DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **33**, Number **2**, February **2013**

doi:10.3934/dcds.2013.33.599

pp. **599–628**

PROPAGATION OF LONG-CRESTED WATER WAVES

Jerry L. Bona

Department of Mathematics, Statistics and Computer Science
The University of Illinois at Chicago
851 S. Morgan St. MC 249
Chicago, IL 60607, USA

Thierry Colin

Université de Bordeaux, Institut de Mathématiques de Bordeaux
UMR CNRS 5251 and INRIA Bordeaux Sud-Ouest
351 cours de la Libération
33405 Talence, France

Colette Guillopé

Université Paris-Est
Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050)
UPEC, UPEMLV, CNRS
94010 Créteil, France

Abstract. The present essay is concerned with a model for the propagation of three-dimensional, surface water waves. Of especial interest will be long-crested waves such as those sometimes observed in canals and in near-shore zones of large bodies of water. Such waves propagate primarily in one direction, taken to be the $x-$direction in a Cartesian framework, and variations in the horizontal direction orthogonal to the primary direction, the $y-$direction, say, are often ignored. However, there are situations where weak variations in the secondary horizontal direction need to be taken into account.

Our results are developed in the context of Boussinesq models, so they are applicable to waves that have small amplitude and long wavelength when compared with the undisturbed depth. Included in the theory are well-posedness results on the long, Boussinesq time scale. As mentioned, particular interest is paid to the lateral dynamics, which turn out to satisfy a reduced Boussinesq system. Waves corresponding to disturbances which are localized in the $x-$direction as well as bore-like disturbances that have infinite energy are taken up in the discussion.

1. **Introduction.** The present study is concerned with surface water waves. Of particular interest will be long-crested waves whose propagation is primarily along one direction, say the $x-$coordinate in a standard $xyz-$Cartesian coordinate system in which the vertical coordinate z increases in the direction opposite to that in which gravity acts. A three-dimensional theory is needed, as variations in the $y-$direction

2010 *Mathematics Subject Classification.* Primary: 35A01, 35A02, 35A09, 35Q35, 35Q51; Secondary: 35Q86, 45G15, 76B03, 76B15, 76B25.

Key words and phrases. Long-crested water waves, Boussinesq systems, bore propagation, initial-boundary-value problems.

are allowed. It is presumed, however, that the variations in the $y-$direction subside as y goes to $\pm\infty$, so that at least formally, a two-dimensional description is appropriate there.

Equations for the temporal evolution of such disturbances will be developed in a spatial domain unbounded in both the $x-$ and the $y-$directions, though a domain with limited extent in the $y-$direction can also be countenanced.

More precisely, we consider a layer of incompressible, irrotational, perfect fluid of undisturbed depth h_0 resting upon a horizontal, featureless bottom represented by the plane

$$\Big\{(x,y,z):\ z=-h_0\Big\}.$$

A typical kind of disturbance of the quiescent system, in which there is no motion whatever, and which fits within the framework of our theory, is a line solitary wave (a traveling wave of elevation that is uniform in the $y-$direction) whose depth structure is simply

$$h(x,y,t)=\varphi_c(x-ct)+h_0, \tag{1}$$

where the speed of propagation $c>0$, say, is a fixed constant. The dependent variable $h(x,y,t)$ is the height of the water column at the point (x,y) at time t. As usual, we let $\eta(x,y,t)=h(x,y,t)-h_0$ be the deviation of the free surface from its rest position and presume that η is a single-valued function of $(x,y,t)\in\mathbb{R}^2\times\mathbb{R}_+$. Of course, unlike h, the deviation η need not be positive to make physical sense.

