha \tilde{u}) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-1} \tilde{u}$$

for some constants c_j . But if $\tilde{u} \in H^{s-3/2}$, there exists a constant C , such that

$$\begin{aligned} & |([\partial_\alpha^{s-1} \tilde{z} - N^{-1}(\partial_\alpha^{s-1} V_1) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-2} \tilde{z}], \\ & \quad [N^{-1}[\partial_\alpha^{s-1}, N](\partial_\alpha \tilde{u}) - N^{-1}(\partial_\alpha N) \partial_\alpha^{s-1} \tilde{u}]_{L^2})| \\ & \leq C\{ \|\partial_\alpha^{s-1} \tilde{z} - N^{-1}(\partial_\alpha^{s-1} V_1) + N^{-1}(\partial_\alpha N) \partial_\alpha^{s-2} \tilde{z}\|_{L^2}^2 \\ & \quad + \|u\|_{H^{s-\frac{3}{2}}}^2 + \|\mathcal{W}\|_{H^s}^2 \}, \end{aligned}$$

where the term $\|\mathcal{W}\|_{\mathcal{H}^s}^2$ results from bounding the derivatives of h_0 and h_2 that occur in $(\partial_\alpha^j N)$.

Next observe that

$$(\partial_t N^{-1})f = -N^{-1}(\partial_t N)N^{-1}f.$$

Since $(\partial_t N) = -\partial_\alpha(\partial_t h_0)\partial_\alpha - (\partial_t h_2)(\partial_\alpha \mathcal{K}_0^{-1})$, and we know that $\partial_t h_0 \in H^{s-5/2}$, while $\partial_t h_2 \in H^{s-7/2}$, we have $\|(\partial_t N^{-1})f\|_{L^2} \leq C(\|\mathcal{W}\|_{\mathcal{H}^s}\|f\|_{H^{s-3}})$, using the fact that $N^{-1} : H^{s-3} \rightarrow H^{s-1}$ and the fact that $N^{-1}(\partial_t N)$ can be extended to a bounded operator from L^2 to L^2 . Thus,

$$\begin{aligned} & |([\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}], (\partial_t N^{-1})(\partial_\alpha^{s-1}V_1))_{L^2}| \\ & \leq C\{\|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 + \|\mathcal{W}\|_{\mathcal{H}^s}^2 + \|V_1\|_{H^{s-3}}^2\}. \end{aligned}$$

The time derivative of $N^{-1}(\partial_\alpha N)$ is computed in like fashion and we find

$$\begin{aligned} & |([\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}], (\partial_t[N^{-1}(\partial_\alpha N)])(\partial_\alpha^{s-2}\tilde{z}))_{L^2}| \\ & \leq C\{\|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 + \|\mathcal{W}\|_{\mathcal{H}^s}^2 + \|\tilde{z}\|_{H^{s-2}}^2\}. \end{aligned}$$

Finally, using the estimate on G_5 from above, we find

$$\begin{aligned} & |([\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}], N^{-1}\partial_\alpha^{s-1}G_5)_{L^2}| \\ & \leq C\{\|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 + \|\mathcal{W}\|_{\mathcal{H}^s}^2\}. \end{aligned}$$

Collecting all of these pieces we see that we have the estimate

$$\begin{aligned} & \frac{1}{2}\partial_t\|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 \\ & \leq C\{\|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 + \|\mathcal{W}\|_{\mathcal{H}^s}^2\}, \quad (16) \end{aligned}$$

where we used the fact that $\|u\|_{H^{s-3/2}} \leq \|\mathcal{W}\|_{\mathcal{H}^s}$.

In a similar way, if we define $\tilde{x} = \partial_\alpha X_2$, we then have the estimate

$$\begin{aligned} & \frac{1}{2}\partial_t\|\partial_\alpha^{s-1}\tilde{x} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{x}\|_{L^2}^2 \\ & \leq C\{\|\partial_\alpha^{s-1}\tilde{x} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{x}\|_{L^2}^2 + \|\mathcal{W}\|_{\mathcal{H}^s}^2\}. \quad (17) \end{aligned}$$

Hence, we introduce the pre-energy

$$\begin{aligned} \mathcal{E}_r(\omega, \omega) &= (\Lambda^r v, \Lambda^r v)_{L^2} + (\Lambda^{r+1/2}u, \Lambda^{r+1/2}u)_{L^2} \\ & \quad + \|\partial_\alpha^{s-1}\tilde{z} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{z}\|_{L^2}^2 \\ & \quad + \|\partial_\alpha^{s-1}\tilde{x} - N^{-1}(\partial_\alpha^{s-1}V_1) + N^{-1}(\partial_\alpha N)\partial_\alpha^{s-2}\tilde{x}\|_{L^2}^2. \end{aligned}$$

Summarizing these estimates shows that there exists a constant C_5 , such that

$$\frac{d}{dt}(\mathcal{E}_r(\omega, \omega) + E_{s-3}(\mathcal{W}, \mathcal{W})) \leq C_5(\mathcal{E}_r(\omega, \omega) + E_s(\mathcal{W}, \mathcal{W})) \quad (18)$$

with $3r = s - \frac{7}{2}$. Moreover, if $\|\mathcal{W}\|_{\mathcal{H}^s}$, is sufficiently small the norm defined by either $\mathcal{E}_r(\omega, \omega) + E_{s-3}(\mathcal{W}, \mathcal{W})$ or $\mathcal{E}_r(\omega, \omega) + E_s(\mathcal{W}, \mathcal{W})$ is equivalent to the \mathcal{H}_e^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]:

$$\begin{aligned} \partial_t Z_{1,j} &= \mathcal{K}_0 U_{1,j}, \\ \partial_t X_{2,j} &= \mathcal{K}_0 U_{1,j} + \mathcal{S}_1(\mathcal{W}_{j-1})U_{1,j-1}, \\ \partial_t U_{1,j} &= V_{1,j}, \\ \partial_t V_{1,j} &= \mathcal{L}(\mathcal{W}_{j-1})U_{1,j} + G_{5,j-1} \end{aligned} \tag{19}$$

with $\mathcal{W}_{e,j}|_{t=0} = \mathcal{W}_{e,0}$ for all $j \in \mathbb{N}$ in the space $C([0, T_0], \mathcal{H}_e^s)$ for a $T_0 > 0$ sufficiently small, and hence Theorem 4 follows immediately. \square

Definition 2. The *compatibility conditions* for (12) are denoted by \mathcal{C}_p . They are the obvious analogues of $\mathcal{C}_{p,X}$.

3. Long wave analysis

In this section we derive the KdV equations in detail and sketch the very similar derivation of the Kawahara equations. Then we construct an approximation for the solutions of (1) and (2) and we estimate the formal error, the so-called residual.

3.1. Derivation of the KdV and Kawahara 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 $\underline{\alpha} = \varepsilon\alpha$ and the amplitudes are of order $\mathcal{O}(\varepsilon^2)$ in the KdV case or $\mathcal{O}(\varepsilon^4)$ in the Kawahara case. We first consider the former. There will be no fast dynamics, that is to say, no dynamics on time scales of $\mathcal{O}(1)$, and so we also introduce the variable for the slow temporal scale $\underline{t} = \varepsilon t$. The amplitude is scaled in such a way that the leading linear and leading nonlinear terms are of the same order. We make the standard ansatz

$$X_1(\varepsilon, \alpha, t) = \varepsilon \underline{X}_1(\underline{\alpha}, \underline{t}) \quad \text{and} \quad X_2(\varepsilon, \alpha, t) = \varepsilon^2 \underline{X}_2(\underline{\alpha}, \underline{t}), \tag{20}$$

and we use the fact that the operator $\mathcal{K}(X)$ possesses the expansion [Cr85, Lemma 3.7, p. 827]

$$\mathcal{K}(X)U = \mathcal{K}_0 U + [X_1, \mathcal{K}_0]\partial_\alpha U - (X_2 + \mathcal{K}_0(X_2\mathcal{K}_0))\partial_\alpha U + \mathcal{O}(X^2)U$$

so that in the long-wave limit we obtain

$$\mathcal{K}_0(\partial_\alpha) = -\varepsilon\partial_{\underline{\alpha}} - \frac{1}{3}\varepsilon^3\partial_{\underline{\alpha}}^3 - \frac{2}{15}\varepsilon^5\partial_{\underline{\alpha}}^5 + \mathcal{O}(\varepsilon^7).$$

Inserting (20) and the expansion for $\mathcal{K}(X)$ into (1) and (2), we find

$$\begin{aligned} \partial_{\underline{t}}^2 \underline{X}_1 + \partial_{\underline{\alpha}} \underline{X}_2 &= -\varepsilon^2((\partial_{\underline{\alpha}} \underline{X}_1)(\partial_{\underline{t}}^2 \underline{X}_1)) + \mu \varepsilon^2 \partial_{\underline{\alpha}}^3 \underline{X}_2 + \mathcal{O}(\varepsilon^4), \\ \partial_{\underline{t}} \underline{X}_2 + \partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1 &= \varepsilon^2(-\frac{1}{3} \partial_{\underline{t}} \partial_{\underline{\alpha}}^3 \underline{X}_1 + (\partial_{\underline{\alpha}} \underline{X}_1)(\partial_{\underline{t}} \partial_{\underline{\alpha}} \underline{X}_1) - \underline{X}_2 \partial_{\underline{t}} \partial_{\underline{\alpha}} \underline{X}_1) + \mathcal{O}(\varepsilon^4). \end{aligned}$$

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

$$\partial_{\underline{t}}^2 \underline{X}_1 + \partial_{\underline{\alpha}} \underline{X}_2 = -\varepsilon^2((\partial_{\underline{\alpha}} \underline{X}_1)(\partial_{\underline{t}}^2 \underline{X}_1)) + \mu \varepsilon^2 \partial_{\underline{\alpha}}^3 \underline{X}_2, \tag{21}$$

$$\partial_{\underline{t}} \underline{X}_2 + \partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1 = \varepsilon^2(-\frac{1}{3} \partial_{\underline{t}} \partial_{\underline{\alpha}}^3 \underline{X}_1 + (\partial_{\underline{\alpha}} \underline{X}_1)(\partial_{\underline{t}} \partial_{\underline{\alpha}} \underline{X}_1) - \underline{X}_2 \partial_{\underline{t}} \partial_{\underline{\alpha}} \underline{X}_1). \tag{22}$$

