Euro. Jnl of Applied Mathematics (2003), *vol.* 14, *pp.* 547–570.
DOI: 10.1017/S0956792503005163 Printed in the United Kingdom

Estimates for the three-wave interaction of surface water waves

GUIDO SCHNEIDER[1] and C. EUGENE WAYNE[2]

[1]*Mathematisches Institut I, Universität Karlsruhe, D-76128 Karlsruhe, Germany*
email: guido.schneider@mathematik.uni-karlsruhe.de
[2]*Department of Mathematics and Center for BioDynamics, Boston University, Cummington St., Boston, MA 02215, USA*
email: cew@math.bu.edu

(*Received 22 March 2002; revised 2 January 2003*)

The equations for three-wave interaction describe the resonant, quadratic, nonlinear interaction of three waves. They are obtained as amplitude equations in an asymptotic reduction of the basic equations of nonlinear optics, fluid mechanics, and plasma physics. These equations are completely integrable and have been the subject of intensive research in the last years. It is the purpose of this paper to prove exact estimates between the approximations obtained via this system and solutions of the original physical system. Although the three-wave interaction model is believed to describe a number of different physical models we restrict attention to its application as a model of the resonant interaction of water waves subject to weak surface tension.

1 Introduction

In this paper we continue our program to justify the use of common models for water waves. In previous work we have examined the equations appropriate to approximate the long-wavelength motion of gravity waves [21], and capillary-gravity waves [22]. In this work we consider the interaction of three wave trains whose frequencies and wave numbers are in resonance. The equations for the so called Three-Wave Interaction (TWI) are given by

$$\begin{aligned}\partial_T A_1 &= c_g(k_1)\partial_X A_1 + i\gamma_1\overline{A_2A_3},\\ \partial_T A_2 &= c_g(k_2)\partial_X A_2 + i\gamma_2\overline{A_1A_3},\\ \partial_T A_3 &= c_g(k_3)\partial_X A_3 + i\gamma_3\overline{A_1A_2},\end{aligned} \tag{1.1}$$

with $T \in \mathbb{R}$, $X \in \mathbb{R}$, $c_g(k_j) \in \mathbb{R}$, $\gamma_j \in \mathbb{R}$, $k_j \in \mathbb{R}$, and $A_j(X,T) \in \mathbb{C}$. The TWI equations describe the resonant quadratic nonlinear interaction of three wave packets modulating some underlying wave trains $e^{i(k_jx+\omega_jt)}$. These wave trains have to satisfy the resonance condition

$$\left.\begin{aligned}k_1+k_2+k_3&=0\\ \omega_1+\omega_2+\omega_3&=0\end{aligned}\right\}, \tag{1.2}$$

with $k_j \in \mathbb{R}$ the spatial wave numbers and ω_j the temporal wave numbers. It is the purpose of this paper to show that the dynamics in the original system behave as predicted by the equations for the three-wave interaction, i.e. to prove exact estimates between the solutions of the water wave problem and the approximations obtained by the equations for the three-wave interaction. In the case of surface water waves moving under the influence of gravity and weak surface tension, the system for the three-wave interaction has been considered previously [3, 6, 17]. For a relatively recent survey of both the theoretical and experimental status of resonant interaction models in the theory of water waves, see Hammack & Henderson [9].

The quadratic resonant three-wave interaction seems to play a big role in the generation of capillary-gravity surface waves. For instance, Janssen [10, 11] argues that three-wave interactions may be important in the wind-induced generation of initial wavelets comparing the formal approximations with experimental data [14]. In the resonant case, a third wave packet of order $\mathcal{O}(\varepsilon)$ can be generated out of two wave packets of order $\mathcal{O}(\varepsilon)$. In the non-resonant case only the generation of order $\mathcal{O}(\varepsilon^2)$ wave packets is possible. This idea goes back to the beginning of the 1960s [16, 19], based on observations at the beginning of the last century, and has been reviewed, for instance, in Hammack & Henderson [9].

Just as with the Korteweg-de Vries or Nonlinear Schrödinger equations, the TWI equations can be derived as an asymptotic approximation for a number of systems, including nonlinear optics, plasma physics, and fluid mechanics, especially for internal waves in two layer fluids, and for surface water waves in case of low surface tension. See Ablowitz & Segur [2] and Craik [8] and the references therein. Although we restrict ourselves to the water wave problem due to our particular interest, we believe that the approximation property holds for the other original systems, too. In fact, we expect that the analogous theorems may be easier to prove in those cases, because the local existence and uniqueness theory is simpler.

Equations (1.1) turned out to be completely integrable, and they have been the subject of intensive research in the last few years (see elsewhere [4, 5, 13, 1] for recent developments in the x-independent case). Unlike the KdV equations, for example, the complete integrability of (1.1) cannot be reduced to an integral equation of Gelfand–Levitan type. Instead, one must solve a Riemann-Hilbert problem.

The water wave problem consists in finding the irrotational flow of an inviscid, incompressible fluid in an infinitely long canal of fixed depth h with impermeable bottom under the influence of gravity and surface tension. We choose coordinates with $x \in \mathbb{R}$ denoting the unbounded direction in the fluid and with $y \geqslant -1$ the coordinate measuring the fluid's (finite) depth. The velocity field of the fluid satisfies Euler's equations and the fluid fills the unknown time-dependent domain $\Omega(t)$ between the bottom $\{(x, -1) | x \in \mathbb{R}\}$ and the free unknown top surface $\Gamma(t) = \{(x, \eta(x,t)) | x \in \mathbb{R}\}$ with $\eta : \mathbb{R} \to \mathbb{R}$. The TWI equations are approximation to water waves that are small perturbations of a flat surface. As such, we assume throughout this paper that the fluid surface can be written as the graph of some function. It follows from our results that if the initial conditions are of the form described by the hypothesis of Theorem 1.1, then the surface can be written as a graph at least over the period when the approximation is valid.

Under these assumptions, there exists a potential $\phi : \Omega(t) \to \mathbb{R}$ such that the velocity field $\underline{u} = (u, v)$ satisfies $u = \partial_x \phi$ and $v = \partial_y \phi$. The potential ϕ and the elevation η of the top surface satisfy

$$\partial_x^2\phi + \partial_y^2\phi = 0, \quad \text{in} \quad \Omega(t), \tag{1.3}$$

$$\partial_y\phi = 0, \quad \text{for} \quad y = -1, \tag{1.4}$$

$$\partial_t\eta = \partial_y\phi - (\partial_x\eta)\partial_x\phi, \quad \text{on} \quad \Gamma(t), \tag{1.5}$$

$$\partial_t\phi = -\frac{1}{2}((\partial_x\phi)^2 + (\partial_y\phi)^2) + \mu\partial_x\left[\frac{\partial_x\eta}{\sqrt{1+(\partial_x\eta)^2}}\right] - \eta, \quad \text{on} \quad \Gamma(t), \tag{1.6}$$

where μ is a parameter proportional to surface tension. Without loss of generality, we set the depth of the fluid and the gravitational constant to one. It is well known, and it will be explained in the next section, that the water wave problem is completely described by the evolution of the elevation of the top surface $\eta = \eta(x,t)$ and the horizontal velocity component $w = w(x,t) = u(x,\eta(x,t),t)$ at the top surface.

Then the equations (1.1) for the three-wave interaction can be derived by the following multiple scaling ansatz. There exist vectors $\varphi_j \in \mathbb{C}^2$ (which depend only upon k_j and can be computed explicitly) such that

$$\begin{pmatrix} w \\ \eta \end{pmatrix} \approx \varepsilon\psi_1(x,t) = \varepsilon A_1(\varepsilon x, \varepsilon t)e^{i(k_1x+\omega_1t)}\varphi_1 + \varepsilon A_2(\varepsilon x, \varepsilon t)e^{i(k_2x+\omega_2t)}\varphi_2 + \varepsilon A_3(\varepsilon x, \varepsilon t)e^{i(k_3x+\omega_3t)}\varphi_3 + c.c.$$

with $0 < \varepsilon \ll 1$ a small parameter, c.c. meaning complex conjugate, and the spatial and temporal wavenumbers k_j and ω_j having to satisfy the resonance condition (1.2), and being related by the linear dispersion relation of the water wave problem

$$\omega^2 = (k + \mu k^3)\tanh k. \tag{1.7}$$

Notation. We denote Fourier transform by $(\mathscr{F}u)(k) = \hat{u}(k) = \frac{1}{2\pi}\int u(x)e^{-ikx}\,dx$. 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 $\|u\|_{C_b^n} = \sum_{j=0}^n \|\partial_x^j u\|_{C_b^0}$, where $\|u\|_{C_b^0} = \sup_{x\in\mathbb{R}} |u(x)|$.

Then our result is as follows.

Theorem 1.1 *Let $s \geqslant 6$ and choose ω_j and k_j to satisfy* (1.2) *and* (1.7)*, 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 $A_1, A_2, A_3 \in C([0,T_0], (H^{s+2}(\mathbb{R},\mathbb{C}))^3)$ be solutions of (1.1) *with*

$$\sup_{T\in[0,T_0]} \|A_j(T)\|_{H^{s+2}} \leqslant C_1$$

for $j = 1,2,3$. Then there are solutions of the water wave problem (1.3)–(1.6) *satisfying*

$$\sup_{t\in[0,\frac{T_0}{\varepsilon}]} \left\| \begin{pmatrix} w \\ \eta \end{pmatrix}(x,t) - \varepsilon\psi_1(x,t) \right\|_{(C_b^{s-2})^2} \leqslant C_2\varepsilon^{3/2}.$$

This theorem also allows us to find the dynamics of the equations for the three-wave interaction in the water wave problem, since the error of order $\mathcal{O}(\varepsilon^{3/2})$ is small compared with the solution and approximation, which are both of order $\mathcal{O}(\varepsilon)$.

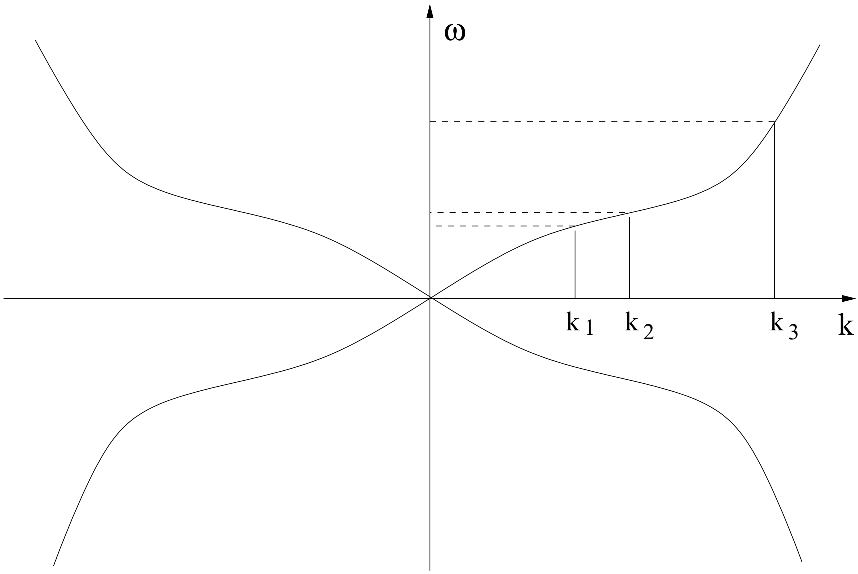

FIGURE 1. Schematic of the linear dispersion relation and possible quadratic resonances.