The theory is designed to allow for disturbances which are more complex than those appearing in (1), disturbances whose initial structure might be

$$\eta_0(x,y)=\eta(x,y,0)=\varphi_c(x)+\psi(x,y), \tag{2}$$

where ψ is not necessarily small, but we might demand that $\psi(x,y)\to 0$ as (x,y) becomes unbounded. Another possibility that lies within the scope of our theory is a line solitary wave modulated in the $y-$direction, *viz.*

$$\eta_0(x,y)=\eta(x,y,0)=\varphi_{c(y)}(x)+\psi(x,y), \tag{3}$$

where ψ is as above, and $c(y)\to c_\pm$ as $y\to\pm\infty$. Here, the constants c_+ and c_- need not be equal.

An early effort at modeling such disturbances in case the initial condition $\eta_0(x,y)$ varies very slowly in the $y-$direction was introduced by Kadomtsev and Petviashvili [18] in their study of the stability of solitary waves to transverse perturbations. Commentary on their model equation and its relation to the water-wave problem, together with suggested improvements of the model to remove an artificial zero-mass condition, can be found in the recent work of Lannes and Saut [20] and Molinet, Saut and Tzvetkov [22].

The wave motion is presumed to fit within the Boussinesq regime, and consequently a Boussinesq system will comprise the governing equations. The study of such systems has been developed by many authors (see Bona, Chen and Saut [7, 8] for a collection of references). Rigorous comparisons with the full Euler equations for the flow of a perfect fluid are also available in Bona, Colin and Lannes [9], while long time existence for the full Euler equations is available in Alvaret-Samaneigo and Lannes [3]. What particularly distinguishes the present work from these previous efforts is the non-trivial behavior of solutions as $|(x,y)|\to+\infty$.

Related theory and some telling numerical simulations have also been worked out by Dougalis, Mitsotakis and Saut [15, 16]. Their theoretical development is set in Sobolev spaces and does not allow for motions that do not evanesce in all directions.

The plan of the paper is the following. Section 2 is devoted to preliminaries, including a precise formulation of the problem. The original initial-boundary-value problem is recast as an integral equation in Section 3 and local well-posedness is established in a variety of function spaces. While Section 4 deals with aspects of the asymptotic behavior of solutions as x or y becomes unboundedly large, Section 5 is concerned with extending the local well-posedness theory to the longer, Boussinesq time scale. This latter theory is developed in the more general context of d coupled equations in n spatial variables. The water-wave problem considered here then falls out as an example of the application of this theory. The body of the paper closes with a summary and a perspective for future developments.

2. **Preliminaries and formulation of the problem.** We commence with a brief indication of the notation in force hereafter, most of which is standard.

2.1. **Notation.** Derivatives with respect to spatial or temporal variables are designated by subscripts x, y, z or t, and also, when convenient, by $\partial_x, \partial_y, \partial_{x_i}, 1 \le i \le n$, or ∂_t. The differential operators Δ, ∇ and D^k for $k \ge 0$ are always taken with respect to the spatial variables (x and y if $n = 2$, $x_1, \cdots, x_n$ for general values of n). We also use the standard multi-index notation ∂^γ, $\gamma \in \mathbb{Z}_+^n$, for n–variable partial derivatives, where $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ is the non-negative integers.

Except for the abbreviations noted below, the norm of an element f in a Banach space X is denoted $\|f\|_X$. If Ω is a measurable set in $\mathbb{R}^n$, $n \ge 1$ a given integer, $L_p(\Omega)$ is the Lebesgue space of p^{th}–power integrable functions on Ω if $1 \le p < \infty$, with the usual modification if $p = \infty$. When Ω is understood from the context, the $L_p(\Omega)$–norm of a function or of a vector-valued function f is written simply $|f|_p$. (If $f = (f_1, \cdots, f_n)$ is a vector-valued function of x, say, then $|f|_p = |f_1|_p + \cdots + |f_n|_p$ where $|f_j|_p$ is the usual L_p–norm of the real-valued function $f_j, j = 1, \cdots, n$.) If the context is in doubt, we write $|f|_{L_p(\Omega)}$.