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

$$\underline{X}_2 + \partial_{\underline{\alpha}} \underline{X}_1 = \varepsilon^2(-\frac{1}{3} \partial_{\underline{\alpha}}^3 \underline{X}_1 + (\partial_{\underline{\alpha}} \underline{X}_1)^2). \tag{23}$$

Differentiating this with respect to time gives

$$\begin{aligned} \partial_{\underline{t}} \underline{X}_2 + \partial_{\underline{t}} \partial_{\underline{\alpha}} \underline{X}_1 &= -\frac{1}{3} \varepsilon^2 \partial_{\underline{\alpha}}^3 \partial_{\underline{t}} \underline{X}_1 + 2\varepsilon^2 (\partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1)(\partial_{\underline{\alpha}} \underline{X}_1) \\ &= -\frac{1}{3} \varepsilon^2 \partial_{\underline{\alpha}}^3 \partial_{\underline{t}} \underline{X}_1 + \varepsilon^2 (\partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1)(\partial_{\underline{\alpha}} \underline{X}_1) - \varepsilon^2 \underline{X}_2 (\partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1) + \varepsilon^4 \mathcal{E}_1^4, \end{aligned} \tag{24}$$

where

$$\mathcal{E}_1^4 = (-\frac{1}{3} \partial_{\underline{\alpha}}^3 \underline{X}_1 + (\partial_{\underline{\alpha}} \underline{X}_1)^2)(\partial_{\underline{\alpha}} \partial_{\underline{t}} \underline{X}_1).$$

This means that by the choice (23) equation (22) is satisfied up to terms of $\mathcal{O}(\varepsilon^4)$.

Inserting (23) into (21) gives

$$\partial_{\underline{t}}^2 \underline{X}_1 - \partial_{\underline{\alpha}}^2 \underline{X}_1 = \varepsilon^2((\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^4 \underline{X}_1 - \partial_{\underline{\alpha}} \underline{X}_1 \partial_{\underline{t}}^2 \underline{X}_1 - \partial_{\underline{\alpha}}((\partial_{\underline{\alpha}} \underline{X}_1)^2)) + \mathcal{O}(\varepsilon^4) \tag{25}$$

and so

$$\begin{aligned} \partial_{\underline{t}}^2 \underline{X}_1 &= (1 + \varepsilon^2 \partial_{\underline{\alpha}} \underline{X}_1)^{-1} (\partial_{\underline{\alpha}}^2 \underline{X}_1 + \varepsilon^2(\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^4 \underline{X}_1 - \varepsilon^2 \partial_{\underline{\alpha}}((\partial_{\underline{\alpha}} \underline{X}_1)^2)) + \mathcal{O}(\varepsilon^4) \\ &= \partial_{\underline{\alpha}}^2 \underline{X}_1 + \varepsilon^2(\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^4 \underline{X}_1 - \varepsilon^2 \partial_{\underline{\alpha}}((\partial_{\underline{\alpha}} \underline{X}_1)^2) - \varepsilon^2 \partial_{\underline{\alpha}} \underline{X}_1 \partial_{\underline{\alpha}}^2 \underline{X}_1 + \mathcal{O}(\varepsilon^4) \\ &= \partial_{\underline{\alpha}}^2 \underline{X}_1 + \varepsilon^2(\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^4 \underline{X}_1 - \varepsilon^2 \frac{3}{2} \partial_{\underline{\alpha}}((\partial_{\underline{\alpha}} \underline{X}_1)^2) + \mathcal{O}(\varepsilon^4). \end{aligned}$$

Corresponding to our variable $Z_1 = \mathcal{K}_0 X_1$, we introduce in the long-wave limit $\underline{Z}_1 = -\partial_{\underline{\alpha}} \underline{X}_1$. From (25), we see that $\underline{Z}_1 = -\partial_{\underline{\alpha}} \underline{X}_1$ satisfies

$$\partial_{\underline{t}}^2 \underline{Z}_1 - \partial_{\underline{\alpha}}^2 \underline{Z}_1 = \varepsilon^2((\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^4 \underline{Z}_1 + \frac{3}{2} \partial_{\underline{\alpha}}^2((\underline{Z}_1)^2)) + \mathcal{O}(\varepsilon^4). \tag{26}$$

Remark 16. Neglecting terms of order $\mathcal{O}(\varepsilon^4)$, system (26) is ill-posed. Using $\partial_{\underline{t}}^2 \underline{Z}_1 - \partial_{\underline{\alpha}}^2 \underline{Z}_1 = \mathcal{O}(\varepsilon^2)$ gives the well-posed system

$$\partial_{\underline{t}}^2 \underline{Z}_1 - \partial_{\underline{\alpha}}^2 \underline{Z}_1 = \varepsilon^2((\frac{1}{3} - \mu) \partial_{\underline{\alpha}}^2 \partial_{\underline{t}}^2 \underline{Z}_1 + \frac{3}{2} \partial_{\underline{\alpha}}^2((\underline{Z}_1)^2)). \tag{27}$$

Equations (25), (26), and (27) have first been derived by Boussinesq (cf. [Bo1877]).

Our goal is to obtain approximation equations which are independent of ε . Therefore, we make the hypothesis that solutions of (26) are of the form

$$Z_1 = A_1(\underline{\alpha} - \underline{t}, \varepsilon^2 t) + A_2(\underline{\alpha} + \underline{t}, \varepsilon^2 t). \tag{28}$$

Letting $T = \varepsilon^2 t$, this yields

$$\begin{aligned} &\varepsilon^2(-2\partial_{\underline{\alpha}}\partial_T(A_1 - A_2)) + \varepsilon^4\partial_T^2(A_1 + A_2) \\ &= \varepsilon^2\left(\left(\frac{1}{3} - \mu\right)\partial_{\underline{\alpha}}^4(A_1 + A_2) + \frac{3}{2}\partial_{\underline{\alpha}}^2((A_1)^2 + (A_2)^2) + r_1\right) + \mathcal{O}(\varepsilon^4) \end{aligned}$$

with $r_1 = 3\partial_{\underline{\alpha}}^2(A_1A_2)$. Since A_1 and A_2 are spatially localized and since the two wave packets only meet on a time scale $\mathcal{O}(1/\varepsilon)$ which is relatively short compared to the overall 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 up to terms of $\mathcal{O}(\varepsilon^4)$, the water-wave problem can completely be described by two decoupled KdV equations

$$\begin{aligned} 2\partial_T A_1 &= -\left(\frac{1}{3} - \mu\right)\partial_{\underline{\alpha}}^3 A_1 - \frac{3}{2}\partial_{\underline{\alpha}}((A_1)^2), \\ 2\partial_T A_2 &= \left(\frac{1}{3} - \mu\right)\partial_{\underline{\alpha}}^3 A_2 + \frac{3}{2}\partial_{\underline{\alpha}}((A_2)^2). \end{aligned} \tag{29}$$

This has been shown in [Sch98] for (27) as a model problem and in [SW00a] for the water-wave problem without surface tension. The model problem in the Kawahara case has been handled in [SW00c].

Remark 17. The Kawahara equations are obtained if we set $\mu = \frac{1}{3} + \nu\varepsilon^2, T = \varepsilon^4 t$,

$$X_1(\varepsilon, \alpha, t) = \varepsilon^3 \underline{X}_1(\underline{\alpha}, t) \quad \text{and} \quad X_2(\varepsilon, \alpha, t) = \varepsilon^4 \underline{X}_2(\underline{\alpha}, t).$$

With exactly the same calculations we end up with two decoupled Kawahara equations

$$\begin{aligned} 2\partial_T A_1 &= \nu\partial_{\underline{\alpha}}^3 A_1 - \frac{2}{90}\partial_{\underline{\alpha}}^5 A_1 - \frac{3}{2}\partial_{\underline{\alpha}}((A_1)^2), \\ 2\partial_T A_2 &= -\nu\partial_{\underline{\alpha}}^3 A_2 + \frac{2}{90}\partial_{\underline{\alpha}}^5 A_2 + \frac{3}{2}\partial_{\underline{\alpha}}((A_2)^2). \end{aligned} \tag{30}$$

The details are left to the reader.

3.2. The approximation and estimates for the residual

It is the aim of this section to construct an approximation $\varepsilon^2\Psi_e$ out of (23) and (29) for the extended system (12). This has to be done in such a way that the extended residual

$$\text{Res}_e(\mathcal{W}_e) = (\text{Res}_{Z_1}(\mathcal{W}_e), \text{Res}_{X_2}(\mathcal{W}_e), \text{Res}_{U_1}(\mathcal{W}_e), \text{Res}_{V_1}(\mathcal{W}_e))$$

with

$$\begin{aligned} \text{Res}_{Z_1}(\mathcal{W}_e) &= -\partial_t Z_1 + \mathcal{K}_0 U_1, \\ \text{Res}_{X_2}(\mathcal{W}_e) &= -\partial_t X_2 + \mathcal{K}_0 U_1 + \mathcal{S}_1(\mathcal{W})U_1, \\ \text{Res}_{U_1}(\mathcal{W}_e) &= -\partial_t U_1 + V_1, \\ \text{Res}_{V_1}(\mathcal{W}_e) &= -\partial_t V_1 + \mathcal{L}(\mathcal{W})U_1 + G_5 \end{aligned}$$

is small. The residual contains all terms which do not drop out after inserting the extended approximation $\varepsilon^2\Psi_e$ into system (12). We additionally define

$$\text{Res}(\mathcal{W}) = (\text{Res}_{Z_1}(\mathcal{W}), \text{Res}_{X_2}(\mathcal{W}), \text{Res}_{U_1}(\mathcal{W})),$$

where in computing $\text{Res}_{U_1}(\mathcal{W})$, we use the fact that $\partial_t U_1$ can also be rewritten as in (34) and hence,

$$\begin{aligned} \text{Res}_{U_1} &= -\partial_t U_1 + (1 - \mathcal{M}_2 Z_1 + (\partial_\alpha X_2)\mathcal{K}(\mathcal{W}))^{-1} \\ &\quad \times (-(\partial_\alpha X_2)(1 + [\partial_t, \mathcal{K}(\mathcal{W})]U_1) \\ &\quad + \mu R(\partial_\alpha X, \partial_\alpha^2 X) + \mu S(\partial_\alpha X, \partial_\alpha^3 X)). \end{aligned}$$