It is easy to see that the resonance condition (1.2) can be satisfied for the dispersion relation (1.7) of the water wave problem for small $\mu > 0$, i.e. for small surface tension (see Figure 1).

In principle, the proof of Theorem 1.1 is based on a simple application of Gronwall's inequality. It has been pointed out [15] that such an estimate can be proved on a time scale $\mathcal{O}(\frac{1}{\varepsilon})$ if the solutions are of order $\mathcal{O}(\varepsilon)$, and if the residual, i.e. the terms which remain after inserting the ansatz into the equations of the original system, can be made sufficiently small, here $\mathcal{O}(\varepsilon^3)$. In the present case, we make the residual small by adding to the original ansatz $\varepsilon\psi_1$ additional terms that are of higher order in ε. That these terms can be chosen is such a way as to make the residual small is proven in Lemma 3.6. Therefore, the equations for the three wave interaction (1.1) do provide accurate and useful approximations of the water wave problem. We note that this fact should not be taken for granted. There are modulation equations (some examples of which are described in Schneider [20]) which, although derived by reasonable formal arguments, do not reflect the true dynamics of the original equations.

The difficulties in proving Theorem 1.1 come largely from the inherent difficulties of the water wave problem, and not from the approximation procedure. So far there is no direct local existence and uniqueness theory for the Eulerian formulation (1.3)–(1.6) of the water wave problem, which does not work with analytic initial conditions. Therefore, to prove Theorem 1.1 we have to work with the Lagrangian formulation of the water wave problem for which a number of local existence and uniqueness theorems in Sobolev spaces exist [7, 18, 24, 25, 26, 27]. This formulation will be introduced in the next section. It has the drawback that several of the variables in this formulation exhibit secular growth, which is difficult to control over the long time scales we must work with. To circumvent this secular growth, we apply a method from Schneider & Wayne [22] where the KdV-equation has been justified as an amplitude equation for the water wave problem with weak surface tension.

Notation. Throughout this paper, we assume $0 < \varepsilon \ll 1$ and denote possibly different constants by the same symbol C. 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 The Lagrangian formulation of the water wave problem

As mentioned in the introduction, there is no existence and uniqueness theory for the Eulerian formulation (1.3)–(1.6) of the water wave problem in Sobolev spaces. Thus, this section is devoted to introducing the Lagrangian formulation of the water wave problem, rewriting this formulation as a quasi-linear system of partial differential equations, and then stating an existence theorem for solutions of this system that we proved in Schneider & Wayne [22]. For fixed time t the free surface of the fluid can be written as

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

It is a Jordan curve which has no intersection with the bottom $\{(\alpha, -1) | \alpha \in \mathbb{R}\}$. In contrast to the Eulerian formulation in the Lagrangian formulation, $\Gamma(t)$ does not have to be a graph over the bottom. Under the assumptions on the flow which we made in the introduction the dynamics of the water problem is completely determined by the evolution of the free surface $\Gamma(t)$, which is governed by (for a careful derivation of the following system of equations, see Yosihara [27]):

$$\partial_t^2 X_1 (1 + \partial_\alpha X_1) + \partial_\alpha X_2 \left(1 + \partial_t^2 X_2\right) = \mu R\left(\partial_\alpha X, \partial_\alpha^2 X\right) + \mu S\left(\partial_\alpha X, \partial_\alpha^3 X\right), \tag{2.1}$$

$$\partial_t X_2 = \mathscr{K}(X) \partial_t X_1, \tag{2.2}$$

where

$$\begin{aligned} X(\alpha, t) &= (X_1(\alpha, t), X_2(\alpha, t)), \\ R\left(\partial_\alpha X, \partial_\alpha^2 X\right) &= -3Q(\partial_\alpha X)^{-5}\left((1 + \partial_\alpha X_1)\partial_\alpha^2 X_1 + (\partial_\alpha X_2)\left(\partial_\alpha^2 X_2\right)\right) \\ &\quad \times \left(-\partial_\alpha X_2 \partial_\alpha^2 X_1 + (1 + \partial_\alpha X_1)\partial_\alpha^2 X_2\right), \\ S\left(\partial_\alpha X, \partial_\alpha^3 X\right) &= Q(\partial_\alpha X)^{-3}\left(-\partial_\alpha X_2 \partial_\alpha^3 X_1 + (1 + \partial_\alpha X_1)\partial_\alpha^3 X_2\right), \\ Q(\partial_\alpha X) &= ((1 + \partial_\alpha X_1)^2 + (\partial_\alpha X_2)^2)^{1/2}. \end{aligned}$$

The operator $\mathscr{K}(X)$ acts linearly on $U_1 = \partial_t X_1$, but depends nonlinearly on X. It is related to the Dirichlet-Neumann operator and its existence is a consequence of the incompressibility and irrotationality of the flow. It is defined by $\mathscr{K}(X)U_1 = \partial_y \phi|_{\Gamma(t)}$, where $\phi : \Omega(t) \to \mathbb{R}$ solves (for fixed t) the boundary value problem

$$\begin{aligned} \Delta\phi &= 0, && \text{in } \ \Omega(t), \\ \partial_y \phi &= 0, && \text{for } y = -1, \\ \partial_x \phi &= U_1, && \text{on } \ \Gamma(t). \end{aligned}$$

The operator $\mathscr{K}(X)$ is of the form $\mathscr{K}(X) = \mathscr{K}_0 + \mathscr{S}_1(X)$, where $\mathscr{K}_0$ is the linear part of the operator $\mathscr{K}(X)$, and has the Fourier symbol $\hat{\mathscr{K}}_0(k) = -i \tanh(k)$. The nonlinear part $\mathscr{S}_1(X)\cdot$ has certain smoothing properties which are summarized in Appendix A.

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

To prove the existence and uniqueness of solutions of (2.1)–(2.2), it is embedded in a quasi-linear system of PDEs. There are various ways of doing this but we will use the notation and formulation from Schneider & Wayne [22].

The quasi-linear system we construct is a four-dimensional system for the variables X_1, X_2, U_1 and $V_1 = \partial_t U_1$. All variables are collected in the vector $\mathscr{V} = (X_1, X_2, U_1, V_1)$. Unfortunately, as explained in Schneider & Wayne [21], the variable X_1 is unbounded (in space) and grows rapidly (in time), making the resulting solutions difficult to control over the long time scales which we need to work with. However, as we also discussed in that reference, the derivatives of X_1 do not suffer from this secular growth, and thus it is advantageous to work with the additional variable $Z_1 = \mathscr{K}_0 X_1$ (which for 'long-wavelength' initial conditions behaves like $Z_1 \approx \partial_\alpha X_1$.) The reasons for this particular choice of variables are discussed in more detail in Schneider & Wayne [21, Remark 2.2]. Somewhat surprisingly, the system of equations for the water wave problem can be rewritten entirely in terms of the four variables (Z_1, X_2, U_1, V_1). We define the vectors of variables $\mathscr{W} = (Z_1, X_2, U_1)$ and $\mathscr{W}_e = (Z_1, X_2, U_1, V_1)$. The vectors $\mathscr{W}$ and $\mathscr{W}_e$ will be in the spaces $\mathscr{H}^s = H^s \times H^s \times H^{s-3/2}$ and $\mathscr{H}^s_e = H^s \times H^s \times H^{s-3/2} \times H^{s-3}$, respectively. We also abuse notation slightly and do not distinguish between operators which depend on $\mathscr{V}$ or $\mathscr{W}$, i.e. for instance we will write $\mathscr{K}(X)$ as either $\mathscr{K}(\mathscr{V})$ or $\mathscr{K}(\mathscr{W})$, depending on the circumstances. (Note that it is not immediately apparent that $\mathscr{K}$ can be expressed in terms of $\mathscr{W}$. This is a consequence of the way in which it depends on X_1, as explained in Schneider & Wayne [21].)

2.1 The quasi-linear system

The quasi-linear system for $\mathscr{W}_e = (Z_1, X_2, U_1, V_1)$ constructed in Schneider & Wayne [22] is then given by

$$\begin{aligned} \partial_t Z_1 &= \mathscr{K}_0 U_1, \\ \partial_t X_2 &= \mathscr{K}_0 U_1 + \mathscr{S}_1(\mathscr{W}) U_1, \\ \partial_t U_1 &= V_1, \\ \partial_t V_1 &= \mathscr{L}(\mathscr{W}) U_1 + G_5. \end{aligned} \tag{2.3}$$

We distinguished the relevant linear and quasi-linear terms in the first column on the right-hand side from the semi-linear ones $\mathscr{S}_1(\mathscr{W})U_1$ and G_5 in the second column. The exact form of the nonlinear term, G_5, is not important for what follows – all we need is the fact that it is quadratic and semi-linear, that is, if $\mathscr{W}_e \in \mathscr{H}^s_e$, then for any $R > 0$ there exists C_R such that if $\|\mathscr{W}_e\|_{\mathscr{H}^s_e} \leqslant R$,

$$\|G_5\|_{H^{s-3}} \leqslant C_R \|\mathscr{W}_e\|^2_{\mathscr{H}^s_e},$$

which can be easily verified using the formulas for G_5 provided in Appendix B. The quasi-linear term $\mathscr{L}(\mathscr{W})$ will be more important in what follows and it has the form:

$$\mathscr{L}(\mathscr{W}) U_1 = -\partial_\alpha \big(h_0 \mathscr{K}_0 \partial_\alpha^2 U_1 \big) - h_2 \partial_\alpha^2 U_1 - \mathscr{K}_0 \partial_\alpha U_1 \tag{2.4}$$

$$h_0 = -\mu((1 + \partial_\alpha X_1)^2 + (\partial_\alpha X_2)^2)^{3/2} \tag{2.5}$$

$$h_2 = \frac{3\mu}{((1 + \partial_\alpha X_1)^2 + (\partial_\alpha X_2)^2)^{5/2}} \big((1 + \partial_\alpha X_1)\big(\partial_\alpha^2 X_2\big) - (\partial_\alpha X_2)\big(\partial_\alpha^2 X_1\big) \big) \tag{2.6}$$

In Appendix A we summarize the estimates on $\mathscr{S}_1$, which imply the semi-linearity of the remaining term in the second equation of (2.3). Rather than studying (2.3) in detail at this point we recommend that the reader goes on and we refer back to (2.3) whenever necessary during the remainder of the proof. For more details and the derivation of this system, we refer to Schneider & Wayne [22]. In Schneider & Wayne [22] the following local existence and uniqueness result was proved.

Theorem 2.1 *For all $s \geqslant 6$ there exists a $C_1 > 0$ such that for all $C_2 \in (0, C_1]$ we have a $T_0 > 0$ such that the following is true. For each initial condition $\mathscr{W}_{e,0} \in \mathscr{H}_e^s$ with $\|\mathscr{W}_{e,0}\|_{\mathscr{H}_e^s} \leqslant C_2$ there exists a unique solution $\mathscr{W}_e \in C([0, T_0], \mathscr{H}_e^s)$ of (2.3) with $\mathscr{W}_e|_{t=0} = \mathscr{W}_{e,0}$.*

Local existence and uniqueness of solutions of the water wave problem (2.1) and (2.2) follows indirectly since (2.1) and (2.2) can be identified as a subsystem of (2.3). In particular not all initial conditions $\mathscr{W}_{e,0}$ of (2.3) lead to solutions of the water wave problem (2.1) and (2.2) – only those which have been computed as described above from $X_1|_{t=0}$, $X_2|_{t=0}$, and $U_1|_{t=0}$. Therefore, we introduce the space $\mathscr{C}_{p,X}$ of functions which satisfy the following compatibility conditions, i.e. the subset of the phase space of (2.3) in which the solutions satisfy (2.1) and (2.2).