If $k \ge 0$ is an integer, $W_p^k(\Omega)$ is the subspace of functions f in $L_p(\Omega)$, whose distributional partial derivatives $\partial^\gamma f$ also lie in $L_p(\Omega)$ for all multi-indices γ with $|\gamma| \le k$, with its usual norm

$$\|f\|_{W_p^k(\Omega)} = \sum_{|\gamma| \le k} |\partial^\gamma f|_{L_p(\Omega)} = \|f\|_{W_p^k},$$

the right-hand notation being preferred when Ω is understood from context. Spaces of vector-valued functions with components in $W_p^k(\Omega)$ will also be considered, and the same notation will be used for the norm of such a vector-valued function, which is simply the sum of the W_p^k–norms of its components. The L_2–based spaces appear frequently and the norm of f in $H^k(\Omega) = W_2^k(\Omega)$ is abbreviated to $\|f\|_k$ if Ω is clearly delineated from context. The space $\mathcal{C}_b(\mathbb{R}^n)$ is the collection of bounded, continuous functions on $\mathbb{R}^n$ with the $L_\infty(\mathbb{R}^n)-$norm, while $\mathcal{C}_0(\mathbb{R}^n)$ is the subset comprised of functions which are null at infinity, *i.e.* $f \in \mathcal{C}_0(\mathbb{R}^n)$ means that f is everywhere continuous and $\lim_{|x|\to+\infty} f(x) = 0$. The space $\mathcal{C}_0(\mathbb{R}^n)$ is a closed linear subspace of the Banach space $\mathcal{C}_b(\mathbb{R}^n)$. For an integer $k \ge 0$, $\mathcal{C}_b^k(\mathbb{R}^n)$ is the subspace of f in $\mathcal{C}_b(\mathbb{R}^n)$ such that $\partial^\gamma f \in \mathcal{C}_b(\mathbb{R}^n)$ for all multi-indices γ with $|\gamma| \le k$. Similarly $\mathcal{C}_0^k(\mathbb{R}^n)$ are those elements f of $\mathcal{C}_0(\mathbb{R}^n)$ such that $\partial^\gamma f \in \mathcal{C}_0(\mathbb{R}^n)$ for all multi-indices γ with $|\gamma| \le k$. These spaces carry their usual norms, namely

$$\|f\|_{\mathcal{C}_b^k(\mathbb{R}^n)} = \|f\|_{\mathcal{C}_0^k(\mathbb{R}^n)} = \sum_{|\gamma| \le k} |\partial^\gamma f|_\infty .$$

Use will also be made of the Hölder spaces $\mathcal{C}_b^\mu(\mathrm{I\!R}^n)$ and $\mathcal{C}_b^{k+\mu}(\mathrm{I\!R}^n)$ for integers $k > 0$ and $0 < \mu < 1$. The latter is the space of $f \in \mathcal{C}_b^k(\mathrm{I\!R}^n)$ such that if γ is a multi-index with $|\gamma| = k$, then $\partial^\gamma f \in \mathcal{C}_b^\mu(\mathrm{I\!R}^n)$ where $\mathcal{C}_b^\mu(\mathrm{I\!R}^n) \subset \mathcal{C}_b(\mathrm{I\!R}^n)$ are those functions for which

$$[f]_\mu = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\mu} < +\infty.$$

Spaces that single out the temporal variable will also appear. If $T > 0$ and if Z is a Banach space, the Banach space $L_p(0, T; Z)$ is the space of measurable mappings $u : [0, T] \longrightarrow Z$ such that $\|u(t)\|_Z$ is in $L_p(0, T)$, with the obvious norm. The closed subspace of $L_\infty(0, T; Z)$ of continuous mappings is denoted $\mathcal{C}([0, T]; Z)$. If $k \geq 0$ is an integer, $\mathcal{C}^k([0, T]; Z)$ are those functions u such that the $Z-$valued distributional derivative $\partial_t^j u$ lies in $\mathcal{C}([0, T]; Z)$, for all j with $0 \leq j \leq k$, with the norm

$$\|u\|_{\mathcal{C}^k([0,T];Z)} = \sum_{j=0}^{k} ||\partial_t^j u||_{\mathcal{C}([0,T];Z)}\,.$$