With these preparations we define the approximations

$$\varepsilon^2\Psi = (\varepsilon^2\psi_{Z_1}, \varepsilon^2\psi_{X_2}, \varepsilon^2\psi_{U_1}) \quad \text{and} \quad \varepsilon^2\Psi_e = (\varepsilon^2\psi_{Z_1}, \varepsilon^2\psi_{X_2}, \varepsilon^2\psi_{U_1}, \varepsilon^3\psi_{V_1}), \tag{31}$$

where

$$\begin{aligned} \varepsilon^2\psi_{Z_1}(\underline{\alpha}, t) &= \varepsilon^2(A_1(\underline{\alpha} - t, \varepsilon^2 t) + A_2(\underline{\alpha} + t, \varepsilon^2 t)), \\ \varepsilon^2\psi_{X_2}(\underline{\alpha}, t) &= \varepsilon^2\psi_{Z_1} + \varepsilon^4(\psi_{Z_1})^2, \\ \varepsilon^2\psi_{U_1}(\underline{\alpha}, t) &= \varepsilon^2\partial_t \mathcal{K}_0^{-1}\psi_{Z_1}(\underline{\alpha}, t), \\ \varepsilon^3\psi_{V_1}(\underline{\alpha}, t) &= \varepsilon^3\partial_t^2 \mathcal{K}_0^{-1}\psi_{Z_1}(\underline{\alpha}, t). \end{aligned} \tag{32}$$

Note that the V_1 component is scaled with ε^3 .

Remark 18. The motivation for the definitions of ψ_{U_1} and ψ_{V_1} are that $\varepsilon^2\psi_{U_1} = \varepsilon^2\partial_t \psi_{X_1}$ where $\psi_{X_1} = \mathcal{K}_0^{-1}\psi_{Z_1}$. But $\mathcal{K}_0^{-1}\psi_{Z_1}$ 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 ψ_{U_1} is well defined. A similar argument applies to ψ_{V_1} . The motivation for the definition of ψ_{X_2} is the formal calculation (23).

The approximation $\varepsilon^2\Psi_e$ has the following properties.

Lemma 5. 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 (29) with

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

Then,

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

Proof. This follows more or less line for line the proof of [SW00a, Lemma 5.4]. \square

The following lemma ensures that the approximation $\varepsilon^2\Psi_e$ defined in (31) 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 6. Fix $s \geq 1$. For all $C_A > 0$ there exist $C_{\text{Res}}, \varepsilon_0 > 0$ 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 (29) 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 (29) with $A_2 = 0$ and

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

Then for

$$Q_1 = \text{Res}_\varepsilon(\varepsilon^2 \Psi_\varepsilon(s)) \quad \text{and} \quad Q_2 = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \varepsilon^5 \partial_{\underline{\alpha}}((\tau_{-t} S_\varepsilon A_1)(\tau_t S_\varepsilon A_2)) \\ -\frac{3}{2} \varepsilon^6 \partial_t \partial_{\underline{\alpha}}((\tau_{-t} S_\varepsilon A_1)(\tau_t S_\varepsilon A_2)) \end{pmatrix}$$

the following estimates hold:

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

and

$$\varepsilon^{7/2} q(t) := \|Q_2(t)\|_{\mathcal{H}_\varepsilon^s} \leq C_{\text{Res}} \varepsilon^{9/2} / (1 + (\varepsilon t)^2).$$

Proof. The proof is very close to the [SW00a, Lemma 5.5]. Recalling that $Z_1 = \mathcal{K}_0 X_1$ and $\underline{Z}_1 = -\partial_{\underline{\alpha}} X_1$, we obtain

$$\begin{aligned} \text{Res}_{X_2}(\mathcal{W}_\varepsilon) &= -\varepsilon^3 \partial_t \underline{X}_2 - \varepsilon^3 \partial_{\underline{\alpha}} U_1 \\ &\quad + \varepsilon^5 \left(-\frac{1}{3} \partial_{\underline{\alpha}}^3 U_1 - \underline{Z}_1 \partial_{\underline{\alpha}} U_1 - \underline{X}_2 \partial_{\underline{\alpha}} U_1\right) + \mathcal{O}(\varepsilon^7), \\ \text{Res}_{U_1}(\mathcal{W}_\varepsilon) &= -\varepsilon^3 \partial_t U_1 - \varepsilon^3 \partial_{\underline{\alpha}} X_2 + \varepsilon^5 (\underline{Z}_1 \partial_t U_1) + \mu \varepsilon^5 \partial_{\underline{\alpha}}^3 X_2 + \mathcal{O}(\varepsilon^7) \\ &= -\varepsilon^3 \partial_t U_1 - \varepsilon^3 \partial_{\underline{\alpha}} X_2 - \varepsilon^5 (\underline{Z}_1 \partial_{\underline{\alpha}} X_2) + \mu \varepsilon^5 \partial_{\underline{\alpha}}^3 X_2 + \mathcal{O}(\varepsilon^7), \\ \text{Res}_{V_1}(\mathcal{W}_\varepsilon) &= -\varepsilon^4 \partial_t V_1 - \varepsilon^4 \partial_t \partial_{\underline{\alpha}} X_2 - \varepsilon^6 (\partial_t \underline{Z}_1)(\partial_{\underline{\alpha}} X_2) \\ &\quad - \varepsilon^6 \underline{Z}_1(\partial_t \partial_{\underline{\alpha}} X_2) - \mu \varepsilon^6 \partial_{\underline{\alpha}}^3 U_1 + \mathcal{O}(\varepsilon^8) \\ &= -\varepsilon^4 \partial_t V_1 + \varepsilon^4 \partial_{\underline{\alpha}}^2 U_1 \\ &\quad - \varepsilon^6 \partial_{\underline{\alpha}} \left(\left(\mu - \frac{1}{3}\right) \partial_{\underline{\alpha}}^3 U_1 - \underline{Z}_1 \partial_{\underline{\alpha}} U_1 - \underline{X}_2 \partial_{\underline{\alpha}} U_1\right) \\ &\quad + \varepsilon^6 (\partial_{\underline{\alpha}} U_1)(\partial_{\underline{\alpha}} X_2) + \varepsilon^6 \underline{Z}_1(\partial_{\underline{\alpha}}^2 U_1) + \mathcal{O}(\varepsilon^8) \\ &= -\varepsilon^4 \partial_t V_1 + \varepsilon^4 \partial_{\underline{\alpha}}^2 U_1 + \varepsilon^6 \left(\frac{1}{3} - \mu\right) \partial_{\underline{\alpha}}^4 U_1 + 2\varepsilon^6 \underline{Z}_1 \partial_{\underline{\alpha}}^2 U_1 \\ &\quad + 2\varepsilon^6 (\partial_{\underline{\alpha}} X_2)(\partial_{\underline{\alpha}} U_1) + \varepsilon^6 (\partial_{\underline{\alpha}} \underline{Z}_1)(\partial_{\underline{\alpha}} U_1) + \varepsilon^6 \underline{X}_2(\partial_{\underline{\alpha}}^2 U_1) + \mathcal{O}(\varepsilon^8). \end{aligned}$$

Inserting $\mathcal{W}_\varepsilon = \varepsilon^2 \Psi_\varepsilon$ gives by construction

$$\begin{aligned} \text{Res}_{Z_1}(\varepsilon^2 \Psi_\varepsilon) &= 0, \\ \text{Res}_{X_2}(\varepsilon^2 \Psi_\varepsilon) &= 0 + \mathcal{O}(\varepsilon^7), \\ \text{Res}_{U_1}(\varepsilon^2 \Psi_\varepsilon) &= -\frac{1}{2} \varepsilon^5 \partial_{\underline{\alpha}}((\tau_{-t} S_\varepsilon A_1)(\tau_t S_\varepsilon A_2)) + \mathcal{O}(\varepsilon^7), \\ \text{Res}_{V_1}(\varepsilon^2 \Psi_\varepsilon) &= -\frac{3}{2} \varepsilon^6 \partial_t \partial_{\underline{\alpha}}((\tau_{-t} S_\varepsilon A_1)(\tau_t S_\varepsilon A_2)) + \mathcal{O}(\varepsilon^8), \end{aligned}$$

where the cancellation in the second line is due to the quadratic terms in ψ_{X_2} . Using [SW00a, Lemma 2.3] which shows that

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

if $|f(k)| \leq C|k|^n$, establishes

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

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

$$\begin{aligned} \|S_\varepsilon(uv)\|_{H^s} &\leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{C_b^s} \|S_\varepsilon(u\rho_1)\|_{H^s} \|S_\varepsilon(v\rho_2)\|_{C_b^s} \\ &\leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{C_b^s} \varepsilon^{-1/2} \|u\rho_1\|_{H^s} \|v\rho_2\|_{C_b^s} \\ &\leq \|S_\varepsilon(\rho_1 \rho_2)^{-1}\|_{C_b^s} \varepsilon^{-1/2} \|u\rho_1\|_{H^s} \|v\rho_2\|_{H^{s+1}}, \end{aligned}$$

we have

$$\begin{aligned} \|(\tau_{-t} S_\varepsilon A_1)(\tau_t S_\varepsilon A_2)\|_{H^s} &\leq C \sup_{\alpha} |((1 + (\varepsilon(\alpha - t))^2)(1 + (\varepsilon(\alpha + t))^2))^{-1}| \\ &\quad \times \varepsilon^{-1/2} \left(\sup_{T \in [0, T_0]} \|A_1\|_{H^s(2)} \right) \left(\sup_{T \in [0, T_0]} \|A_2\|_{H^{s+1}(2)} \right) \\ &\leq C_{\text{Res}} \varepsilon^{-1/2} / (1 + (\varepsilon t)^2), \end{aligned}$$

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

For the case $\mu = \frac{1}{3} + \nu \varepsilon^2$ the details are again left to the reader. The results are summarized in the following lemma.