Definition 2.1 *We define the supset $\mathscr{C}_{p,X}$ of $\mathscr{H}_e^s$ to be those elements of $\mathscr{H}_e^s$ for which there exists $\chi_1 \in H^3$ such that*

$$\begin{aligned}\mathscr{C}_{p,X} = \{\mathscr{W}_e = (\phi_0, \phi_1, \phi_2, \phi_3) \mid\ & (a)\ \phi_0 = \mathscr{K}_0\chi_1, \text{ and if} \\ & \phi_4 \equiv \mathscr{K}(\chi_1, \phi_1)\phi_3 + [\partial_t, \mathscr{K}(\chi_1, \phi_1)]\phi_2 \text{ then} \\ & b)\ (1 + \partial_\alpha\chi_1)\phi_3 + (\partial_\alpha\phi_1)(1 + \phi_4) \\ & \quad = \mu R\big(\partial_\alpha(\chi_1, \phi_1), \partial_\alpha^2(\chi_1, \phi_1)\big) + \mu S\big(\partial_\alpha(\chi_1, \phi_1), \partial_\alpha^3(\chi_1, \phi_1)\big).\end{aligned}$$

Note that in the definition of ϕ_4, the quantity $[\partial_t, \mathscr{K}(\chi_1, \phi_1)]$ is a function of ϕ_0, ϕ_1 and ϕ_2 since as we remarked above, $\mathscr{K}(X_1, X_2)$ can be re-expressed as a function of $Z_1 = \mathscr{K}_0^{-1}X_1$ and X_2, and in addition $\partial_t\chi_1 = \phi_2$ and $\partial_t\phi_1 = \mathscr{K}(\chi_1, \phi_1)\phi_2$. Note further that if $\mathscr{W}|_{t=0} \in \mathscr{C}_{p,X}$ it follows $\mathscr{W}(t) \in \mathscr{C}_{p,X}$ for all $t > 0$ due to the construction of (2.3).

3 The formal approximation

Our approach to prove Theorem 1.1 is the same as that used in Schneider & Wayne [21, 22], namely we first derive, via a formal perturbation expansion, equations whose solutions we believe provide a good approximation to the water wave problem and then prove that the difference between the approximation provided by the solutions of these modulation equations and the true solution of the water wave problem remains small for the time scales of interest. The present section is devoted to deriving the formal modulation equations for the three-wave interaction model.

We write our approximation in the form of the ansatz:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \approx \varepsilon\psi_1 + \varepsilon^2\psi_0 + \varepsilon^2\psi_2 \tag{3.1}$$

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

with

$$\varepsilon\psi_1 = \begin{pmatrix} \sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon a_j A_j(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j \\ \sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon b_j A_j(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j \end{pmatrix}$$

$$\varepsilon^2\psi_0 = \begin{pmatrix} \sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon^2 B_j^0(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j \\ \sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon^2 C_j^0(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j \end{pmatrix}$$

$$\varepsilon^2\psi_2 = \begin{pmatrix} \sum_{\substack{j=-3\\ j\neq 0}}^{3} \sum_{\substack{\ell=-3\\ \ell\neq 0}}^{3} \varepsilon^2 B_{j\ell}(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j\mathscr{E}_\ell \\ \sum_{\substack{j=-3\\ j\neq 0}}^{3} \sum_{\substack{\ell=-3\\ \ell\neq 0}}^{3} \varepsilon^2 C_{j\ell}(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j\mathscr{E}_\ell \end{pmatrix},$$

where $\mathscr{E}_j = e^{i(k_j\alpha+\omega_j t)}$, $\overline{A}_j = A_{-j}$, $k_j = -k_{-j}$, $\omega_j = -\omega_{-j}$, and where k_j and ω_j are chosen such that the resonance condition (1.2) and the linear dispersion relation (1.7) are satisfied.

Remark 3.1 We note that we do not necessarily expect that (3.1) gives an approximation to the water wave problem correct to $\mathcal{O}(\varepsilon^3)$. While one can presumably derive refinements of the TWI-system which describe the $\mathcal{O}(\varepsilon^2)$ terms in X_1, X_2. This may involve additional terms besides B_j and C_j.

As we show below, the amplitude functions A_j satisfy the three-wave interaction equations and thus the term $\varepsilon\psi_1$ in (3.1) is precisely the desired approximation in Theorem 1.1. Note that so long as B_j^0, C_j^0, and $B_{j\ell}$, and $C_{j\ell}$ remain of $\mathcal{O}(1)$ for $0 \leqslant t \leqslant T_0/\varepsilon$, we can choose them to have any value we like without altering the fact that the leading order approximation is $\varepsilon\psi_1$. We will choose these coefficients to eliminate certain unwanted terms in the residual and also to simplify the proof of the approximation theorem in the next section.

To derive the equations satisfied by A_j, B_j^0, C_j^0, $B_{j\ell}$ and $C_{j\ell}$, we insert (3.1) into (2.1)–(2.2) and expand the resulting equations in powers of ε. To identify the various terms we need to know the form of the linear and bilinear terms in $\mathscr{K}(\mathscr{W})$. For this we use the following lemma.

Lemma 3.2 *The operator $\mathscr{K}(X)$ possesses the expansion*

$$\mathscr{K}(X)U_1 = \mathscr{K}_0 U_1 + B_1(X)U_1 + S_2(\mathscr{V})U_1.$$

with

$$\begin{aligned} B_1(X)U_1 &= ([X_1, \mathscr{K}_0]\partial_\alpha U_1) - (X_2 + \mathscr{K}_0(X_2\mathscr{K}_0))\partial_\alpha U_1 \\ &= ([X_1, \mathscr{K}_0]\partial_\alpha U_1) - \left(1 + \mathscr{K}_0^2\right)(X_2\partial_\alpha U_1) - \mathscr{K}_0([X_2, \mathscr{K}_0]\partial_\alpha U_1), \end{aligned} \tag{3.2}$$

$$\mathscr{S}_2(X)U_1 = \mathcal{O}(\|\mathscr{V}\|^2)U_1. \tag{3.3}$$

Proof See Schneider & Wayne [21, Lemma 3.8, Remark 3.9] (based on Craig [7, Lemma 3.7, p. 827]). □

Remark 3.3 Note that, to compare Lemma 3.2 with the expressions in Schneider & Wayne [21], we must recall the notation used in that reference. As mentioned in the previous section in the quasi-linear system (2.3), we replace the variable X_1 by $Z_1 = \mathscr{K}_0 X_1$. To rewrite the equations in terms of these new variables we also define the operator

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

whose properties are summarized in Lemma A.4. To express the term $\partial_\alpha X_1$ in terms of Z_1, we also define the operator

$$\mathscr{M}_2 \cdot = -\partial_\alpha (\mathscr{K}_0)^{-1} \cdot$$

which is a map from H^{s+1} to H^s. Note that, with the aid of $\mathscr{M}_1$, we can rewrite the quadratic term in (3.2) as

$$B_1(\mathscr{W})U_1 = \mathscr{M}_1(Z_1, \partial_\alpha U_1) - \left(1 + \mathscr{K}_0^2\right)(X_2 \partial_\alpha U_1) - \mathscr{K}_0([X_2, \mathscr{K}_0]\partial_\alpha U_1),$$

i.e. as a function of $\mathscr{W} = (Z_1, X_2, U_1, V_1)$, rather than as a function of (X_1, X_2). As explained in Schneider & Wayne [21], this is true not only of the quadratic term in $\mathscr{K}(X)$, but also of the higher order terms as well, so that we can write $\mathscr{K}(X)$ as $\mathscr{K}(\mathscr{W})$. It is this observation that allows us to write (2.3) as a quasi-linear system for $\mathscr{W}$ rather than for $\mathscr{V} = (X_1, X_2, U_1, V_1)$.

Remark 3.4 The estimate on $\mathscr{S}_2$ in (3.3) is not just formal. If $s \geqslant 7/2$, one has the estimate

$$\|\mathscr{S}_2(\mathscr{W})U_1\|_{H^s} \leqslant C\|\mathscr{W}\|^2_{\mathscr{H}^s}\|U_1\|_{H^3}.$$

See Schneider & Wayne [21, Corollary 3.16].

We expand (2.1)–(2.2) with the aid of Lemma 3.2, retaining explicitly the linear and quadratic terms:

$$\partial_t^2 X_1 + \partial_\alpha X_2 - \mu \partial_\alpha^3 X_2 = \mathscr{B}_1(X_1, X_2) + \mathscr{O}(\|\mathscr{W}\|^3), \tag{3.4}$$

$$\partial_t X_2 - \mathscr{K}_0 \partial_t X_1 = \mathscr{B}_2(X_1, X_2) + \mathscr{O}(\|\mathscr{W}\|^3) \tag{3.5}$$

with

$$\begin{aligned}\mathscr{B}_1(X_1, X_2) = &-\left(\partial_t^2 X_1\right)\partial_\alpha X_1 - (\partial_\alpha X_2)\partial_t^2 X_2 \\ &- 3\mu\left(\partial_\alpha^2 X_1\right)\partial_\alpha^2 X_2 - \mu(\partial_\alpha X_2)\partial_\alpha^3 X_1 \\ &- 2\mu(\partial_\alpha X_1)\partial_\alpha^3 X_2, \\ \mathscr{B}_2(X_1, X_2) = &\ X_1 \mathscr{K}_0 \partial_\alpha \partial_t X_1 - \mathscr{K}_0(X_1 \partial_\alpha \partial_t X_1) \\ &- X_2 \partial_\alpha \partial_t X_1 - \mathscr{K}_0((X_2 \mathscr{K}_0)\partial_\alpha \partial_t X_1).\end{aligned}$$

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

We now turn to the derivation of the three-wave interaction model. If we insert (3.1) into (3.4)–(3.5), and consider first the terms proportional to $\varepsilon\mathscr{E}_j$ for $j=\pm1,\pm2,\pm3$, we see that a_j and b_j must satisfy

$$\begin{pmatrix} -\omega_j^2 & ik_j+\mu ik_j^3 \\ -\hat{\mathscr{K}}_0(k_j) & 1 \end{pmatrix}\begin{pmatrix} a_j \\ b_j \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3.6}$$

which leads to the linear dispersion relation (1.7). Provided (1.7) holds, we can choose

$$\begin{pmatrix} a_j \\ b_j \end{pmatrix}=\begin{pmatrix} 1 \\ \hat{\mathscr{K}}_0(k_j) \end{pmatrix}, \tag{3.7}$$

to satisfy (3.6).

We next consider the terms arising from $\varepsilon\psi_1+\varepsilon^2\psi_0$ that are proportional to $\varepsilon^2\mathscr{E}_j$. These yield:

$$\begin{aligned} 2i\omega_j\partial_T A_j-\omega_j^2B_j^0+\hat{\mathscr{K}}_0(k_j)\partial_X A_j+ik_jC_j^0 & \\ +3\mu k_j^2\hat{\mathscr{K}}_0(k_j)\partial_X A_j+\mu ik_j^3C_j^0 &= \text{nonlinear terms}, \\ i\omega_jC_j^0+\hat{\mathscr{K}}_0(k_j)\partial_T A_j-\hat{\mathscr{K}}_0(k_j)\partial_T A_j & \\ -i\omega_j\hat{\mathscr{K}}_0(k_j)B_j^0-\omega_j\frac{d\hat{\mathscr{K}}_0}{dk}(k_j)\partial_X A_j &= \text{nonlinear terms}. \end{aligned}$$

To obtain an equation for A_j alone, we choose $B_j^0=0$, and obtain from the second equation

$$C_j^0=-i\frac{d\hat{\mathscr{K}}_0}{dk}(k_j)\partial_X A_j+\text{nonlinear terms}$$

and so

$$2i\omega_j\partial_T A_j+\left(\hat{\mathscr{K}}_0(k_j)\left(1+3\mu k_j^2\right)+\frac{d\hat{\mathscr{K}}_0}{dk}(k_j)\left(k_j+\mu k_j^3\right)\right)\partial_X A_j=\text{nonlinear terms}$$

which is equivalent to

$$\partial_T A_j=c_g(k_j)\partial_X A_j+\text{nonlinear terms},$$

where $c_g(k_j)=\frac{dw}{dk}(k_j)$ is the group velocity of wave packets with spatial wavenumber k_j.