It will be convenient of use the abbreviations

$$\mathcal{X}_T = \Big(\mathcal{C}([0, T]; \mathcal{C}_b(\mathrm{I\!R}^2))\Big)^3, \quad \mathcal{X}_{T,0} = \Big(\mathcal{C}([0, T]; \mathcal{C}_0(\mathrm{I\!R}^2))\Big)^3,$$

and for $\mu \in (0, 1)$,

$$\mathcal{X}_T^\mu = \Big(\mathcal{C}([0, T]; \mathcal{C}_b^\mu(\mathrm{I\!R}^2))\Big)^3, \quad \mathcal{X}_{T,0}^\mu = \Big(\mathcal{C}([0, T]; \mathcal{C}_0^\mu(\mathrm{I\!R}^2))\Big)^3;$$

also, for integers $k, l \geq 0$,

$$\mathcal{X}_T^{k,l} = \Big(\mathcal{C}^k([0, T]; \mathcal{C}_b^l(\mathrm{I\!R}^2))\Big)^3,$$

and similarly for $\mathcal{X}_{T,0}^{k,l}$, and $\mathcal{X}_T^{k,l+\mu}$, for $0 \leq \mu < 1$.

2.2. **The central problem.** As mentioned at the outset, a homogeneous layer of perfect fluid of depth h_0 is presumed to be resting on the plane $\{(x, y, z) : \ z = -h_0\}$. It is assumed that the wave motion resulting from a disturbance of the equilibrium has a resulting free surface that is a graph over the flat bottom. In this circumstance, the free surface may be described by the function $\eta = \eta(x, y, t)$ as indicated already in Section 1. With the additional assumptions that the fluid is incompressible (a good assumption for air or water in ordinary circumstances) and the flow irrotational (a presumption that is often a good one on large scales), a classical mathematical formulation of the water-wave problem is the system

$$\left\{\begin{array}{lcll} \beta\Delta\phi + \phi_{zz} & = & 0 & \text{in } \{-1 \leq z \leq \alpha\eta\}, \\ \phi_z & = & 0 & \text{at } \{z = -1\}, \\ \eta_t + \alpha\nabla\phi \cdot \nabla\eta & = & \dfrac{1}{\beta}\phi_z & \text{on } \{z = \alpha\eta\}, \\ \phi_t + \dfrac{1}{2}\Big(\alpha|\nabla\phi|^2 + \dfrac{\alpha}{\beta}(\phi_z)^2\Big) + \eta & = & 0 & \text{on } \{z = \alpha\eta\}, \end{array}\right. \tag{4}$$

where Δ and ∇ are the obvious differential operators with respect to the variables x and y. The variables in these equations have been scaled using the scheme

$$\tilde{x} = \ell x, \ \tilde{y} = \ell y, \ \tilde{z} = h_0 z, \ \tilde{\eta} = A\eta, \ \tilde{t} = \frac{\ell}{c_0} t\,,$$

where those adorned with a tilde are the original, dimensional quantities, $A = \max_{x,y,t} |\tilde{\eta}|$ is the maximum amplitude encountered in the wave motion, ℓ is the smallest wavelength for which the flow has significant energy and $c_0 = \sqrt{gh_0}$ is the kinematic wave velocity, with g the gravity constant. The unknown function $\phi = \phi(x, y, z, t)$ is the velocity potential, whose existence follows from incompressibility and irrotationality, and which is scaled via $\tilde{\phi} = \ell g A\phi/c_0$. The velocity field U is therefore given by $U = (\nabla\phi, \phi_z) = (\phi_x, \phi_y, \phi_z)$ where as above, ∇ denotes the gradient operator in the x–y variables.