Lemma 7. Fix $s \geq 1$. For all $C_A > 0$ there exist $C_\psi, C_{\text{Res}}, \varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following is true. Let $A_1, A_2 \in C([0, T_1], H^{s+8}(2))$ be solutions of (30) with

$$\sup_{T \in [0, T_1]} \|(A_1, A_2)(T)\|_{H^{s+8}(2)} \leq C_A$$

or let $A_1, A_2 \in C([0, T_1], H^{s+8})$ be solutions of (30) with $A_2 = 0$ and

$$\sup_{T \in [0, T_1]} \|(A_1, A_2)(T)\|_{H^{s+8}} \leq C_A.$$

Then we find, for

$$\begin{aligned}
 \varepsilon^4 \psi_{Z_1}(\underline{\alpha}, \underline{t}) &= \varepsilon^4 (A_1(\underline{\alpha} - \underline{t}, \varepsilon^4 \underline{t}) + A_2(\underline{\alpha} + \underline{t}, \varepsilon^4 \underline{t})), \\
 \varepsilon^4 \psi_{X_2}(\underline{\alpha}, \underline{t}) &= \varepsilon^4 \psi_{Z_1} + \varepsilon^8 (\psi_{Z_1})^2, \\
 \varepsilon^4 \psi_{U_1}(\underline{\alpha}, \underline{t}) &= \varepsilon^4 \partial_t \mathcal{K}_0^{-1} \psi_{Z_1}(\underline{\alpha}, \underline{t}), \\
 \varepsilon^5 \psi_{V_1}(\underline{\alpha}, \underline{t}) &= \varepsilon^5 \partial_t^2 \mathcal{K}_0^{-1} \psi_{Z_1}(\underline{\alpha}, \underline{t}),
 \end{aligned} \tag{33}$$

that

$$\sup_{t \in [0, T_1/\varepsilon^5]} \|\Psi_e(t)\|_{\mathcal{H}_e^s} \leq C\Psi.$$

Moreover, for

$$Q_1 = \text{Res}_\varepsilon(\varepsilon^4 \Psi_e(s)) \quad \text{and} \quad Q_2 = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \varepsilon^9 \partial_{\underline{\alpha}}((\tau_{-\underline{t}} S_\varepsilon A_1)(\tau_{\underline{t}} S_\varepsilon A_2)) \\ -\frac{3}{2} \varepsilon^{10} \partial_{\underline{t}} \partial_{\underline{\alpha}}((\tau_{-\underline{t}} S_\varepsilon A_1)(\tau_{\underline{t}} S_\varepsilon A_2)) \end{pmatrix},$$

the following estimates hold:

$$\sup_{t \in [0, T_1/\varepsilon^5]} \|Q_1 - Q_2\|_{H_e^s} \leq C_{\text{Res}} \varepsilon^{21/2}$$

and

$$\varepsilon^{11/2} q(t) := \|Q_2(t)\|_{H_e^s} \leq C_{\text{Res}} \varepsilon^{17/2} / (1 + (\varepsilon t)^2).$$

Remark 19. We recall here that we use T_1 rather than T_0 to denote the existence interval of solutions of the Kawahara equations in the function space $H^{s+8}(2)$ because in contrast to the case of the KdV equations we are unaware of any results of the global well-posedness of the Kawahara equations in these weighted function spaces.

4. The error estimates

Now we are ready to formulate our main result. For (1) and (2) written as the first order system

$$\begin{aligned}
 \partial_t Z_1 &= \mathcal{K}_0 U_1, \\
 \partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{S}_1(X) U_1, \\
 \partial_t U_1 &= -(1 - \mathcal{M}_2 Z_1 + (\partial_\alpha X_2) \mathcal{K}_0 \\
 &\quad + (\partial_\alpha X_2) \mathcal{S}_1(X))^{-1} [(\partial_\alpha X_2)(1 + [\partial_t, \mathcal{S}_1(X)] U_1) \\
 &\quad - \mu R(\partial_\alpha X, \partial_\alpha^2 X) - \mu S(\partial_\alpha X, \partial_\alpha^3 X)]
 \end{aligned} \tag{34}$$

in the variables collected in \mathcal{W} , we show that there exist solutions which behave in approximately the same way as predicted by the approximation $\varepsilon^2 \Psi$ defined in (32) and constructed via the solutions of the two decoupled KdV equations (29).

Theorem 5. Fix $s \geq 6$ and let $\beta = 7/2$. Then 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 (29) 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 (29) 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 there is a unique solution $\mathcal{W} = \varepsilon^2 \Psi + \varepsilon^\beta R \in C([0, T_0/\varepsilon^3], \mathcal{H}^s)$ of (34) which satisfies

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

Remark 20. This is not in contradiction with Section 2. Local existence and uniqueness of solutions for (34) follows indirectly since (34) is a subsystem of (12), namely the system of all solutions of (12) in \mathcal{C}_p .

Proof of Theorem 1. From [SW00a, Lemma 2.5] we have the estimate

$$\sup_{t \in [0, T_0/\varepsilon_3]} |X_1|_{\alpha=0}(t) \leq C_2 \varepsilon^{1/2}.$$

Also, [SW00a, Lemma 2.6] ensures that in the long-wave limit the initial conditions ϕ_1 and ϕ_2 of (34) can be separated into the initial conditions $A_1|_{T=0}$ and $A_2|_{T=0}$ of the two decoupled KdV equations and that we have a reparametrization of the physical solution with an initial error $\varepsilon^{7/2} R|_{t=0}$ such that Theorem 5 can be applied. Then by [SW00a, Lemma 6.3] we construct via (32) a formal approximation of the solutions of (34) for all $t \in [0, T_0/\varepsilon^3]$ with $T_0 > 0$ finite, but arbitrarily large. Then Theorem 5 ensures that the water-wave problem (34) in the long-wave limit really can be described approximately by two decoupled KdV equations (29). So the only step remaining in the proof of Theorem 1 is to establish the validity of Theorem 5. \square

Preliminaries. The operator $\mathcal{S}_1(X)$ is split into its quadratic and its higher order terms.

Lemma 8. The operator $\mathcal{K}(\mathcal{W})$ possesses the expansion

$$\mathcal{K}(\mathcal{W})U_1 = \mathcal{K}_0 U_1 + B_1(\mathcal{W})U_1 + \mathcal{S}_2(\mathcal{W})U_1$$

with

$$\begin{aligned} B_1(\mathcal{W})U_1 &= \mathcal{M}_1(Z_1, \partial_\alpha U_1) - (X_2 + \mathcal{K}_0(X_2 \mathcal{K}_0)) \partial_\alpha U_1 \\ &= \mathcal{M}_1(Z_1, \partial_\alpha U_1) - (1 + \mathcal{K}_0^2)(X_2 \partial_\alpha U_1) - \mathcal{K}_0([X_2, \mathcal{K}_0] \partial_\alpha U_1), \\ \mathcal{S}_2(\mathcal{W})U_1 &= \mathcal{O}(\|\mathcal{W}\|^2)U_1. \end{aligned}$$

Proof. See [SW00a, Lemma 3.8] and [SW00a, Remark 3.9]. \square

Remark 21. Due to Lemma 3 and Lemma 4 the operator B_1 is smoothing. Thus,

$$\|B_1(\mathcal{W})U_1\|_{H^s} \leq C\|\mathcal{W}\|_{H^s}\|U_1\|_{H^3}.$$

This quantity can also be estimated with the aid of Remark 11, obtaining

$$\|B_1(\mathcal{W})U_1\|_{H^s} \leq C \min(\|\hat{\mathcal{W}}\|_{L^1(s)}\|U_1\|_{H^3}, \|\mathcal{W}\|_{H^s}\|\hat{U}\|_{L^1(s)}). \quad (35)$$

As noted in Remark 11, this latter estimate is useful if we wish to minimize the loss of powers of ε due to the scaling of the L^2 norms.

As a direct consequence of the smoothing properties of B_1 and Lemma 1 we have the following corollary.

Corollary 3. *Assume the situation of Lemma 1. Then, for all $s \geq 6$,*

$$\begin{aligned} \|\mathcal{S}_2(\mathcal{W})U_1\|_{H^s} &\leq C\|\mathcal{W}\|_{\mathcal{H}^s}^2\|U_1\|_{H^3}, \\ \|\partial_t(\mathcal{S}_2(\mathcal{W})U_1)\|_{H^{s-3}} &\leq C\|\mathcal{W}\|_{\mathcal{H}^s}^2(\|U_1\|_{H^3} + \|V_1\|_{H^3}), \\ \|[\partial_t^2, \mathcal{S}_2(\mathcal{W})]U_1\|_{H^{s-3}} &\leq C\|\mathcal{W}_e\|_{\mathcal{H}^s}^2(\|U_1\|_{H^4} + \|V_1\|_{H^4}), \\ \|\partial_\alpha(\mathcal{S}_2(\mathcal{W})U_1)\|_{H^{s-1}} &\leq C\|\mathcal{W}\|_{\mathcal{H}^s}^2\|U_1\|_{H^3}. \end{aligned}$$

With these preparations we can begin the **Proof of Theorem 5**.