Finally, we turn to the computation of the nonlinear terms. Our strategy is to substitute (3.1) into (3.4)–(3.5) and use the coefficients $B_{j\ell}$ and $C_{j\ell}$ in $\varepsilon^2\psi_2$ to eliminate as many of the terms of $\mathcal{O}(\varepsilon^2)$ in the nonlinearity as possible. Making this substitution and equating the coefficients of $\varepsilon^2\mathscr{E}_j\mathscr{E}_\ell$ in the linear terms which arise from $\varepsilon^2\psi_2$ with the coefficient of $\varepsilon^2\mathscr{E}_j\mathscr{E}_\ell$ in the nonlinear term, we obtain the equations:

$$\begin{pmatrix} -(\omega_j+\omega_\ell)^2 & i(k_j+k_\ell)+\mu i(k_j+k_\ell)^3 \\ -\hat{\mathscr{K}}_0(k_j+k_\ell) & 1 \end{pmatrix}\begin{pmatrix} B_{j\ell} \\ C_{j\ell} \end{pmatrix}=\begin{pmatrix} \beta_{j\ell} \\ \gamma_{j\ell} \end{pmatrix}A_jA_\ell, \tag{3.8}$$

where the coefficients $\beta_{j\ell}$ and $\gamma_{j\ell}$ are coefficients whose computation in terms of a_j, b_j, ω_j and k_j is straightforward but tedious.

We consider the equations (3.8) in three groups:

Group I

$$(j,\ell) \in \{(1,1),(2,2),(3,3),(-1,-1),(-2,-2),(-3,-3)\\ (1,-2)(1,-3),(-2,1),(-3,1),(2,-1),(-1,2),\\ (2,-3),(-3,2),(3,-2),(-2,3)\}$$

For (j,ℓ) taking any of these values, $\det\begin{pmatrix} -(\omega_j+\omega_\ell)^2 & i(k_j+k_\ell)+\mu i(k_j+k_\ell)^3 \\ -\hat{\mathscr{K}}_0(k_j+k_\ell) & 1 \end{pmatrix} \neq 0$. Thus, we set $B_{j,\ell}$ and $C_{j,\ell}$ equal to the (unique) solution of (3.8) in this case.

Group II

$$(j,\ell) \in \{(1,-1),(2,-2),(3,-3),(-1,1),(-2,2),(-3,3)\}$$

In this case, $\det\begin{pmatrix} -(\omega_j+\omega_\ell)^2 & i(k_j+k_\ell)+\mu i(k_j+k_\ell)^3 \\ -\hat{\mathscr{K}}_0(k_j+k_\ell) & 1 \end{pmatrix} = 0$ but an explicit calculation shows that $\beta_{j,\ell} = \gamma_{j,\ell} = 0$, so we set $B_{j\ell} = C_{j\ell} = 0$ in this case. This is motivated by the fact that the nonlinear terms for these values of (j,ℓ) are of order $\mathcal{O}(\varepsilon^3)$, which in turn results from the fact that the nonlinear terms in the long wave limit (i.e. in the asymptotic limit in which the KdV-equation is the correct modulation equation) contains at least one derivative. These derivatives act on functions which vary slowly in time and space, and as a result these terms are of higher order in ε.

Group III (the resonant terms)

$$(j,\ell) \in \{(1,2),(2,3),(1,3),(-1,-2),(-2,-3),(-1,-3)\}$$

In this case, $\det\begin{pmatrix} -(\omega_j+\omega_\ell)^2 & i(k_j+k_\ell)+\mu i(k_j+k_\ell)^3 \\ -\hat{\mathscr{K}}_0(k_j+k_\ell) & 1 \end{pmatrix} = 0$, but $\beta_{j,\ell}$ and $\gamma_{j,\ell}$ are nonzero.

Thus, these terms cannot be eliminated from the nonlinearity, so we set $B_{j\ell} = C_{j\ell} = 0$ in this case as well.

If we now use the fact that

$$\overline{\mathscr{E}}_1\overline{\mathscr{E}}_2 = \mathscr{E}_3,\ \overline{\mathscr{E}}_1\overline{\mathscr{E}}_3 = \mathscr{E}_2,\ \overline{\mathscr{E}}_2\overline{\mathscr{E}}_3 = \mathscr{E}_1, \tag{3.9}$$

we see that all terms of order ε and ε^2 that result from substituting (3.1) into (3.4) and (3.5) will vanish provided A_j satisfies:

$$\begin{aligned} \partial_T A_1 &= c_g(k_1)\partial_X A_1 + i\gamma_1 \overline{A}_2\overline{A}_3, \\ \partial_T A_2 &= c_g(k_2)\partial_X A_2 + i\gamma_2 \overline{A}_1\overline{A}_3, \\ \partial_T A_3 &= c_g(k_3)\partial_X A_3 + i\gamma_3 \overline{A}_1\overline{A}_2, \end{aligned} \tag{3.10}$$

with constants $\gamma_j \in \mathbb{R}$. The explicit formulas for the γ_j's in terms of the k_j and ω_j can be found in Craik [8]. We have $\gamma_j \in \mathbb{R}$ due to the conservation of energy.

For the proof of the Approximation Theorem 1.1, the approximation has to be extended to the variables Z_1, U_1 and V_1, too. We define $\varepsilon\Psi_{X_1}$ as the first component of (3.1), $\varepsilon\Psi_{X_2}$ as the second component of (3.1), $\Psi_{U_1} = \partial_t\Psi_{X_1}$, and $\Psi_{V_1} = \partial_t^2\Psi_{X_1}$. Recalling the definition of Z_1 we define $\Psi_{Z_1} = \mathscr{K}_0\Psi_{X_1}$.

Remark 3.5 We have no need for the explicit form of Ψ_{Z_1} in what follows, but we note that by taking advantage of the long-wavelength form of the amplitude functions A_j, B_j^0, C_j^0, $B_{j\ell}$ and $C_{j\ell}$, one can write out an expansion for Ψ_{Z_1} in powers of ε. For example, the terms of $\mathcal{O}(\varepsilon)$ are simply

$$\sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon \hat{\mathscr{K}}_0(k_j) A_j(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j.$$

The approximations to the solution are collected in the vector

$$\varepsilon\Psi = \begin{pmatrix} \varepsilon\Psi_{Z_1} \\ \varepsilon\Psi_{X_2} \\ \varepsilon\Psi_{U_1} \\ \varepsilon\Psi_{V_1} \end{pmatrix}. \tag{3.11}$$

This approximation allows us to make the formal error, the so-called residual

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

with

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

small. By the calculations defining the coefficients A_j, B_j^0, C_j^0, $B_{j\ell}$, and $C_{j\ell}$, we have

Lemma 3.6 *Fix $s \geqslant 6$. For all $C_A > 0$ there exist C_Ψ, C_{Res}, $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following is true. Let $A_1, A_2, A_3 \in C([0, T_0], (H^{s+2}(\mathbb{R}, \mathbb{C}))^3)$ be solutions of* (1.1) *satisfying*

$$\sup_{T\in[0,T_0]} \|A_j(T)\|_{H^{s+2}} \leqslant C_A$$

for $j = 1, 2, 3$. Then we have

$$\sup_{t\in[0,T_0/\varepsilon]} \|\Psi(t)\|_{C_b^s} \leqslant \sup_{t\in[0,T_0/\varepsilon]} \|\hat{\Psi}(t)\|_{L^1(s)} \leqslant C_\Psi \tag{3.12}$$

and

$$\sup_{t\in[0,T_0/\varepsilon]} \|\mathrm{Res}(\varepsilon\Psi(t))\|_{\mathscr{H}_e^s} \leqslant C_{\mathrm{Res}}\varepsilon^{5/2},$$

where $\|\hat{u}\|_{L^1(s)} = \|\hat{u}\rho^s\|_{L^1}$ with $\rho(k) = (1 + k^2)^{1/2}$.

Remark 3.7 It may be surprising at first sight that we obtain a bound on the residual of $C_{\mathrm{Res}}\varepsilon^{5/2}$ rather than $C\varepsilon^3$ that the formal calculation led us to expect. This is simply a result of the way the L^2 norms scale for long-wavelength functions, i.e. if $A \in L^2$, and $(\mathscr{S}_\varepsilon A)(x) = A(\varepsilon x)$, then $\|\mathscr{S}_\varepsilon A\|_{L^2} = \varepsilon^{-1/2}\|A\|_{L^2}$. In contrast we have $\|u\|_{C_b^0} = \|\mathscr{S}_\varepsilon u\|_{C_b^0}$, and since $\mathscr{F}(\mathscr{S}_\varepsilon A) = \varepsilon^{-1}\mathscr{S}_{1/\varepsilon}(\mathscr{F}A)$ we have $\|\hat{u}\|_{L^1} = \|\varepsilon^{-1}\mathscr{S}_{1/\varepsilon}\hat{u}\|_{L^1}$. The estimate (3.12) is used, for instance, to estimate $\|\psi R\|_{H^s} \leqslant C\|\psi\|_{C_b^s}\|R\|_{H^s}$ without loss of powers in ε. See also Remark A.6.

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

4 The error estimates

Now we are ready to formulate our main result. For (2.1) and (2.2) written as the first order system

$$\begin{aligned}
\partial_t Z_1 &= \mathscr{K}_0 U_1, \\
\partial_t X_2 &= \mathscr{K}_0 U_1 + \mathscr{S}_1(X) U_1, \\
\partial_t U_1 &= -(1 - \mathscr{M}_2 Z_1 + (\partial_\alpha X_2)\mathscr{K}_0 + (\partial_\alpha X_2)\mathscr{S}_1(X))^{-1}\left[(\partial_\alpha X_2)(1 + [\partial_t, \mathscr{S}_1(X)]U_1)\right. \\
&\qquad \left. - \mu R\left(\partial_\alpha X, \partial_\alpha^2 X\right) - \mu S\left(\partial_\alpha X, \partial_\alpha^3 X\right)\right]
\end{aligned} \tag{4.1}$$

in the variables collected in $\mathscr{W}$ we show that there exist solutions which behave in approximately the same way as predicted by the approximation $\varepsilon\Psi$ defined in (3.11) and constructed via the solutions of the equations for the three-wave interaction (1.1).

Theorem 4.1 *Fix $s \geqslant 6$. 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 $A = (A_1, A_2, A_3) \in C([0, T_0], (H^{s+2})^3)$ be solutions of* (1.1) *with*

$$\sup_{T \in [0, T_0]} \|(A_1, A_2, A_3)\|_{(H^{s+2})^3} \leqslant C_A$$

and let $\mathscr{W}|_{t=0} = \varepsilon\Psi|_{t=0} + \varepsilon^\beta R|_{t=0} \in \mathscr{H}^s$ with $\|R|_{t=0}\|_{\mathscr{H}^s} \leqslant C_0$ and $\beta = 3/2$. Then we have a unique solution $\mathscr{W} = \varepsilon\Psi + \varepsilon^\beta R \in C([0, T_0/\varepsilon], \mathscr{H}^s)$ of (4.1)*, which satisfies*

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

Remark 4.2 Local existence and uniqueness of solutions for (4.1) follows indirectly, since (4.1) is a subsystem of (2.3), namely the system of all solutions of (2.3) in $\mathscr{C}_{p,X}$.