The Boussinesq regime is characterized by the parameters

$$\alpha = \frac{A}{h_0} \quad \text{and} \quad \beta = \Big(\frac{h_0}{\ell}\Big)^2$$

where A and ℓ are as above. Assume now that both α and β are small compared to one, and that the Stokes number $S = \alpha/\beta$ is of order one. Of course, this can be imposed upon the initial data for the problem, but it must also be presumed that it continues to hold as the wave evolves in time. That such a presumption can be inferred from conditions on the intitial data is a consequence of the recent work [3]. In the circumstances just delineated, a formal expansion of the velocity potential in the vertical coordinate, followed by ignoring all terms of quadratic order or higher in the quantities α and β, leads to the set of *abcd*–systems (coupled systems of three nonlinear evolution equations, see [7, 8]),

$$\begin{cases} V_t + \nabla\eta + \dfrac{\alpha}{2}\nabla|V|^2 + \beta\Big(a\,\Delta\nabla\eta - b\,\Delta V_t\Big) & = 0, \\ \eta_t + \nabla\cdot V + \alpha\nabla\cdot(\eta V) + \beta\Big(c\,\Delta\nabla\cdot V - d\,\Delta\eta_t\Big) & = 0. \end{cases} \tag{5}$$

The coefficients a, b, c and d are

$$a = \frac{1-\theta^2}{2}\mu, \quad b = \frac{1-\theta^2}{2}(1-\mu), \quad c = \Big(\frac{\theta^2}{2} - \frac{1}{6}\Big)\lambda, \quad d = \Big(\frac{\theta^2}{2} - \frac{1}{6}\Big)(1-\lambda),$$

where λ and μ are real parameters that, formally, may be chosen without restriction, and θ lies in the interval $[0, 1]$. The dependent variable $z = \eta(x, y, t)$ is the deviation of the free surface from its rest position $(x, y, 0)$ at the time t, as already discussed. (Thus the free surface lies at the point $(x, y, \eta(x, y, t))$ at time t for all $(x, y) \in \mathbb{R}^2$.) The variable $V = V^\theta = (u^\theta, v^\theta)$ is the horizontal velocity field at the height θ above the bottom.

Solutions of the well-posed subclass of these systems are approximations of the solutions of the Euler system. Indeed, they provide direct approximations of the deviation of the free surface and of the horizontal velocity field $V = V^\theta$ at the height θ above the bottom (at the vertical coordinate $z = \theta - 1$), where θ has a fixed value (again, with $0 \le \theta \le 1$, since the scaled height is measured in depths). A small additional calculation using the formula

$$V^\sigma(x, y, t) = \Big(1 - \frac{(1-\theta)^2 - (1-\sigma)^2}{2}\beta^2\Delta\Big)V^\theta(x, y, t) \tag{6}$$

yields an approxmiation to the horizontal velocity field at the height σ above the bottom. At the Boussinesq level of approximation, the vertical velocity is quadratic in the small parameters α and β, and so ignored.

From the perspective of the practical use of these types of systems, the most convenient choice is to take $\theta = \sqrt{2/3}$ and $\lambda = \mu = 0$ so that (5) reduces to

$$\begin{cases} \eta_t + \nabla\cdot V + \alpha\nabla\cdot(\eta V) - \dfrac{\beta}{6}\,\Delta\eta_t = 0, \\ V_t + \nabla\eta + \dfrac{\alpha}{2}\nabla|V|^2 - \dfrac{\beta}{6}\,\Delta V_t = 0, \end{cases} \tag{7}$$