Remark 22. 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 standing hypothesis:

(HS) *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 $\sup_{\tau \in (0,t)} \|R_e\|_{\mathcal{H}^s} \leq C_R$.*

We proceed as follows. As explained in [SW00a], and as can be seen just by examining the linearized equations, outside the set \mathcal{C}_p there is a secular growth of the solutions of (12). However, this growth occurs only in the size of the components of \mathcal{W} – not for the derivatives of these functions. In order to control the error over the very long time scales of interest here we must use the property $\mathcal{W}(t) \in \mathcal{C}_p$ by taking (34) for the evolution of \mathcal{W} and then using (12) for the evolution of the derivatives of \mathcal{W} and for V_1 . We follow Remark 5 and write a solution $\mathcal{W} = (Z_1, X_2, U_1)$ of (34) as a sum of the approximation $\varepsilon^2\Psi$ and an error $\varepsilon^\beta R$ with

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

and we write a solution $\mathcal{W}_e = (Z_1, X_2, U_1, V_1)$ of (12) as a sum of the approximation $\varepsilon^2\Psi_e$ and an error $\varepsilon^\beta R_e$ with

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

We expand (34) up to terms of quadratic order. With

$$\begin{aligned} Q &= (1 + 2\partial_\alpha X_1 + (\partial_\alpha X_1)^2 + (\partial_\alpha X_2)^2)^{1/2} = 1 + \partial_\alpha X_1 + \text{h.o.t.} \\ &= 1 - \mathcal{M}_2 Z_1 + \text{h.o.t.}, \\ R &= -3(\partial_\alpha^2 X_1)(\partial_\alpha^2 X_2) + \text{h.o.t.} = 3(\partial_\alpha \mathcal{M}_2 Z_1)(\partial_\alpha^2 X_2) + \text{h.o.t.}, \\ S &= (1 + 3\mathcal{M}_2 Z_1 + \text{h.o.t.})(\partial_\alpha X_2)(\partial_\alpha^2 \mathcal{M}_2 Z_1) + (1 - \mathcal{M}_2 Z_1)\partial_\alpha^3 X_2 \\ &= \partial_\alpha^3 X_2 + (\partial_\alpha X_2)(\partial_\alpha^2 \mathcal{M}_2 Z_1) + 2(\mathcal{M}_2 Z_1)\partial_\alpha^3 X_2 + \text{h.o.t.} \end{aligned}$$

we obtain

$$\begin{aligned} \partial_t Z_1 &= \mathcal{K}_0 U_1, \\ \partial_t X_2 &= \mathcal{K}_0 U_1 + \mathcal{M}_1(Z_1, \partial_\alpha U_1) - (X_2 + \mathcal{K}_0(X_2 \mathcal{K}_0))\partial_\alpha U_1 + \mathcal{O}(\|\mathcal{W}\|^3), \\ \partial_t U_1 &= -\partial_\alpha X_2 - (\mathcal{M}_2 Z_1)\partial_\alpha X_2 + (\partial_\alpha X_2)\mathcal{K}_0 \partial_\alpha X_2 \\ &\quad + \mu \partial_\alpha^3 X_2 + \mu(\partial_\alpha X_2)(\partial_\alpha^2 \mathcal{M}_2 Z_1) \\ &\quad + 2\mu(\mathcal{M}_2 Z_1)\partial_\alpha^3 X_2 + 3\mu(\partial_\alpha \mathcal{M}_2 Z_1)(\partial_\alpha^2 X_2) + \mathcal{O}(\|\mathcal{W}\|^3). \end{aligned} \tag{36}$$

Now insert the ansatz $\mathcal{W} = \varepsilon^2 \Psi + \varepsilon^\beta R$ into (36), which yields

$$\begin{aligned} \partial_t R_{Z_1} &= \mathcal{K}_0 R_{U_1}, \\ \partial_t R_{X_2} &= \mathcal{K}_0 R_{U_1} + \varepsilon^2 \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha R_{U_1}) - \varepsilon^2(\psi_{X_2} + \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0))\partial_\alpha R_{U_1} \\ &\quad + \varepsilon^3 \mathcal{M}_1(R_{Z_1}, \partial_\alpha \psi_{U_1}) - \varepsilon^3(R_{X_2} + \mathcal{K}_0(R_{X_2} \mathcal{K}_0))\partial_\alpha \psi_{U_1} + \mathcal{N}_1 \\ &= \mathcal{K}_0 R_{U_1} + \varepsilon^2 \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha R_{U_1}) - \varepsilon^2(\psi_{X_2} + \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0))\partial_\alpha R_{U_1} + \mathcal{N}_2 \\ &= \mathcal{K}_0 R_{U_1} + \varepsilon^2 s_1(\Psi, \partial_\alpha R_{U_1}) + \mathcal{N}_2, \\ \partial_t R_{U_1} &= -\partial_\alpha R_{X_2} - \varepsilon^2(\mathcal{M}_2 \psi_{Z_1})\partial_\alpha R_{X_2} + \varepsilon^3(\partial_\alpha \psi_{X_2})\mathcal{K}_0 \partial_\alpha R_{X_2} \\ &\quad - \varepsilon^3(\mathcal{M}_2 R_{Z_1})\partial_\alpha \psi_{X_2} + \varepsilon^3(\partial_\alpha R_{X_2})\mathcal{K}_0 \partial_\alpha \psi_{X_2} \\ &\quad + \mu \partial_\alpha^3 R_{X_2} + \mu \varepsilon^3(\partial_\alpha \psi_{X_2})\partial_\alpha^2 \mathcal{M}_2 R_{Z_1} + \mu \varepsilon^4(\partial_\alpha R_{X_2})\partial_\alpha^2 \mathcal{M}_2 \psi_{Z_1} \\ &\quad + 2\mu \varepsilon^5(\mathcal{M}_2 R_{Z_1})\partial_\alpha^3 \psi_{X_2} + 2\mu \varepsilon^2(\mathcal{M}_2 \psi_{Z_1})\partial_\alpha^3 R_{X_2} \\ &\quad + 3\mu \varepsilon^3(\mathcal{M}_2 \partial_\alpha \psi_{Z_1})(\partial_\alpha^2 R_{X_2}) + 3\mu \varepsilon^4(\mathcal{M}_2 \partial_\alpha R_{Z_1})(\partial_\alpha^2 \psi_{X_2}) + \mathcal{N}_3 \\ &= -\partial_\alpha R_{X_2} - \varepsilon^2(\mathcal{M}_2 \psi_{Z_1})\partial_\alpha R_{X_2} \\ &\quad + \mu \partial_\alpha^3 R_{X_2} + 2\mu \varepsilon^2(\mathcal{M}_2 \psi_{Z_1})\partial_\alpha^3 R_{X_2} + \mathcal{N}_4 \end{aligned} \tag{37}$$

with

$$\begin{aligned} \|\mathcal{N}_1\|_{H^s} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \|\mathcal{N}_2\|_{H^s} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \|\mathcal{N}_3\|_{H^{s-3}} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \|\mathcal{N}_4\|_{H^{s-3}} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \varepsilon^2 s_1(\Psi, \partial_\alpha R_{U_1}) &= \varepsilon^2 \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha R_{U_1}) - \varepsilon^2(\psi_{X_2} + \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0))\partial_\alpha R_{U_1} \end{aligned}$$

under the hypothesis (HS).

Remark 23. Note that in order to avoid losing any further powers of ε in the estimates of \mathcal{N}_2 and \mathcal{N}_4 , we use estimate (35) to bound terms like $\varepsilon^3 \mathcal{M}_1(R_{Z_1}, \partial_\alpha \psi_{U_1}) + \varepsilon^3 (R_{X_2} + \mathcal{K}_0(R_{X_2} \mathcal{K}_0)) \partial_\alpha \psi_{U_1}$ that appear in these expressions. This allows us to retain the factor of ε^3 in front of the $\|R\|_{\mathcal{H}^s}$ term which is necessary to apply Gronwall’s inequality over the long time intervals of interest here.

The term s_1 is bilinear in its arguments with the regularity of ψ , i.e., $\partial_\alpha R_{U_1}$ is smoothed by this operator according to Lemma 3 and Lemma 4. This term of order $\mathcal{O}(\varepsilon^2)$ in the equations for $\partial_t R_{X_2}$ can be eliminated by the following procedure. We replace $\partial_\alpha R_{U_1}$ in $\varepsilon^2 s_1(\psi, \partial_\alpha R_{U_1})$ by

$$\partial_\alpha R_{U_1} = (\partial_\alpha \mathcal{K}_0^{-1})(\partial_t R_{X_2} - \varepsilon^2 s_1(\Psi, \partial_\alpha R_{U_1}) + \mathcal{N}_2)$$

and so

$$\partial_t R_{X_2} = \mathcal{K}_0 R_{U_1} + \varepsilon^2 s_1(\Psi, (\partial_\alpha \mathcal{K}_0^{-1}) \partial_t R_{X_2}) + \mathcal{N}_{2a},$$

where \mathcal{N}_{2a} obeys the same estimates as \mathcal{N}_2 . With $\partial_t \Psi = \mathcal{O}(\varepsilon)$ we obtain

$$\partial_t (R_{X_2} - \varepsilon^2 s_1(\Psi, (\partial_\alpha \mathcal{K}_0^{-1}) R_{X_2})) = \mathcal{K}_0 R_{U_1} + \mathcal{N}_{2b},$$

where \mathcal{N}_{2b} obeys the same estimates as \mathcal{N}_2 . Then we make the change of variables

$$R_{Y_2} = R_{X_2} - \varepsilon^2 s_1(\Psi, (\partial_\alpha \mathcal{K}_0^{-1}) R_{X_2}), \tag{38}$$

or equivalently

$$R_{X_2} = R_{Y_2} + \varepsilon^2 s_1(\Psi, (\partial_\alpha \mathcal{K}_0^{-1}) R_{Y_2}) + \mathcal{O}(\varepsilon^4),$$

which is of the form identity plus something small and smooth in R_{Y_2} . Thus, we obtain the system

$$\begin{aligned} \partial_t R_{Z_1} &= \mathcal{K}_0 R_{U_1}, \\ \partial_t R_{Y_2} &= \mathcal{K}_0 R_{U_1} + \mathcal{N}_5, \\ \partial_t R_{U_1} &= -\partial_\alpha R_{Y_2} + \mu \partial_\alpha^3 R_{Y_2} + \varepsilon^2 s_2(\Psi, R_{Y_2}) + \mathcal{N}_{4a} \end{aligned} \tag{39}$$

with

$$\begin{aligned} \|\mathcal{N}_{4a}\|_{H^{s-3}} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \|\mathcal{N}_5\|_{H^s} &\leq C(\varepsilon^3 \|R\|_{\mathcal{H}^s} + \varepsilon^\beta \|R\|_{\mathcal{H}^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t)), \\ \varepsilon^2 s_2(\Psi, R_{Y_2}) &= -\varepsilon^2 \partial_\alpha (s_1(\Psi, \mathcal{K}_0^{-1} \partial_\alpha R_{Y_2})) - \varepsilon^2 \psi_{Z_1} \partial_\alpha R_{Y_2} + 2\mu \varepsilon^2 \psi_{Z_1} \partial_\alpha^3 R_{Y_2} \end{aligned}$$

which holds under (HS). Here we used $\hat{\mathcal{N}}_2(k) = 1 + \mathcal{O}(k^2)$ for small $|k|$ and the fact that ψ_{Z_1} is a scaled function.