Remark 4.3 Theorem 1.1 is an immediate consequence of Theorem 4.1. The estimates for the Eulerian variables $w = w(x, t)$ and $\eta = \eta(x, t)$ defined by

$$w(\tilde{X}_1(\alpha, t), t) = \partial_t X_1(\alpha, t) \qquad \text{and} \qquad \eta(\tilde{X}_1(\alpha, t), t) = X_2(\alpha, t)$$

follow in a fashion very similar to that of Schneider & Wayne [21, Corollary 1.5], though the fast oscillation of the $\mathscr{E}_j$ with respect to the time scale on which the amplitudes A_j change make it possible to avoid the use of weighted Sobolev spaces here. Given solutions $A = (A_1, A_2, A_3) \in C([0, T_0], (H^{s+2})^3)$ of (1.1), construct $\varepsilon\psi_1 + \varepsilon^2\psi_0 + \varepsilon^2\psi_2$ as in (3.1) and then $\varepsilon\Psi$ as in (3.11). Choose as initial conditions for (4.1) $\mathscr{W}|_{t=0} = \varepsilon\Psi|_{t=0} + \varepsilon^\beta R|_{t=0}$, and let $\mathscr{W} = \varepsilon\Psi + \varepsilon^\beta R$ be the solution of (4.1) constructed in Theorem 4.1. Note that $X_1(\alpha, t) = X_1(\alpha, 0) + \int_0^t U_1(\alpha, s)\, ds$. Theorem 4.1 implies that $U_1 = \varepsilon\Psi_{U_1} + \varepsilon^\beta R_{U_1}$, with $\beta = 3/2$. Since $\Psi_{U_1} = \partial_t \Psi_{X_1}$, equation (3.1) implies that

$$\Psi_{U_1}(\alpha, t) = \sum_{\substack{j=-3 \\ j \neq 0}}^{3} \varepsilon(i\omega_j) a_j A_j(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j + \varepsilon^2 \sum_{\substack{j=-3 \\ j \neq 0}}^{3} a_j \partial_T A_j(\varepsilon\alpha, \varepsilon t)\mathscr{E}_j + \varepsilon^2(\partial_t\psi_0)_1 + \varepsilon^2(\partial_t\psi_2)_1, \tag{4.2}$$

where $(\partial_t\psi_j)_1$ denotes the first component of the vector $\partial_t\psi_j$. Note that from the estimates

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

on A_j (and the fact that $\partial_T A_j$ satisfies (1.1)), we have that there exists $C > 0$ such that

$$\left\| \varepsilon^2 \sum_{\substack{j=-3\\ j\neq 0}}^{3} a_j \partial_T A_j(\varepsilon\cdot, \varepsilon t)\mathscr{E}_j \right\|_{H^{s+1}} \leqslant C\varepsilon^{3/2}$$

for all $0 \leqslant t \leqslant T_0/\varepsilon$. (The 'loss' of half a power of ε is again just a reflection of the way in which the Sobolev norms scale when evaluated on long-wavelength functions.) Similarly, $\|\varepsilon^2(\partial_t\psi_0)_1\|_{H^{s+1}} \leqslant C\varepsilon^{3/2}$ and $\|\varepsilon^2(\partial_t\psi_2)_1\|_{H^{s+1}} \leqslant C\varepsilon^{3/2}$. Thus, combining these estimates with the bounds on R_{U_1} coming from Theorem 4.1, we see that we can write

$$\int_0^t U_1(\alpha, s)\, ds = \sum_{\substack{j=-3\\ j\neq 0}}^{3} \varepsilon(i\omega_j)a_j \int_0^t A_j(\varepsilon\alpha, \varepsilon s)\mathscr{E}_j\, ds + \int_0^t U_1^{rem}(\alpha, s)\, ds,$$

where $\| \int_0^t U_1^{rem}(\cdot, s)\, ds\|_{H^s} \leqslant Ct\varepsilon^{3/2}$. Turning our attention to the integral involving A_j we see that

$$\begin{aligned}\int_0^t A_j(\varepsilon\alpha, \varepsilon s)e^{i(k_j\alpha+\omega_j s)}\, ds &= \int_0^t A_j(\varepsilon\alpha, \varepsilon s)\frac{1}{i\omega_j}\partial_s\left(e^{i(k_j\alpha+\omega_j s)}\right) ds\\ &= \frac{1}{i\omega_j}A_j(\varepsilon\alpha, \varepsilon s)e^{i(k_j\alpha+\omega_j s)}\Big|_0^t - \frac{1}{i\omega_j}\varepsilon\int_0^t \partial_T A_j(\varepsilon\alpha, \varepsilon s)e^{i(k_j\alpha+\omega_j s)}\, ds. \qquad (4.3)\end{aligned}$$

Recalling again the estimates on A_j, and the way in which the H^s norms of long-wavelength functions scale, we have

$$\left\| \varepsilon \sum_{\substack{j=-3\\ j\neq 0}}^{3} (i\omega_j)a_j \int_0^t A_j(\varepsilon\cdot, s)\mathscr{E}_j\, ds \right\|_{H^{s+1}} \leqslant C\left(\sqrt{\varepsilon} + t\varepsilon^{3/2}\right).$$

Combining this with the estimates above, we see that for $0 \leqslant t \leqslant T_0/\varepsilon$, we have $\|X_1(\cdot, t) - X_1(\cdot, 0)\|_{C_b^{s-2}} \leqslant C\sqrt{\varepsilon}$. Thus, by the inverse function theorem the function $\tilde{X}_1(\alpha, t) = \alpha + X_1(\alpha, t)$ has an inverse $\tilde{X}_1^{-1}(x, t) = x + \Xi(x, t)$ with

$$\sup_{t\in[0,T_0/\varepsilon]} \|\Xi(\cdot, t)\|_{C_b^{s-2}} \leqslant C\sqrt{\varepsilon}.$$

Thus, if we denote by $\varepsilon\psi_{1,2}$ the second component of the vector $\varepsilon\psi_1$, and note that it is equal to the order ε term in $\varepsilon\Psi_{X_2}$ we have

$$\begin{aligned}\sup_{t\in[0,T_0/\varepsilon]} \|\eta(\cdot, t) - \varepsilon\psi_{1,2}(\cdot, t)\|_{C_b^{s-2}} &\leqslant \sup_{t\in[0,T_0/\varepsilon]} \big(\|X_2(\cdot, t) - \varepsilon\psi_{1,2}(\cdot, t)\|_{C_b^{s-2}}\\ &\qquad + \|X_2(\cdot, t) - \eta(\cdot, t)\|_{C_b^{s-2}}\big)\\ &= \sup_{t\in[0,T_0/\varepsilon]} \left(\|X_2(\cdot, t) - \varepsilon\psi_{1,2}(\cdot, t)\|_{C_b^{s-2}} + \|X_2(\cdot, t) - X_2\big(\tilde{X}_1^{-1}(\cdot, t), t\big)\|_{C_b^{s-2}}\right)\\ &\leqslant C\varepsilon^{3/2} + C\varepsilon^{3/2}.\end{aligned}$$

The estimate on w is similar and Theorem 1.1 follows.

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

Proof of Theorem 4.1 The proof is very close to that of Schneider & Wayne [22], where the validity of the KdV and Kawahara equations has been established for long-wavelength initial data, though as mentioned in the introduction the proof is actually somewhat simpler here due to the shorter time scales involved.

We write a solution $\mathscr{W} = (Z_1, X_2, U_1)$ of (4.1) as a sum of the approximation $\varepsilon\Psi$ and an error $\varepsilon^\beta\mathscr{R}$ with

$$\varepsilon\Psi = \left(\varepsilon\Psi_{Z_1}, \varepsilon\Psi_{X_2}, \varepsilon\Psi_{U_1}\right) \text{ and } \varepsilon^\beta\mathscr{R} = \left(e^\beta R_{Z_1}, \varepsilon^\beta R_{X_2}, \varepsilon^\beta R_{U_1}\right),$$

and we write a solution $\mathscr{W}_e = (Z_1, X_2, U_1, V_1)$ of (2.3) as a sum of the approximation

$$\varepsilon\Psi_e = \left(\varepsilon\Psi_{Z_1}, \varepsilon\Psi_{X_2}, \varepsilon\Psi_{U_1}\varepsilon\Psi_{V_1}\right) \text{ and } \varepsilon^\beta\mathscr{R}_e = \left(e^\beta R_{Z_1}, \varepsilon^\beta R_{X_2}, \varepsilon^\beta R_{U_1}\varepsilon^\beta R_{V_1}\right).$$

To control the growth of the solutions in an optimal way, it is advantageous to consider (4.1) in parallel with (2.3). As we will see, we can use (4.1) to control the low order derivatives of the solution, and consider (2.3) only to control the highest order derivatives. We expand the right-hand side of both systems in Taylor polynomials, in (4.1) considering explicitly only the linear terms, while in (2.3) we must retain explicitly all the quasi-linear terms.

Throughout the rest of this section, we assume that the following standing hypothesis holds:

(HS) *For all* $C_R > 0$ *there exist constants* $\varepsilon_0 > 0$ *and* $C > 0$ *such that the following estimates hold for all* $\varepsilon \in (0, \varepsilon_0)$ *and* $t \geqslant 0$ *as long as* $\sup_{\tau\in(0,t)} \|\mathscr{R}_e(\tau)\|_{\mathscr{H}_e^s} \leqslant C_R$.

This assumption will be validated with the energy estimate at the end of the section, and (HS) should be viewed as boot strap assumption.

If we first insert the ansatz $\mathscr{W} = \varepsilon\Psi + \varepsilon^\beta\mathscr{R}$ into (4.1) and expand, identifying explicitly the linear terms, we obtain

$$\begin{aligned} \partial_t R_{Z_1} &= \mathscr{K}_0 R_{U_1}, \\ \partial_t R_{X_2} &= \mathscr{K}_0 R_{U_1} + \mathscr{N}_1, \\ \partial_t R_{U_1} &= -\partial_\alpha R_{X_2} + \mu\partial_\alpha^3 R_{X_2} + \mathscr{N}_2, \end{aligned} \tag{4.4}$$

where with the aid of Lemma 3.2, Remark 3.4 and Remark A.6, we see that provided (HS) holds:

$$\|\mathscr{N}_1\|_{H^s} \leqslant C\left(\varepsilon\|\mathscr{R}\|_{\mathscr{H}^s} + \varepsilon^\beta\|\mathscr{R}\|^2_{\mathscr{H}^s} + \varepsilon C_{\text{Res}}\right), \tag{4.5}$$

$$\|\mathscr{N}_2\|_{H^{s-3}} \leqslant C\left(\varepsilon\|\mathscr{R}\|_{\mathscr{H}^s} + \varepsilon^\beta\|\mathscr{R}\|^2_{\mathscr{H}^s} + \varepsilon C_{\text{Res}}\right). \tag{4.6}$$

(See also Schneider & Wayne [22, equations (35) and (36)], where similar estimates are derived.)