where $V = V\sqrt{2/3}$. This is the so-called BBM–BBM Boussinesq system (see e.g. [2], [4], [16]). Some reasons why this is a good choice among the three-parameter family (5) is the ease with which non-homogeneous boundary conditions can be imposed and accurate numerical schemes devised (see *e.g.* the discussions in [4], [5], [10] and [6]). The zeroes on the right-hand side are in reality the terms in the formal expansion of the original variables that are neglected in coming to the Boussinesq approximation. These terms are of second order, which is to say, of order α^2, $\alpha\beta$ and β^2. (If the Stokes number $S = 1$, then of course all these terms are identical.) A simple rescaling of V, η, (x, y) and t allows one to dispense with the parameters α and β and the values $1/6$ appearing above. Performing these changes of variables and writing V in terms of its components, $V = (u(x, y, t), v(x, y, t))$, the system (7) satisfied by (η, u, v) is, in detail,

$$\begin{cases} \eta_t + u_x + v_y + (\eta u)_x + (\eta v)_y - \eta_{xxt} - \eta_{yyt} & = & 0, \\ u_t + \eta_x + uu_x + vv_x - u_{xxt} - u_{yyt} & = & 0, \\ v_t + \eta_y + uu_y + vv_y - v_{xxt} - v_{yyt} & = & 0, \end{cases} \tag{8}$$

posed in $\mathbb{R}^2 \times \mathbb{R}_+$, with initial conditions

$$\eta(x, y, 0) = \eta_0(x, y), \quad u(x, y, 0) = u_0(x, y), \quad v(x, y, 0) = v_0(x, y),$$

say, defined for $(x, y) \in \mathbb{R}^2$.

The behavior at infinity that captures the type of wave motion in view here is that the free surface is asymptotically constant in the $x-$direction and that variations with respect to the $y-$variable vanish in the limit of large $|y|$. For example, we might ask that

$$\begin{cases} \eta \to 0, & (u, v) \to (0, 0), & \text{as } x \to \pm\infty, \\ v \to 0, & \partial_y \to 0, & \text{as } y \to \pm\infty, \\ \eta \to \eta^\pm, & u \to u^\pm, & \text{as } y \to \pm\infty, \end{cases} \tag{9}$$

where the functions $\eta^\pm = \eta^\pm(x, t)$ and $u^\pm = u^\pm(x, t)$ will turn out to be solutions to the reduced systems,

$$\begin{cases} (\eta^\pm)_t + (u^\pm)_x + (\eta^\pm u^\pm)_x - (\eta^\pm)_{xxt} & = 0, \\ (u^\pm)_t + (\eta^\pm)_x + u^\pm u^\pm_x - (u^\pm)_{xxt} & = 0, \end{cases} \tag{10}$$

set in $\mathbb{R} \times \mathbb{R}_+$, with initial conditions

$$\eta^\pm(x, 0) = \eta_0^\pm(x) = \lim_{y\to\pm\infty} \eta_0(x, y), \quad u^\pm(x, 0) = u_0^\pm(x) = \lim_{y\to\pm\infty} u_0(x, y), \tag{11}$$

for $x \in \mathbb{R}$. Appropriate compatibility conditions on the auxiliary data are that

$$\begin{cases} \eta_0(x, y), \quad u_0(x, y), \quad v_0(x, y) \to 0, & \text{as } x \to \pm\infty, \\ \eta_0^\pm(x), \quad u_0^\pm(x) \to 0, & \text{as } x \to \pm\infty. \end{cases} \tag{12}$$

This corresponds to a disturbance that is long-crested, but localized in the $x-$direction. We could equally well ask that

$$\begin{cases} \eta, \eta^\pm \to 0 \ \text{ as } \ x \to +\infty, \\ \eta, \eta^\pm \to 1 \ \text{ as } \ x \to -\infty, \end{cases}$$

with all the other asymptotic behaviors as in (9), a specification which corresponds to bore propagation.