It is easy to see that the $H^{s-3/2}$ norm of R_U cannot be estimated by the term \mathcal{N}_4 since we would lose regularity. Therefore, to estimate the highest derivative we have to use the equation $\partial_t U_1 = V_1$ and extend (34) by the fourth equation of (12). For clarity, we recall that solutions of (12) which lie in the subset \mathcal{C}_p describe the same solutions of the water-wave problem as (34). We also recall that while general solutions of (12) are expected to grow “rapidly” with time, which is why

we cannot use that form of the equations for all our estimates, the derivatives of \mathcal{W}_ε are *not* expected to exhibit this secular growth and hence we can use (12) to control the derivatives of R_U without losing control of our estimates of the remainder.

When we introduce the normal form transformation (38), the vector R is changed to $\tilde{R} = (R_{Z_1}, R_{Y_2}, R_{U_1})$ and extended to $\tilde{R}_e = (R_{Z_1}, R_{Y_2}, R_{U_1}, R_{V_1})$. For $\varepsilon > 0$ sufficiently small, R can be estimated by \tilde{R} in \mathcal{H}^s and vice versa.

Writing $V_1 = \varepsilon^3 \psi_{V_1} + \varepsilon^\beta R_{V_1}$ we find formally that

$$\begin{aligned} \partial_t R_{U_1} &= R_{V_1}, \\ \partial_t R_{V_1} &= -\partial_\alpha \mathcal{K}_0 R_{U_1} + \mu \partial_\alpha^3 \mathcal{K}_0 R_{U_1} + \varepsilon^2 \partial_t s_2(\psi, R_{Y_2}) + \mathcal{O}(\varepsilon^3), \end{aligned}$$

where we have used the above calculations. Looking more carefully at this system and retaining all terms whose H^{s-3} norm cannot be bounded by $\|R_e\|_{\mathcal{H}_\varepsilon^s}$, we obtain

$$\partial_t R_{V_1} = -\partial_\alpha (h_0 \partial_\alpha^2 \mathcal{K}_0 R_{U_1}) - h_2 \partial_\alpha^2 R_{U_1} - \mathcal{K}_0 \partial_\alpha R_{U_1} + \varepsilon^2 \partial_t s_2(\Psi, R_{Y_2}) + \mathcal{N}_8,$$

where we have

$$\|\mathcal{N}_8\|_{H^{s-3}} \leq C(\varepsilon^3 \|R_e\|_{\mathcal{H}_\varepsilon^s} + \varepsilon^\beta \|R_e\|_{\mathcal{H}_\varepsilon^s}^2 + \varepsilon^3 C_{\text{Res}} + q(t))$$

under (HS). As in the equation for R_{X_2} , we eliminate the terms of $\mathcal{O}(\varepsilon^2)$ by a smooth near identity change of coordinates. In order to do so let

$$R_{W_1} = R_{V_1} - \varepsilon^2 s_2(\psi, R_{Y_2}).$$

We finally find

$$\partial_t R_{W_1} = \partial_\alpha (h_0 \partial_\alpha^2 \mathcal{K}_0 R_{U_1}) + h_2 \partial_\alpha^2 R_{U_1} - \mathcal{K}_0 \partial_\alpha R_{U_1} + \mathcal{N}_9,$$

where \mathcal{N}_9 obeys the same estimates as \mathcal{N}_8 .

Now collect the new variables in $\tilde{R}_e = (R_{Z_1}, R_{Y_2}, R_{U_1}, R_{W_1})$. Again for small ε the vector \tilde{R}_e can be estimated by R_e and vice versa.

With these preparations we now start to obtain the estimates for the error. In a first step, the variable R_{Z_1} is estimated in terms of the remaining variables. In the second step we construct a new scalar product $\mathbf{E}_s(\cdot, \cdot)$ for the $(R_{X_2}, R_{U_1}, R_{V_1})$ variables. It is equivalent to the usual $\mathcal{H}_l^s = H^s \times H^{s-3/2} \times H^{s-3}$ scalar product. We define $R_l = (R_{Y_2}, R_{U_1}, R_{V_1})$ the “lower” part of R_e . As explained in Remark 14 the main part of $\mathbf{E}_s(\cdot, \cdot)$ must be constructed so that in the computation of $\frac{d}{dt} \mathbf{E}_s(R_l, R_l)$ all quasilinear terms from (12) cancel and $\frac{d}{dt} \mathbf{E}_s(R_l, R_l)$ can be estimated in terms of $\mathbf{E}_s(R_l, R_l)$ in such a way that Gronwall’s inequality can be applied. On the other hand, in order to apply the ideas of Remark 5 we have to modify $\mathbf{E}_s(\cdot, \cdot)$ further so that the influence of the linear terms of order $\mathcal{O}(\varepsilon^2 \|R_l\|_{\mathcal{H}_\varepsilon^s})$, when measured in the norm defined by $\mathbf{E}_s(\cdot, \cdot)$, is of order $\mathcal{O}(\varepsilon^3)$. Finally, the (R_{Y_2}, R_{U_1}) part of the scalar product $E_s(\cdot, \cdot)$ has to be modified so that the linear terms of order $\mathcal{O}(\|R_e\|_{\mathcal{H}_\varepsilon^s})$ coming from the term \mathcal{N}_5 in the equation for $\partial_t R_{Y_2}$ cancel.

We begin by estimating R_{Z_1} in terms of R_ℓ . From the first two equations for the error it follows that

$$\partial_t R_{Y_2} - \partial_t R_{Z_1} = \mathcal{N}_{10},$$

where

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

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

$$\forall C_A, T_0 \exists C_4, C_5 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0) : \sup_{t \in [0, T_0/\varepsilon^3]} \|R_{Z_1}(t)\|_{H^s} \leq C_4 + C_5 \|R_l(t)\|_{H_l^s} \quad (40)$$

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

For $\mathcal{W}_e \in \mathcal{C}_p$ we can relate $V_1 = \partial_t U_1$ to $\mu \partial_\alpha^3 X_2 - \partial_\alpha X_2$. Thus, by the implicit function theorem we obtain

$$\mu \partial_\alpha^3 R_{X_2} - \partial_\alpha R_{X_2} = R_{V_1} + \tilde{\mathcal{N}}_{12}$$

or equivalently

$$\mu \partial_\alpha^3 R_{Y_2} - \partial_\alpha R_{Y_2} = R_{W_1} + \mathcal{N}_{12}$$

with

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

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

$$\|\partial_\alpha^s R_{Y_2}\|_{L^2} \leq \mu^{-1} (\|\partial_\alpha^{s-3} R_{W_1}\|_{L^2} + \|\partial_\alpha^{s-2} R_{Y_2}\|_{L^2} + \|\partial_\alpha^{s-3} \mathcal{N}_{12}\|_{L^2}), \quad (41)$$

so we can control $\|R_{Y_2}\|_{H^s}$ by $\|R_{Y_2}\|_{L^2}$, $\|R_{U_1}\|_{H^{s-3/2}}$, and $\|R_{W_1}\|_{H^{s-3}}$ by applying an interpolation inequality to $\|\partial_\alpha^{s-2} R_{Y_2}\|_{L^2}$.

Remark 24. Note that the preceding estimates are not uniform in μ as μ tends toward zero.

We look more closely at the term $-\varepsilon^2 \partial_\alpha s_1(\psi, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2})$ which appears on the right-hand side of the equation for $\partial_t R_{U_1}$ before we start to construct the final energy. It can be rewritten as

$$\begin{aligned} & \partial_\alpha s_1(\Psi, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \\ &= \partial_\alpha \{ \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) - \psi_{X_2} \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} - \mathcal{K}_0(\psi_{X_2} \mathcal{K}_0) \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} \} \\ &= \partial_\alpha \{ \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \\ & \quad - (1 + \mathcal{K}_0^2)(\psi_{X_2} \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) + \mathcal{K}_0[\mathcal{K}_0, \psi_{X_2}] \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} \}. \end{aligned}$$

The last term can be estimated to be of order $\mathcal{O}(\varepsilon)$ (considered as an element of H^{s-1}) by using the following lemma which shows that commutators with the scaled function Ψ gain an order of ε .

Lemma 9. *Let $R \in H^2$ and let $\Psi \in H^{s+1}$ be the approximation defined above. Then there exists a $C > 0$ such that, for all $\varepsilon \in (0, 1)$,*

$$\|[\mathcal{K}_0, \Psi]R\|_{H^s} \leq C\varepsilon \|\partial_X \Psi\|_{H^s} \|R\|_{H^2}.$$

Proof. See [SW00a, Section 6; e.g., (6.8)]. \square

By the preceding estimates it remains to estimate $\|R_{Y_2}\|_{L^2}$, $\|R_{U_1}\|_{H^{s-3/2}}$ and $\|R_{V_1}\|_{H^{s-3}}$. After these preliminary simplifications we come to the final **energy estimates**.

Note that the operator $\mathcal{M}_2 = -\partial_\alpha \mathcal{K}_0^{-1}$ is self-adjoint and positive. Thus we can take its square root and we find

$$\begin{aligned} \partial_t \int (\mathcal{M}_2^{1/2} R_{Y_2})^2 d\alpha / 2 &= \int R_{Y_2} (\mathcal{M}_2 \partial_t R_{Y_2}) d\alpha \\ &= \int R_{Y_2} (\mathcal{M}_2 (\mathcal{K}_0 R_{U_1} + \mathcal{N}_5)) d\alpha \\ &= \int (\partial_\alpha R_{Y_2}) R_{U_1} d\alpha + \mathcal{N}_{e1}, \end{aligned}$$

where

$$|\mathcal{N}_{e1}| \leq C(\varepsilon^3 \|R_t\|_{\mathcal{H}_l^s}^2 + \varepsilon^\beta \|R_t\|_{\mathcal{H}_l^s}^3 + \varepsilon^3 C_{res} \|R_t\|_{\mathcal{H}_l^s} + q(t) \|R_t\|_{\mathcal{H}_l^s})$$

under (HS). Similarly we find that, for the positive self-adjoint operator $\mathcal{M}_3 = \mu \mathcal{K}_0^{-1} \partial_\alpha^3$,

$$\begin{aligned} \partial_t \int (\mathcal{M}_3^{1/2} R_{Y_2})^2 d\alpha / 2 &= \int R_{Y_2} (\mathcal{M}_3 \partial_t R_{Y_2}) d\alpha \\ &= \int R_{Y_2} (\mathcal{M}_3 (\mathcal{K}_0 R_{U_1} + \mathcal{N}_5)) d\alpha \\ &= - \int (\mu \partial_\alpha^3 R_{Y_2}) R_{U_1} d\alpha + \mathcal{N}_{e2}, \end{aligned}$$

where \mathcal{N}_{e2} obeys the same estimates as \mathcal{N}_{e1} .