Note that we cannot use these equations to control the long-time behavior of the $H^{s-3/2}$ norm of R_{U_1} since we would lose regularity. Thus, as we did in previous work [21, 22], we extend (4.4) by considering also the evolution of $V_1 = \partial_t U_1$. To prevent misunderstanding, we remind the reader that solutions of (2.3) which lie in $\mathscr{C}_{p,X}$ are also solutions of the

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

water wave problem (4.1). Inserting our ansatz $\mathscr{W}_e = \varepsilon\Psi_e + \varepsilon^\beta \mathscr{R}_e$ into (4.1) we obtain

$$\begin{aligned}
\partial_t R_{Z_1} &= \mathscr{K}_0 R_{U_1} \\
\partial_t R_{X_2} &= \mathscr{K}_0 R_{U_1} + \mathscr{N}_1 \\
\partial_t R_{U_1} &= R_{V_1}, \\
\partial_t R_{V_1} &= -\partial_\alpha\left(h_0 \partial_\alpha^2 \mathscr{K}_0 R_{U_1}\right) - h_2 \partial_\alpha^2 R_{U_1} - \mathscr{K}_0 \partial_\alpha R_{U_1} + \mathscr{N}_3,
\end{aligned} \tag{4.7}$$

where we have retained all the quasi-linear terms and where $\mathscr{N}_3$ obeys

$$\|\mathscr{N}_3\|_{H^{s-3}} \leqslant C\left(\varepsilon \|\mathscr{R}_e\|_{\mathscr{H}_e^s} + \varepsilon^\beta \|\mathscr{R}_e\|^2_{\mathscr{H}_e^s} + \varepsilon C_{\mathrm{Res}}\right)$$

under the standing hypothesis (HS) (and $\mathscr{N}_1$ satisfies the same estimate as before).

We also recall that the coefficients h_0 and h_2 were defined in (2.5) and (2.6). Solutions of this system are controlled using the energy function constructed in Schneider & Wayne [22] and the general outline of what follows is similar to the argument used there. As a first step, the variable R_{Z_1} is estimated in terms of the others. 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 $\mathscr{H}_\ell^s = H^s \times H^{s-3/2} \times H^{s-3}$-scalar product. We define $\mathscr{R}_\ell = (R_{X_2}, R_{U_1}, R_{V_1})$, the lower part of $\mathscr{R}_e$. The main part of this energy functional is chosen such that in the computation of $\frac{d}{dt}\mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)$ all quasi-linear terms from (2.3) cancel, and we can estimate $\frac{d}{dt}\mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)$ in terms of $\mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)$. Once this is achieved the desired estimate on the norm of $\mathscr{W}_e$ follows easily from Gronwall's inequality.

We begin by estimating R_{Z_1} in terms of $\mathscr{R}_\ell$. From the first two equations for the error, it follows that

$$\partial_t R_{X_2} - \partial_t R_{Z_1} = \mathscr{N}_1$$

where $\mathscr{N}_1$ satisfies (4.5). Integration with respect to time and Gronwall's inequality yields the estimate

$$\begin{aligned}
&\forall C_A, T_0 \exists C_4, C_5 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0): \\
&\sup_{t\in[0,T_0/\varepsilon]} \|R_{Z_1}(t)\|_{H^s} \leqslant C_4 + C_5 \|\mathscr{R}_\ell(t)\|_{\mathscr{H}_\ell^s},
\end{aligned}$$

as long as $\sup_{t\in[0,T_0/\varepsilon]} \|\mathscr{R}_e(t)\|_{H_e^s} \leqslant C_R$.

Recall that we are interested in solutions of (4.7) that lie in $\mathscr{C}_{p,X}$, i.e. solutions that are actually solutions of the original water wave problem. From Definition 2.1, we see that for $\mathscr{W}_e \in \mathscr{C}_{p,X}$ 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

$$\mu \partial_\alpha^3 R_{X_2} - \partial_\alpha R_{X_2} = R_{V_1} + \mathscr{N}_4$$

with

$$\|\mathscr{N}_4\|_{H^{s-3}} \leqslant C\left(\varepsilon \|\mathscr{R}\|_{\mathscr{H}^s} + \varepsilon^\beta \|\mathscr{R}\|^2_{\mathscr{H}^s} + \varepsilon C_{\mathrm{Res}}\right).$$

In particular, for solutions in $\mathscr{C}_{p,X}$, we can estimate the highest derivatives $\partial_\alpha^s R_{X_2}$ appearing

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

in the following by

$$\left\|\partial_\alpha^s R_{X_2}\right\|_{L^2} \leqslant \mu^{-1}\left(\left\|\partial_\alpha^{s-3} R_{V_1}\right\|_{L^2} + \left\|\partial_\alpha^{s-2} R_{X_2}\right\|_{L^2} + \left\|\partial_\alpha^{s-3} \mathscr{N}_4\right\|_{L^2}\right),$$

so we can control $\|R_{X_2}\|_{H^s}$ by $\|R_{X_2}\|_{L^2}$, $\|R_{U_1}\|_{H^{s-3/2}}$, and $\|R_{V_1}\|_{H^{s-3}}$.

Given the preceeding estimates, all that remains is to estimate $\|R_{X_2}\|_{L^2}$, $\|R_{U_1}\|_{H^{s-3/2}}$, and $\|R_{V_1}\|_{H^{s-3}}$. For this purpose, we use the energy functional constructed in Schneider & Wayne [22], whose definition we now recall.

Since the operator $\mathscr{M}_2 = -\partial_\alpha \mathscr{K}_0^{-1}$ is self-adjoint and positive, we can take its square root and find

$$\begin{aligned}\partial_t \int \left(\mathscr{M}_2^{1/2} R_{X_2}\right)^2 \, d\alpha/2 &= \int R_{X_2}\left(\mathscr{M}_2 \partial_t R_{X_2}\right) d\alpha \\ &= \int R_{X_2}\left(\mathscr{M}_2\left(\mathscr{K}_0 R_{U_1} + \mathscr{N}_1\right)\right) d\alpha \\ &= \int \left(\partial_\alpha R_{X_2}\right) R_{U_1} \, d\alpha + \mathscr{N}_{e1},\end{aligned} \tag{4.8}$$

where

$$|\mathscr{N}_{e1}| \leqslant C\left(\varepsilon \|\mathscr{R}_\ell\|_{\mathscr{H}_e^s}^2 + \varepsilon^\beta \|\mathscr{R}_\ell\|_{\mathscr{H}_\ell^s}^3 + \varepsilon C_{\text{Res}} \|\mathscr{R}_\ell\|_{\mathscr{H}_\ell^s}\right)$$

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

$$\partial_t \int \left(\mathscr{M}_3^{1/2} R_{X_2}\right)^2 \, d\alpha/2 = -\int \left(\mu \partial_\alpha^3 R_{X_2}\right) R_{U_1} \, d\alpha + \mathscr{N}_{e2}, \tag{4.9}$$

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

Next using the evolution equation for R_{U_1} we see that

$$\begin{aligned}\partial_t \int \left(R_{U_1}\right)^2 d\alpha/2 &= \int R_{U_1}(\partial_t R_{U_1}) \, d\alpha \\ &= \int R_{U_1}\left(-\partial_\alpha R_{X_2} + \mu \partial_\alpha^3 R_{X_2} + \mathscr{N}_2\right) d\alpha \\ &= -\int R_{U_1}\left(\partial_\alpha R_{X_2}\right) d\alpha + \mu \int R_{U_1}\left(\partial_\alpha^3 R_{X_2}\right) d\alpha + \mathscr{N}_{e3},\end{aligned} \tag{4.10}$$

where $\mathscr{N}_{e3}$ obeys the same estimates as $\mathscr{N}_{e1}$.

The first two terms on the right-hand side cancel with the corresponding terms in the time derivatives of $\int (\mathscr{M}_2^{1/2} R_{X_2})^2 \, d\alpha$ and $\int (\mathscr{M}_3^{1/2} R_{X_2})^2 \, d\alpha$.

We define a pair of *skew-symmetric* operators, λ_1 and λ_2 related to $\mathscr{M}_2$, namely:

$$\lambda_1^2 = -\mathscr{K}_0 \partial_\alpha, \tag{4.11}$$

$$\lambda_2^2 = -\mathscr{M}_2 = \partial_\alpha \mathscr{K}_0^{-1}. \tag{4.12}$$

Recalling the definitions of h_0 and h_2 from above, we define:

$$\Lambda \cdot = -\partial_\alpha \lambda_1 (h_0 \partial_\alpha \lambda_1) \cdot - \lambda_1 \lambda_2 (h_2 \lambda_1 \lambda_2) \cdot - \lambda_1^2 \cdot . \tag{4.13}$$

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

Then, just as in Schneider & Wayne [22], we treat the quasi-linear terms by introducing $R_U = \lambda_1 R_{U_1}$, and $R_V = \lambda_1 R_{V_1}$ and considering

$$\frac{1}{2}\frac{d}{dt}\left((\Lambda^r R_V, \Lambda^r R_V)_{L^2} + \left(\Lambda^{r+\frac{1}{2}} R_U, \Lambda^{r+\frac{1}{2}} R_U\right)_{L^2}\right)$$
$$= (\Lambda^r R_V, \Lambda^r \lambda_1 \mathcal{N}_3)_{L^2} + (\Lambda^r R_V, \partial_t(\Lambda^r) R_V)_{L^2} + \left(\Lambda^{r+\frac{1}{2}} R_U, \partial_t\left(\Lambda^{r+\frac{1}{2}}\right) R_U\right)_{L^2}, \quad (4.14)$$

where $3r = s - \frac{7}{2}$. The detailed calculation to obtain the relation (4.14), i.e. the cancellation of the quasi-linear terms can be found in Schneider & Wayne [22, Section 2]. With the estimates for $\mathcal{N}_3$ from above, it follows that

$$|(\Lambda^r R_V, \Lambda^r \lambda_1 \mathcal{N}_3)_{L^2}| \leqslant C\left(\varepsilon \|\mathcal{R}_\ell\|^2_{\mathcal{H}^s_\ell} + \varepsilon^\beta \|\mathcal{R}_\ell\|^3_{\mathcal{H}^s_\ell} + \varepsilon C_{\text{Res}} \|\mathcal{R}_\ell\|_{\mathcal{H}^s_\ell}\right). \quad (4.15)$$

By the calculations of Schneider & Wayne [22, Section 4], it is known that the time derivatives $\partial_t(\Lambda^r)$ and $\partial_t(\Lambda^{r+\frac{1}{2}})$ can be estimated as follows

$$|(\Lambda^r R_V, \partial_t(\Lambda^r) R_V)_{L^2}| + \left|\left(\Lambda^{r+\frac{1}{2}} R_U, \partial_t\left(\Lambda^{r+\frac{1}{2}}\right) R_U\right)\right|$$
$$\leqslant C\left((\Lambda^r R_V, \Lambda^r R_V)_{L^2} + \left(\Lambda^{r+\frac{1}{2}} R_U, \Lambda^{r+\frac{1}{2}} R_U\right)_{L^2}\right.$$
$$\left. + (R_{Z_1}, R_{Z_1})_{H^s} + (R_{X_2}, R_{X_2})_{H^s} + (R_{U_1}, R_{U_1})_{H^{s-\frac{3}{2}}} + (R_{V_1}, R_{V_1})_{H^{s-3}}\right), \quad (4.16)$$

i.e. these terms can be considered as semi-linear. We need slightly more, however, namely that these terms are $\mathcal{O}(\varepsilon)$. To see why this is so, note that using the form of Λ in (4.13) we have:

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

All these terms are bounded in a very similar fashion – we shall look at one of them explicitly, and then leave the rest as an exercise.