The reduced system (10) is an approximation of the two-dimensional Euler equations which has been studied at some length in Bona and Chen [6]. Local well-posedness of this system has been established, and if the data are regular, regularity theory of solutions has been developed as well. More details follow in the next section. The system (8)–(12) approximates the three-dimensional Euler equations for the specific flow problem considered here; in particular $\eta(x, y, t)$ yields the approximation $h(x, y, t) = \eta(x, y, t) + 1$ of the scaled water depth at the point (x, y) at time t, whilst $(u(x, y, t), v(x, y, t))$ describes the scaled horizontal velocity at the point $(x, y, \sqrt{2/3} - 1)$ in the fluid domain at time t. Approximations of the velocity field to the same accuracy at other depths are easily obtained using the formula (6) (see again [7, 8]).

3. **Local well-posedness and regularity.** We begin by pointing out an interesting fact about the system (8). If for some $T > 0$, (η, u, v) is a solution in $\mathcal{X}_T$ in the sense of distributions that has initial value $(\eta_0, u_0, v_0) \in \mathcal{C}_b(\mathbb{R}^2)^3$, then it is unique within $\mathcal{X}_T$. Moreover, as will appear in Section 4, the behavior as $|(x, y)| \to +\infty$ of the solution (η, u, v) emanating from (η_0, u_0, v_0) is completely determined by the behavior at infinity of (η_0, u_0, v_0).

The first task is to indicate the validity of the uniqueness assertion and then, armed with this information, proceed to the main development. For some $T > 0$ let $W = (\eta, u, v) \in \mathcal{X}_T$ and write the system (8) in the form

$$\Big(I - \Delta\Big) W_t = F(W), \tag{13}$$

where $F(W) = (F_\eta(W), F_u(W), F_v(W))$ is a vector-valued function whose components are first-order partial derivatives, with respect to x and y, of quadratic polynomials in η, u and v. Since both W and the right-hand side of (13) have components that are tempered distributions, the Fourier transform in the spatial variables makes sense and it is determined that

$$(1 + k^2 + m^2)\hat{W}_t = \hat{F} = ik\hat{G} + im\hat{H},$$

where the circumflex connotes the Fourier transform taken componentwise, (k, m) are the variables dual to (x, y) and $F = \partial_x G + \partial_y H$. A calculation reveals that

$$W_t = K_x \star G(W) + K_y \star H(W), \tag{14}$$

where the convolution is applied componentwise and $K(x, y) = \frac{1}{2\pi} K_0(\sqrt{x^2 + y^2})$ with K_0 being the zeroth-order Bessel function of the third type, also known as the Macdonald function. It is well known that, with its usual normalization (see [1]), $K_0(z)$ is an even function of z, monotone decreasing for $z > 0$ and such that

$$\begin{aligned}
&K_0(z) \text{ is a } \mathcal{C}^\infty\text{-function, except at } z = 0,\\
&K_0(z) \sim \sqrt{\frac{\pi}{2|z|}}\, e^{-|z|} \text{ as } z \to \pm\infty,\\
&K_0(z) \sim -\ln|z| \text{ as } z \to 0,\\
&K_0' \text{ decreases exponentially to 0 at } \pm\infty \text{ and}\\
&K_0'(z) \sim -\frac{1}{z} \text{ as } z \to 0.
\end{aligned} \tag{15}$$

From the properties (15) it is seen at once that both K_x and K_y lie in $L_p(\mathbb{R}^2)$ for $1 \le p < 2$. Hence, the distributional convolutions in (14) are in fact classical. As the right-hand side of (14) lies in $\mathcal{X}_T$, it follows that W lies in $\mathcal{X}_T^{1,0}$. Integrating (14) over $[0,t]$ for $t \le T$ allows us to adduce the formula

$$W(t) = W(0) + \int_0^t \Big(K_x \star G(W(s)) + K_y \star H(W(s)) \Big)\, ds = \mathcal{A}(W)(t) \tag{16}$$

that W must satisfy.