Next using the evolution equation for U_1 we see that:

$$\begin{aligned} \partial_t \int (R_{U_1})^2 d\alpha / 2 &= \int R_{U_1} (\partial_t R_{U_1}) d\alpha \\ &= \int R_{U_1} (-\partial_\alpha R_{Y_2} + \mu \partial_\alpha^3 R_{Y_2} - \varepsilon^2 \psi_{Z_1} \partial_\alpha R_{Y_2} \\ &\quad + 2\mu \varepsilon^2 \psi_{Z_1} \partial_\alpha^3 R_{Y_2} + \varepsilon^2 \partial_\alpha \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \\ &\quad - \varepsilon^2 \partial_\alpha (1 + \mathcal{K}_0^2)(\psi_{X_2} (\mathcal{K}_0^{-1} \partial_\alpha R_{Y_2})) + \tilde{\mathcal{N}}_4) d\alpha, \end{aligned}$$

where we have used $\psi_{Z_1} = \psi_{X_2} + \mathcal{O}(\varepsilon)$ and where $\tilde{\mathcal{N}}_4$ equals \mathcal{N}_4 plus the order epsilon terms that we discarded from $\partial_\alpha s_1(\psi, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2})$ in our preliminary investigations. The term $\int R_{U_1} \tilde{\mathcal{N}}_4 d\alpha$ obeys the same estimates as \mathcal{N}_{e1} .

The first two terms on the right-hand side cancel with the corresponding term in the time derivatives of $\int (\mathcal{M}_2^{1/2} R_{Y_2})^2 d\alpha$ and $\int (\mathcal{M}_3^{1/2} R_{Y_2})^2 d\alpha$.

For the third term, by using $\partial_\alpha \psi_{Z_1} = \mathcal{O}(\varepsilon)$, we obtain

$$-\varepsilon^2 \int R_{U_1} \psi_{Z_1} \partial_\alpha R_{Y_2} d\alpha = \varepsilon^2 \int (\partial_\alpha R_{U_1}) \psi_{Z_1} R_{Y_2} d\alpha + \mathcal{O}(\varepsilon^3).$$

With $\partial_t R_{Y_2} = \mathcal{K}_0 R_{U_2} + \mathcal{O}(\varepsilon)$, this term can be rewritten

$$\varepsilon^2 \int \partial_t (\partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \psi_{Z_1} R_{Y_2} d\alpha + \mathcal{O}(\varepsilon^3).$$

The operator $\mathcal{M}_3^2 = -\partial_\alpha \mathcal{K}_0^{-1}$ is positive and self-adjoint. Hence

$$\begin{aligned} \varepsilon^2 \int \partial_t (\mathcal{M}_3^2 R_{Y_2}) \psi_{Z_1} R_{Y_2} d\alpha &= \varepsilon^2 \int \partial_t (\mathcal{M}_3 R_{Y_2}) [\mathcal{M}_3, \psi_{Z_1}] R_{Y_2} d\alpha \\ &\quad + \varepsilon^2 \int \partial_t (\mathcal{M}_3 R_{Y_2}) \psi_{Z_1} (\mathcal{M}_3 R_{Y_2}) d\alpha \\ &= \frac{1}{2} \varepsilon^2 \partial_t \int \psi_{Z_1} (\mathcal{M}_3 R_{Y_2})^2 d\alpha + \mathcal{O}(\varepsilon^3), \end{aligned}$$

where we again exploited the fact that commutators with the approximation Ψ gain one order of ε and where we used $\partial_t \psi_{Z_1} = \mathcal{O}(\varepsilon)$. In a similar fashion we obtain, for the fourth term,

$$2\varepsilon^2 \mu \int R_{U_1} \psi_{Z_1} \partial_\alpha^3 R_{Y_2} d\alpha = -\varepsilon^2 \partial_t \int \psi_{Z_1} (\mathcal{M}_3 \partial_\alpha R_{Y_2})^2 d\alpha + \mathcal{N}_{e2}.$$

For the fifth term, by partial integration and by using $\partial_t R_{Y_2} = \mathcal{K}_0 R_{U_1} + \mathcal{O}(\varepsilon)$, we obtain

$$\begin{aligned} -\varepsilon^2 \int (\partial_\alpha R_{U_1}) \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) d\alpha \\ = -\varepsilon^2 \int (\partial_t \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) d\alpha + \mathcal{O}(\varepsilon^3). \end{aligned}$$

The last expression defines a trilinear mapping $\mathcal{M}_5(\Psi)[R_1, R_2]$ which is given in Fourier space by

$$\mathcal{M}_5(\widehat{\Psi})[\widehat{R}_1, R_2] = \iint \widehat{R}_1(k) \frac{\widehat{\mathcal{K}}_0(l) - \widehat{\mathcal{K}}_0(k)}{\widehat{\mathcal{K}}_0(l-k)} \widehat{\psi}_{Z_1}(k-l) \widehat{R}_2(l) dl dk + \text{c.c.}$$

This operator is symmetric in the arguments R_1 and R_2 . Therefore, we finally obtain

$$\begin{aligned} \varepsilon^2 \int R_{U_1} \partial_\alpha \mathcal{M}_1(\psi_{Z_1}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) d\alpha \\ = \frac{1}{2} \varepsilon^2 \partial_t \mathcal{M}_5(\Psi)[\partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}] + \mathcal{N}_{e3}. \end{aligned}$$

Finally, we again use

$$\partial_t R_{Y_2} = \mathcal{K}_0 R_{U_2} + \mathcal{O}(\varepsilon)$$

to rewrite the sixth term as as

$$\begin{aligned}
 & -\varepsilon^2 \int R_{U_1} \partial_\alpha (1 + \mathcal{K}_0^2) (\psi_{X_2} (\mathcal{K}_0^{-1} \partial_\alpha R_{Y_2})) d\alpha \\
 & = \varepsilon^2 \int \partial_t (\partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) (1 + \mathcal{K}_0^2) (\psi_{X_2} (\mathcal{K}_0^{-1} \partial_\alpha R_{Y_2})) d\alpha + \mathcal{O}(\varepsilon^3).
 \end{aligned}$$

Since $\mathcal{M}_4^2 = (1 + \mathcal{K}_0^2)$ is again a positive self-adjoint operator, and using the fact that commutators with Ψ gain an additional order of ε , we again have

$$\begin{aligned}
 & \int (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \mathcal{M}_4 (\psi_{X_2} \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) d\alpha \\
 & = \int (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{U_1}) \psi_{X_2} \mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} d\alpha \\
 & \quad + \int (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{U_1}) [\mathcal{M}_4, \psi_{X_2}] \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} d\alpha \\
 & = \int (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}) \psi_{X_2} \mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2} d\alpha + \mathcal{O}(\varepsilon).
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 & -\varepsilon^2 \int R_{U_1} \partial_\alpha (1 + \mathcal{K}_0^2) (\psi_{X_2} (\mathcal{K}_0^{-1} \partial_\alpha R_{Y_2})) d\alpha \\
 & = \varepsilon^2 \partial_t \int \psi_{X_2} (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2})^2 d\alpha + \mathcal{N}_{e4}.
 \end{aligned}$$

Furthermore, under (HS), the remainder terms \mathcal{N}_{ej} satisfy the estimates

$$|\mathcal{N}_{ej}| \leq C(\varepsilon^3 \|R_l\|_{\mathcal{H}_l^s}^2 + \varepsilon^\beta \|R_l\|_{\mathcal{H}_l^s}^3 + (\varepsilon^3 C_{\text{Res}} + q(t)) \|R_l\|_{\mathcal{H}_l^s}). \tag{42}$$

To reiterate, the point of these definitions is that the energy functional must be chosen so that both the quasilinear terms and the terms of $\mathcal{O}(\varepsilon^2)$ in the time derivative cancel, and by adding to the energy the terms of order ε^2 on the right-hand side of the above expressions we can arrange for this to be the case.

For the quasilinear terms, we proceed as in the proof of Theorem 4. We introduce $R_u = \lambda_1 R_{U_1}$ and $R_w = \lambda_1 R_{W_1}$ and then we consider

$$\begin{aligned}
 & \frac{1}{2} \frac{d}{dt} ((\Lambda^r R_v, \Lambda^r R_v)_{L^2} + (\Lambda^{r+1/2} R_u, \Lambda^{r+1/2} R_u)_{L^2}) \\
 & = (\Lambda^r R_v, \Lambda^r \lambda_1 \mathcal{N}_9)_{L^2} + (\Lambda^r R_v, \partial_t (\Lambda^r) R_v)_{L^2} \\
 & \quad + (\Lambda^{r+1/2} R_u, \partial_t (\Lambda^{r+1/2}) R_u)_{L^2},
 \end{aligned}$$

where as in Section 2, $3r = s - \frac{7}{2}$. The estimates for \mathcal{N}_9 can be found above – in particular, we have

$$|(\Lambda^r R_v, \Lambda^r \lambda_1 \mathcal{N}_9)_{L^2}| \leq C(\varepsilon^3 \|R_l\|_{\mathcal{H}_l^s}^2 + \varepsilon^\beta \|R_l\|_{\mathcal{H}_l^s}^3 + (\varepsilon^3 C_{\text{Res}} + q(t)) \|R_l\|_{\mathcal{H}_l^s}).$$

Thus it remains only to consider the behavior of $\partial_t \Lambda^r$ and $\partial_t \Lambda^{r+1/2}$. While the estimates of Section 2 insure that these terms are bounded, we must also determine

how they depend on ε . The terms $\partial_t h_0$ and $\partial_t h_2$ which appear in these expressions will be combinations of expressions of the form $(\varepsilon \Psi_e + \varepsilon^\beta R_e)$. If a time derivative falls on Ψ_e , we get a factor of ε^3 due to the scaling of ψ_e with ε^2 and the fact that time derivatives of ψ_e give an additional ε . On the other hand, if the derivative falls on R_e , we get a factor of ε^β , with $\beta > 3$. Therefore $\partial_t(\Lambda^{r+1/2})$ and $\partial_t(\Lambda^r)$ obey the same estimates as \mathcal{N}_9 .