Consider, for example, $(\Lambda^r R_V, \Lambda^{r-1}(\partial_\alpha \lambda_1 (\partial_t h_0) \partial_\alpha \lambda_1) R_V)_{L^2}$. From the formula for h_0, we see that

$$\partial_t h_0 = \frac{3\mu[(1 + \mathcal{M}_2 Z_1)(\mathcal{M}_2(\partial_t Z_1)) + (\partial_\alpha X_2)(\partial_\alpha \partial_t X_2)]}{[(1 + (\mathcal{M}_2 Z_1))^2 + (\partial_\alpha X_2)^2]^{5/2}}$$
$$= \frac{3\mu[(1 + \mathcal{M}_2 Z_1)(\mathcal{M}_2(\mathcal{K}_0 U_1)) + (\partial_\alpha X_2)(\partial_\alpha[\mathcal{K}_0 U_1 + \mathcal{S}_1(\mathcal{W}) U_1])]}{[(1 + (\mathcal{M}_2 Z_1))^2 + (\partial_\alpha X_2)^2]^{5/2}}. \quad (4.17)$$

Now recalling that $Z_1 = \varepsilon \Psi_{Z_1} + \varepsilon^\beta R_{Z_1}$, $X_2 = \varepsilon \Psi_{X_2} + \varepsilon^\beta R_{X_2}$ and $U_1 = \varepsilon \Psi_{U_1} + \varepsilon^\beta R_{U_1}$, we see that

$$\|\Lambda^{r-1} \partial_\alpha \lambda_1 (\partial_t h_0)\|_{L^\infty} \leqslant C\left(C_{\text{Res}} \varepsilon + \varepsilon \|\mathcal{R}_e\|_{\mathcal{H}^s_e} + \varepsilon^\beta \|\mathcal{R}_e\|^2_{\mathcal{H}^s_e}\right),$$

under hypothesis **(HS)**. Thus,

$$|(\Lambda^r R_V, \Lambda^{r-1}(\partial_\alpha \lambda_1 (\partial_t h_0) \partial_\alpha \lambda_1) R_V)_{L^2}| \leqslant C\left(C_{\text{Res}} \varepsilon \|\mathcal{R}_e\|_{\mathcal{H}^s_e} + \varepsilon \|\mathcal{R}_e\|^2_{\mathcal{H}^s_e} + \varepsilon^\beta \|\mathcal{R}_e\|^3_{\mathcal{H}^s_e}\right).$$

The remaining terms can all be bounded in a similar fashion, and if we take advantage of the fact that the H^s norm of R_{Z_1} can be controlled by $\|\mathcal{R}_\ell\|_{\mathcal{H}^s_\ell}$, we can combine these

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

estimates with (4.15) to obtain

$$\frac{1}{2}\frac{d}{dt}\left((\Lambda^r R_V, \Lambda^r R_V)_{L^2} + \left(\Lambda^{r+\frac{1}{2}} R_U, \Lambda^{r+\frac{1}{2}} R_U\right)_{L^2}\right) \leqslant C\left(C_{\mathrm{Res}}\varepsilon \|\mathscr{R}_\ell\|_{\mathscr{H}^s_\ell} + \varepsilon \|\mathscr{R}_\ell\|^2_{\mathscr{H}^s_\ell} + \varepsilon^\beta \|\mathscr{R}_\ell\|^3_{\mathscr{H}^s_\ell}\right).$$

With these considerations, we define the energy

$$\begin{aligned}\mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell) = &\int \left(\mathscr{M}_2^{1/2} R_{X_2}\right)^2 + \int \left(\mathscr{M}_3^{1/2} R_{X_2}\right)^2 + \int \left(R_{U_1}\right)^2 \\ &+ (\Lambda^r R_V, \Lambda^r R_V)_{L^2} + \left(\Lambda^{r+\frac{1}{2}} R_U, \Lambda^{r+\frac{1}{2}} R_U\right)_{L^2}.\end{aligned}$$

Recalling that the H^s norm of R_{X_2} can be controlled by the L^2 norm of R_{X_2}, the $H^{s-3/2}$ norm of R_{U_1} and the H^{s-3} norm of R_{V_1} and recalling that the H^s norm of R_{Z_1} can be controlled by all others we see that the scalar product defined by $\mathbf{E}_s(\cdot,\cdot)$ is equivalent to the usual $\mathscr{H}^s_e$ 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

$$\|\mathscr{W}\|^2_{\mathscr{H}^s_e} \leqslant c_1 \mathbf{E}_s(\mathscr{W}_\ell, \mathscr{W}_\ell) \leqslant c_2 \|\mathscr{W}\|^2_{\mathscr{H}^s_e}. \tag{4.18}$$

Therefore, combining (4.8), (4.9), (4.10), (4.14), (4.15) and (4.16), we see that there are constants $C_1 = C_1(C_\psi, C_{\mathrm{Res}}, c_j)$, $C_2 = C_2(C_\psi, C_{\mathrm{Res}}, C_R, c_j)$ and $C_3 = C_3(C_\psi, C_{\mathrm{Res}}, c_j)$, such that

$$\begin{aligned}\frac{1}{2}\partial_t \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell) &\leqslant \varepsilon C_1 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell) + \varepsilon^\beta C_2 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)^{\frac{3}{2}} + \varepsilon C_3 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)^{\frac{1}{2}} \\ &\leqslant \varepsilon C_1 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell) + \varepsilon^\beta C_2 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)^{\frac{3}{2}} + \varepsilon C_3 + \varepsilon C_3 \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell).\end{aligned}$$

Applying Gronwall's inequality shows, for all $t \in [0, T_0/\varepsilon]$,

$$\begin{aligned}\mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)(t) &\leqslant \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)(0) + \int_0^t e^{(C_1+C_3+1)\varepsilon\tau}\, d\tau \cdot C_3 \varepsilon \\ &\leqslant \mathbf{E}_s(\mathscr{R}_\ell, \mathscr{R}_\ell)(0) + C_3 T_0 e^{(C_1+C_3+1)T_0} =: c_1^{-1} C_R,\end{aligned}$$

where we have chosen $\varepsilon_0 > 0$ so small that $\varepsilon^{\beta-1} C_2(C_R) \cdot c_2 C_R^2 < 1$. This yields with (4.18) that

$$\sup_{t \in [0, T_0/\varepsilon]} \|\mathscr{R}_e\|^2_{\mathscr{H}^s_e} \leqslant C_R.$$

This completes the proof of Theorem 4.1 □

5 Conclusion

In this paper, we have proved that the dynamics of three resonant wave packets in the 2D water wave problem, can approximately be described by the equations for the so called three-wave interaction. To do so, we established estimates between the formal approximation obtained via the TWI-system and true solutions of the Lagrangian formulation of the 2D water wave problem.

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

There are two natural generalizations of this result. First, the proof of such estimates for the 3D water wave problem. A major difficulty which has to be overcome in the 3D case is the proof of a local existence and uniqueness theorem which so far is not available for the situation considered here.

The second generalization is the justification of the so called four-wave-interaction system, i.e. the description of four resonant wave packets with spatial wavenumbers k_j and temporal wavenumbers ω_j satisfying

$$k_1 + k_2 + k_3 + k_4 = 0 \qquad \text{and} \qquad \omega_1 + \omega_2 + \omega_3 + \omega_4 = 0.$$

There is a new qualitative difficulty due to the longer time scale, which is $\mathcal{O}(\frac{1}{\varepsilon^2})$ in contrast to $\mathcal{O}(\frac{1}{\varepsilon})$ for the TWI-case. We expect that for a proof normal form transforms are necessary to eliminate the quadratic terms in the Lagrangian formulation of the 2D water wave problem.

Acknowledgement

The work of Guido Schneider was partially supported by the Deutsche Forschungsgemeinschaft DFG under the grant Schn520/3-1. The work of Eugene Wayne was supported in part by the NSF through the grant DMS-0103915. In addition, this paper was completed while the second author was visiting the Division of Applied Mathematics at Brown University whose hospitality is gratefully acknowledged. Finally, both authors would like to thank David Sattinger for very useful information about the complete integrability of the TWI-system.

Appendix A Some estimates on the operator $\mathcal{K}(X)$

In this appendix, we summarize several estimates which are used in previous sections. For more details we refer to Schneider & Wayne [21], and the literature cited there.

As already said the operator $\mathcal{K}(X)$ is of the form $\mathcal{K}(X) = \mathcal{K}_0 + \mathcal{S}_1(X)$, where the operator $\mathcal{S}_1(X)\cdot$ has certain smoothing properties. As a rule a term with $\mathcal{S}_1(X)\cdot$ in front will have the regularity of X. More precisely, in terms of the variables $\mathcal{W}$ defined in § 2.

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

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

Proof See Schneider & Wayne [21, Lemma 3.14]. □

Lemma A.2 *Assume the situation of Lemma* A.1. *Then for all $s \geqslant 6$ we have*

$$\|\partial_t(\mathcal{S}_1(\mathcal{W})U_1)\|_{H^{s-3}} \leqslant C\|\mathcal{W}\|_{\mathcal{H}^s}(\|U_1\|_{H^3} + \|V_1\|_{H^3}),$$
$$\left\|[\partial_t^2, \mathcal{S}_1(\mathcal{W})]U_1\right\|_{H^{s-3}} \leqslant C\|\mathcal{W}_e\|_{\mathcal{H}_e^s}(\|U_1\|_{H^3} + \|V_1\|_{H^3}),$$
$$\|\partial_\alpha(\mathcal{S}_1(\mathcal{W})U_1)\|_{H^{s-1}} \leqslant C\|\mathcal{W}\|_{\mathcal{H}^s}\|U_1\|_{H^3}.$$

Proof See Schneider & Wayne [21, Lemma 3.15]. □

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

In the quasi-linear system commutators $[a, \mathscr{K}_0]\cdot$ play a big role. As a rule $[a, \mathscr{K}_0]u$ smooths u and has the regularity of the function a.

Lemma A.3 *Let $r \geqslant 0$, $q > 1/2$ and $0 \leqslant p \leqslant q$. Then there exists a $C > 0$ such that*

$$\|[a, \mathscr{K}_0]u\|_{H^r} \leqslant C\|a\|_{H^{r+p}}\|u\|_{H^{q-p}}.$$

Proof See Schneider & Wayne [21, Lemma 3.12]. □

To avoid the secular growth in the variable X_1, we introduced the variable $Z_1 = \mathscr{K}_0 X_1$ and we associated to Z_1 the operator

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

which satisfies

Lemma A.4 *Let $r \geqslant 0$, $q > 1/2$ and $0 \leqslant p \leqslant q$. Then there exists a $C > 0$ such that*

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

Proof See Schneider & Wayne [21, Corollary 3.13]. □

Remark A.5 $\mathscr{M}_1$ is well defined, even though $\mathscr{K}_0$ is not invertible in general, due to the commutator in its definition.

To express the term $\partial_\alpha X_1$ in terms of Z_1, we defined additionally the operator

$$\mathscr{M}_2\cdot = -\partial_\alpha(\mathscr{K}_0)^{-1}\cdot$$

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

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

Remark A.6 Examining the expression for for $\|[a, \mathscr{K}_0]u\|_{H^r}$ on Schneider & Wayne [21, p. 1499], (which is expressed in terms of the Fourier transforms of a and u), one also sees that it can be bounded by

$$\|[a, \mathscr{K}_0]u\|_{H^r} \leqslant C \min\left(\|\hat{a}\|_{L^1(r+p)}\|u\|_{H^{q-p}}, \|a\|_{H^{r+p}}\|\hat{u}\|_{L^1(q-p)}\right), \tag{A 1}$$

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

$$\|\mathscr{M}_1(a, u)\|_{H^r} \leqslant C \min\left(\|\hat{a}\|_{L^1(r+p)}\|u\|_{H^{q-p}}, \|a\|_{H^{r+p}}\|\hat{u}\|_{L^1(q-p)}\right). \tag{A 2}$$

Estimates of this kind were also used in Schneider & Wayne [21] – see, for example, equation (6.8).