With all these preliminary considerations out of the way the calculation of the time derivative proceeds rapidly. We define

$$\begin{aligned} \mathbf{E}_s(R_l, R_l) &= \int (\mathcal{M}_2^{1/2} R_{Y_2})^2 + \int (\mathcal{M}_3^{1/2} R_{Y_2})^2 + \int (R_{U_1})^2 \\ &\quad + \varepsilon^2 \mathcal{M}_5(\Psi)[\partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}, \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2}] \\ &\quad + \int \psi_{Z_1} (\mathcal{M}_3 R_{Y_2})^2 d\alpha + 2\mu \int \psi_{Z_1} (\partial_\alpha \mathcal{M}_3 R_{Y_2})^2 d\alpha \\ &\quad + \varepsilon^2 \int \psi_{X_2} (\mathcal{M}_4 \partial_\alpha \mathcal{K}_0^{-1} R_{Y_2})^2 d\alpha \\ &\quad + (\Lambda^r R_v, \Lambda^r R_v)_{L^2} + (\Lambda^{r+1/2} R_u, \Lambda^{r+1/2} R_u)_{L^2}. \end{aligned}$$

We easily see by using (40) and (41) that the scalar product $\mathbf{E}_s(\cdot, \cdot)$ is equivalent to the usual \mathcal{H}_l^s scalar product, i.e., there exist positive constants c_1 and c_2 and an $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ we have

$$\|\mathcal{W}_l\|_{\mathcal{H}_l^s}^2 \leq c_1 \mathbf{E}_s(\mathcal{W}_l, \mathcal{W}_l) \leq c_2 \|\mathcal{W}_l\|_{\mathcal{H}_l^s}^2. \tag{43}$$

Therefore, we can follow line for line [SW00a] and sum up our above estimates. We see that there exist constants $C_1 = C_1(C_\Psi, C_{\text{Res}}, c_j)$, $C_2 = C_2(C_\Psi, C_{\text{Res}}, C_R, c_j)$, and $C_3 = C_3(C_\Psi, C_{\text{Res}}, c_j)$, such that

$$\begin{aligned} \frac{1}{2} \partial_t \mathbf{E}_s(R_l, R_l) &\leq \varepsilon^3 C_1 \mathbf{E}_s(R_l, R_l) + \varepsilon^\beta C_2 \mathbf{E}_s(R_l, R_l)^{3/2} \\ &\quad + (\varepsilon^3 C_3 + q(t)) \mathbf{E}_s(R_l, R_l)^{1/2} \\ &\leq \varepsilon^3 C_1 \mathbf{E}_s(R_l, R_l) + \varepsilon^\beta C_2 \mathbf{E}_s(R_l, R_l)^{3/2} \\ &\quad + (C_3 \varepsilon^3 + q(t)) + (\varepsilon^3 C_3 + q(t)) \mathbf{E}_s(R_l, R_l), \end{aligned}$$

Thus, for $y(T) = \mathbf{E}_s(R_l(t), R_l(t))$ with $T = \varepsilon^3 t$ we obtain the differential inequality

$$\begin{aligned} \dot{y} &= (C_1 + C_3 + \varepsilon^{-3} q(t))y + \varepsilon^\beta C_2 y^{3/2} + C_3 + \varepsilon^{-3} q(t) \\ &\leq (C_1 + C_3 + 1 + \varepsilon^{-3} q(t))y + C_3 + \varepsilon^{-3} q(t). \end{aligned}$$

The application of Gronwall’s inequality shows, for all $t \in [0, T_0/\varepsilon^3]$, that

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

where we have chosen $\varepsilon_0 > 0$ so small that $\varepsilon^{\beta-3} C_2(C_R)(c_2 C_R)^2 \leq 1$, where $C_R = 2 \limsup_{\varepsilon \rightarrow 0} y_\varepsilon(T_0) = \mathcal{O}(1)$. With (38), (40), (41), and (43), this yields

$$\sup_{t \in [0, T_0/\varepsilon^3]} \|R_\varepsilon\|_{\mathcal{H}_t^\varepsilon}^2 \leq c_1 C_1.$$

This completes the proof. \square

Remark 25. With the ansatz $\mathcal{W}_\varepsilon = \varepsilon^4 \psi + \varepsilon^\beta R$, where $\beta = \frac{11}{2}$, the error estimates for the approximation by the Kawahara equations follow more or less line for line the proofs of Theorem 5, and Theorem 2 follows. Again the details are left to the reader.

Acknowledgements. We are grateful for discussions with WALTER CRAIG. G. S. would like to thank the Mathematics Department of Boston University and E. W. would like to thank the Mathematical Institute of the University of Stuttgart for their kind hospitality. E. W. thanks Robert Pierce for discussions about water waves. The work of E. W. was supported in part by the NSF through the grant DMS-9803164.

References

- [BBM72] T. B. BENJAMIN, J. BONA & J. J. MAHONY. Model equations for long waves in nonlinear dispersive systems. *Philos. Trans. Roy. Soc. London Ser. A* **272**, (1972) 47–78.
- [Be82] T. B. BENJAMIN. The solitary wave with surface tension *Quart. Appl. Math* **40**, (1982/83) 231–234.
- [Bo1877] M. J. BOUSSINESQ. Essai sur la théorie des eaux courantes. *Mémoires présentés par divers savants à l'Académie des Sciences Inst. France (séries 2)* **23**, (1877) 1–680.
- [BGT96] B. BUFFONI, M. D. GROVES & J. TOLAND. A plethora of solitary gravity-capillary water waves with nearly critical Bond and Froude numbers. *Philos. Trans. Roy. Soc. London* **354**, (1996) 575–607.
- [BG99] B. BUFFONI & M. D. GROVES. A multiplicity result for solitary gravity-capillary waves in deep water via critical-point theory. *Arch. Rational Mech. Anal.* **146**, (1999) 183–220.
- [BS71] J. G. B. BYATT-SMITH. An integral equation for unsteady surface waves and a comment on the Boussinesq equation. *J. Fluid Mech.* **49**, (1971) 625–633.
- [Cr85] W. CRAIG. An existence theory for water waves and the Boussinesq and Korteweg–de Vries scaling limits. *Commun. Part. Diff. Eqns.* **10**, (1985) 787–1003.
- [DS71] N. DUNFORD & J. T. SCHWARTZ. *Linear operators*. Part I–III. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. Reprint of the 1958–1971 original.
- [Il90] A. T. IL'ICHEV. Theory of nonlinear waves described by fifth-order evolution equations. *Fluid Dynam.* (translated from *Izv. Akad. Nauk SSSR Mekh. Zhidk. Gaza*) **25**, (1990) 247–252.
- [Kal89] L. A. KALYAKIN. Long wave asymptotics. Integrable equations as asymptotic limits of non-linear systems. *Russian Math. Surveys* **44**, (1989) 3–42.
- [KN86] T. KANO & T. NISHIDA. A mathematical justification for Korteweg–de Vries equation and Boussinesq equation of water surface waves. *Osaka J. Math.* **23**, (1986) 389–413.
- [Kat75] T. KATO. *Quasi-linear equations of evolution with applications to partial differential equations*. LNM 448, Springer 1974, pp. 25–70.
- [Li78] J. LIGHTHILL. *Waves in Fluids*. Cambridge University Press (1978).

- [Lo64] R. R. LONG. The initial–value problem for long waves of finite amplitude. *J. Fluid. Mech.* **20**, (1964) 161–170.
- [Ma88] A. V. MARCHENKO. Long waves in shallow liquid under ice cover. *J. Appl. Math. Mech.* (translated from Prikl. Mat. Mekh.) **52**, (1988) 180–183.
- [Ma91] F. MATTIOLI. Decomposition of the Boussinesq equations for shallow water waves into a set of coupled KdV equations. *Physics of Fluids A. Fluid Dynamics* **3**, (1991) 2355–2359.
- [Me79] R. E. MEYER. On the meaning of the Boussinesq and Korteweg–de Vries equations. *Bull. Calcutta Math. Soc.* **71**, (1979) 121–130.
- [Na74] V. I. NALIMOV. The Cauchy–Poisson problem. (In Russian). *Dynamika Splosh. Sredy* **18**, (1974) 104–210.
- [Sh76] M. SHINBROT. The initial value problem for surface waves under gravity, I. The simplest case. *Indiana U. Math. J.* **25**, (1976) 281–300.
- [Sch98] G. SCHNEIDER. The long wave limit for a Boussinesq equation. *SIAM Journal of Applied Mathematics* **58**, (1998) 1237–1245.
- [SW00a] G. SCHNEIDER & C. E. WAYNE. The long wave limit for the water wave problem. I. the case of zero surface tension. *Comm. Pure Appl. Math.* **53**, (2000) 1475–1535.
- [SW00b] G. SCHNEIDER & C. E. WAYNE. Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In *Proceeding of Equadiff '99; International Conference on Differential Equations*, vol. 1, eds. B. FIEDLER, K. GRÖGER & J. SPREKELS, World Scientific (2000), pp. 390–404.
- [SW00c] G. SCHNEIDER & C. E. WAYNE. Kawahara dynamics in dispersive media. *Physica D* **152/153**, (2001) 384–394.
- [Wu97] SIJUE WU. Well-posedness in Sobolev spaces of the full water wave problem in 2D. *Invent. Math.* **130**, (1997) 39–72.
- [Wu99] SIJUE WU. Well-posedness in Sobolev spaces of the full water wave problem in 3D. *Journal of the AMS* **12**, (1999) 445–495.
- [Yo82] H. YOSIHARA. Gravity waves on the free surface of an incompressible perfect fluid of finite depth. *RIMS Kyoto* **18**, (1982) 49–96.
- [Yo83] H. YOSIHARA. Capillary–gravity waves for an incompressible ideal fluid. *J. Math. Kyoto Univ.* **23**, (1983) 649–694.

Mathematisches Institut
 Universität Bayreuth
 95440 Bayreuth, Germany
 e-mail: guido.schneider@uni-bayreuth.de

and

Department of Mathematics
 and Center for BioDynamics
 Boston University
 111 Cummington St.
 Boston, MA 02215, USA
 e-mail: cew@math.bu.edu

(Accepted October 25, 2001)

Published online April 11, 2002 – © Springer-Verlag (2002)