We can use these commutator estimates, along with the expansion in Lemma 3.2 to obtain an estimate of $\mathscr{S}_1(\mathscr{W})$ without the loss of half a power of ε usually associated with the H^s norms. Note that if we write $\mathscr{S}_1(\varepsilon\Psi)U = B_1(\varepsilon\Psi)U + \mathscr{S}_2(\varepsilon\Psi)U$, bound the commutators the expression for B_1 in (3.2) by the estimates in (A 1) and (A 2), and take advantage of the fact that $(1+\mathscr{K}_0^2)$ is infinitely smoothing, and if we bound $\mathscr{S}_2(\varepsilon\Psi)U$ by the estimate in Remark 3.4, then we obtain

$$\|\mathscr{S}_1(\mathscr{W})U\|_{H^s} \leqslant C\varepsilon\|U\|_{H^3},$$

where the constant C depends on the norm of Ψ, but is independent of ε. Similar estimates hold the inequalities in Lemma A.2. See also Remark 3.7.

Appendix B The quasi-linear system

In this appendix, we give the detailed form of the various terms in the quasi-linear system (2.3). This form of the water wave problem was derived in Section 2 of Schneider & Wayne [22], which the reader may consult for additional details. Recall that the quasi-linear system had the form

$$\begin{aligned}
\partial_t Z_1 &= \mathscr{K}_0 U_1,\\
\partial_t X_2 &= \mathscr{K}_0 U_1 + \mathscr{S}_1(\mathscr{W})U_1,\\
\partial_t U_1 &= V_1,\\
\partial_t V_1 &= \mathscr{L}(\mathscr{W})U_1 + G_5.
\end{aligned} \tag{B 1}$$

The quasi-linear term $\mathscr{L}(\mathscr{W})$ was discussed in § 2, so here we just give the (long) expression for G_5.

$$\begin{aligned}
G_5 &= -G_4 + \mathscr{K}_0\partial_\alpha U_1,\\
G_4 &= (f_3 - H_1)^{-1}\big(G_3 + H_3\big(h_0 f_3\mathscr{K}_0\partial_\alpha^3 U_1\big) + h_1 a_1 f_3\partial_\alpha^2 U_1 + h_1 a_2 f_3\mathscr{K}_0\partial_\alpha^2 U_1\big)\big),\\
G_3 &= (f_1 - f_2\mathscr{K}_0)G_2 + [(f_1 - f_2\mathscr{K}_0), h_0](f_1 + f_2\mathscr{K}_0)\mathscr{K}_0\partial_\alpha^3 U_1\\
&\quad + [(f_1 - f_2\mathscr{K}_0), h_1 a_1](f_1 + f_2\mathscr{K}_0)\partial_\alpha^2 U_1\\
&\quad + [(f_1 - f_2\mathscr{K}_0), h_1 a_2](f_1 + f_2\mathscr{K}_0)\mathscr{K}_0\partial_\alpha^2 U_1\\
&\quad - h_0 H_2\mathscr{K}_0\partial_\alpha^3 U_1 - h_1 a_1 H_2\partial_\alpha^2 U_1 - h_1 a_2 H_2\mathscr{K}_0\partial_\alpha^2 U_1,\\
G_2 &= G_1 - h_0 f_2\big(1 + \mathscr{K}_0^2\big)\partial_\alpha^3 U_1\\
&\quad + 3\mu Q^{-5}(f_1\partial_\alpha f_1 + f_2\partial_\alpha f_2)f_2\big(1 + \mathscr{K}_0^2\big)\partial_\alpha^2 U_1,\\
G_1 &= \partial_t U_1\big(1 + \mathscr{K}_0^2\big)\partial_\alpha U_1 + (1 + \partial_t U_2)\partial_\alpha(\mathscr{S}_1(X)U_1),\\
&\quad + (\partial_\alpha X_2)\big(\big[\partial_t^2, \mathscr{S}_1(X)\big]U_1\big) + ((1 + \partial_t U_2) - \partial_t U_1\mathscr{K}_0)\mathscr{K}_0\partial_\alpha U_1\\
&\quad - \mu\partial_1 R\big(\partial_\alpha X, \partial_\alpha^2 X\big)\partial_\alpha U - \mu\partial_1 S\big(\partial_\alpha X, \partial_\alpha^3 X\big)\partial_\alpha U\\
&\quad - \mu Q(\partial_\alpha X)^{-3}\big((1 + \partial_\alpha X_1)\partial_\alpha^3(\mathscr{S}_1(X)U_1)\big)\big)\\
&\quad + 3\mu Q^{-5}\big((-f_2\partial_\alpha f_1 + f_1\partial_\alpha f_2)f_2\partial_\alpha^2(\mathscr{S}_1(X)U_1)\big)
\end{aligned}$$

https://doi.org/10.1017/S0956792503005163 Published online by Cambridge University Press

$$
\begin{aligned}
&\quad + 3\mu Q^{-5}\big((f_1\partial_\alpha f_1 + f_2\partial_\alpha f_2)f_1\partial_\alpha^2(\mathscr{S}_1(X)U_1)\big),\\
H_2\cdot &= f_2[\mathscr{K}_0, f_1]\cdot - f_2\mathscr{K}_0[\mathscr{K}_0, f_2]\cdot + f_2\big(1+\mathscr{K}_0^2\big)(f_2\cdot),\\
H_1\cdot &= f_2[\mathscr{K}_0, f_1]\cdot - f_2\mathscr{K}_0[\mathscr{K}_0, f_2]\cdot + f_2\big(1+\mathscr{K}_0^2\big)(f_2\cdot) - (f_1 - f_2\mathscr{K}_0)f_2\mathscr{S}_1(X)\cdot,\\
h_1 &= 3\mu Q(\partial_\alpha X)^{-5},\\
h_0 &= -\mu Q(\partial_\alpha X)^{-3} = -\mu\big(f_1^2 + f_2^2\big)^{-3/2},\\
a_2 &= f_1\partial_\alpha f_1 + f_2\partial_\alpha f_2,\\
a_1 &= (-f_2\partial_\alpha f_1 + f_1\partial_\alpha f_2),\\
f_3 &= f_1^2 + f_2^2,\\
f_2 &= \partial_\alpha X_2,\\
f_1 &= (1 + \partial_\alpha X_1).
\end{aligned}
$$

The key fact, which the patient reader can verify without difficulty using these formulas and the estimates on $\mathscr{K}$ from Appendix A is that G_5 is quadratic, and semi-linear, that is, if $\mathscr{W}_e \in \mathscr{H}_e^s$, then for any $R > 0$ there exists C_R such that if $\|\mathscr{W}_e\|_{\mathscr{H}_e^s} \leqslant R$,

$$\|G_5\|_{H^{s-3}} \leqslant C_R \|\mathscr{W}_e\|^2_{\mathscr{H}_e^s}.$$

References

[1] Alber, M., Luther, G., Marsden, J., & Robbins, J. (1998) Geometric phases, reduction and Lie-Poisson structure for the resonant three-wave interaction. *Physica D*, **123**(1–4), 271–290.

[2] Ablowitz, M. J. & Segur, H. (1981) Solitons and the inverse scattering transform. *SIAM Stud. Appl. Math.* **4**.

[3] Ball, K. F. (1964) Energy transfer between external and internal gravity waves *J. Fluid Mech.* **19**, 465–478.

[4] Beals, R. & Coifman, R. R. (1984) Scattering and inverse scattering for first order systems. *Comm. Pure Appl. Math.* **37**, 39–90.

[5] Beals, R. & Sattinger, D. (1991) On the complete integrability of completely integrable systems. *Comm. Math. Phys.* **138**, 409–436.

[6] Benney, D. J. & Newell, A. C. (1967) The propagation of nonlinear wave envelopes *J. Math. Phys.* **46**, 133–139.

[7] Craig, W. (1985) An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits. *Commun. Part. Diff. Eqns.* **10**, 787–1003.

[8] Craik, A. D. D. (1985) *Wave interactions and fluid flows.* Cambridge University Press.

[9] Hammack, J. L. & Henderson, D. M. (1993) Resonant interactions among surface water waves. *Ann. Rev. Fluid Mech.* **25**, 55–97.

[10] Janssen, P. (1986) The period-doubling of gravity-capillary waves. *J. Fluid Mech.* **172**, 531–545.

[11] Janssen, P. (1987) The initial evolution of gravity-capillary waves. *J. Fluid Mech.* **184**, 581–597.

[12] Kano, T. & Nishida, T. (1979) Sur les ondes de surface de l'eau avec une justification mathematique des equations des ondes en eau profonde. *J. Math. Kyoto Univ.* **19**, 335–370.

[13] Kaup, D. J. (1980) A method for solving the separable initial-value problem of the full three-dimensional three-wave interaction. *Stud. Appl. Math.* **62**, 75–83.

[14] Kawai, T. (1979) Generation of initial wavelets by instability of a coupled shear flow and their evolution to wind waves. *J. Fluid Mech.* **93**, 661–703.

[15] Kirrmann, P., Schneider, G. & Mielke, A. (1992) The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. Roy. Soc. Edin. A*, **122**, 85–91.

[16] Longuet-Higgins, M. S. & Smith, N. D. (1966) An experiment on third order resonant wave interactions. *J. Fluid Mech.* **25**, 417–435.

[17] McGoldrick, L. F. (1965) Resonant interactions among capillary-gravity waves. *J. Fluid Mech.* **21**, 305–331.

[18] Nalimov, V. I. (1974) The Cauchy-Poisson problem. *Dynamika Splosh. Sredy*, **18**, 104–210 (in Russian).

[19] Phillips, O. M. (1960) On the dynamics of unsteady gravity waves of finite amplitude. Part I. The elementary interactions. *J. Fluid Mech.* **9**, 193–217.

[20] Schneider, G. (1995) Validity and Limitation of the Newell-Whitehead equation. *Mathematische Nachrichten*, **176**, 249–263.

[21] Schneider, G. & Wayne, C. E. (2000) The long wave limit for the water wave problem. I. the case of zero surface tension. *Comm. Pure Appl. Math.* **53**(12), 1475–1535.

[22] Schneider, G. & Wayne, C. E. (2002) The rigorous approximation of long-wavelength capillary-gravity waves. *Arch. Rat. Mech. and Anal.* **162**, 247–285.

[23] Shinbrot, M. (1976) The initial value problem for surface waves under gravity, I. The simplest case. *Indiana U. Math. J.* **25**, 281–300.

[24] Wu, S. (1997) Well-posedness in Sobolev-spaces of the full water wave problem in 2D. *Invent. Math.* **130**, 39–72.

[25] Wu, S. (1999) Well-posedness in Sobolev-spaces of the full water wave problem in 3D. *J. Am. Math. Soc.* **12**, 445–495.

[26] Yosihara, H. (1982) Gravity waves on the free surface of an incompressible perfect fluid of finite depth. *RIMS Kyoto*, **18**, 49–96.

[27] Yosihara, H. (1983) Capillary-gravity waves for an incompressible ideal fluid. *J. Math. Kyoto Univ.* **23**, 649–694.