If the right-hand side of (16) is viewed as an operator which maps $\mathcal{X}_T$ into itself, then it is a straightforward consequence of the Minkowski integral inequality that if $V, W \in \mathcal{X}_T$, then

$$\|\mathcal{A}(V) - \mathcal{A}(W)\|_{\mathcal{X}_T} \le CT\Big(1 + \|V\|_{\mathcal{X}_T} + \|W\|_{\mathcal{X}_T}\Big)\|V - W\|_{\mathcal{X}_T}\,, \tag{17}$$

where the constant C may be taken to be the maximum of the $L_1(\mathbb{R}^2)$–norms of K_x and K_y, both of which have the value

$$\int_{\mathbb{R}^2} \Big|\partial_x K(x,y)\Big|\, dx\, dy = \frac{1}{2\pi}\int_{\mathbb{R}^2} \Big|\partial_x K_0\Big(\sqrt{x^2+y^2}\Big)\Big|\, dx\, dy = \frac{2}{\pi}\int_0^\infty K_0(r)\, dr = 1$$

(see [17], p. 736). Because of this estimate, if R and T' are chosen so that

$$R = 2|W(0)|_\infty \quad \text{and} \quad T' \le \min\Big\{\frac{1}{2(1+2R)}, T\Big\}, \tag{18}$$

then $\mathcal{A}$ is a contraction mapping of the ball B_R of radius R about the origin in the space $\mathcal{X}_{T'} = \mathcal{C}([0,T'];\mathcal{C}_b(\mathbb{R}^2))^3$.

Similar considerations reveal that a solution W of (13) in $\mathcal{X}_T$ is, near any point $t \in (0,T)$, the fixed point of a contraction mapping in a closed ball about the origin in a space of the form $C([t_0,t_1];\mathcal{C}_b(\mathbb{R}^2))^3$ for suitable values of t_0 and t_1 with $t_0 < t < t_1$.

From these observations, several things follow at once. These are summarized in the following Theorem.

Theorem 3.1. *(i) For any given $W(0) = (\eta_0, u_0, v_0) \in \mathcal{C}_b(\mathbb{R}^2)^3$, there exists a solution $W = (\eta, u, v) \in \mathcal{X}_{T'}$ of the integral equation (16) where $T' = T'(||W(0)||_{\mathcal{C}_b(\mathbb{R}^2)^3})$ is as in (18).*

(ii) Any solution $W \in \mathcal{X}_T$ of the integral equation (16) must lie in $\mathcal{X}_T^{1,0}$. In fact, for any μ with $0 \le \mu < 1$, and any integer k, W_t lies in $\mathcal{X}_T^{k,\mu}$.

(iii) Any solution of the integral equation (16) in $\mathcal{X}_T$ is a distributional solution of the initial-value problem (8), and conversely, any distributional solution of (8) in $\mathcal{X}_T$ is a solution of (16) on its time interval of existence.

(iv) Solutions of (8) or (16) are unique in $\mathcal{X}_T$ and they are given locally near any time t_0 as the fixed point of a contraction mapping of a ball centered at the origin in $\mathcal{C}([t_0, t_0+\delta];\mathcal{C}_b(\mathbb{R}^2))^3$ for small enough values of δ.

(v) A solution $W \in \mathcal{X}_T$ depends continuously upon its initial value in $\mathcal{C}_b(\mathbb{R}^2)^3$.

Proof. (i) Local existence for (16) follows from the contraction mapping principle as already indicated.

(ii) The fact that W is $\mathcal{C}^1$ in time and $W_t \in \mathcal{X}_T$ is clear from (16). In fact, since $(I-\Delta)^{-1}$ maps $\mathcal{C}_b(\mathbb{R}^2)$ to $\mathcal{C}_b^{1+\mu}(\mathbb{R}^2)$ for any $\mu \in [0,1)$, the further regularity is also clear. Since W_t lies in $\mathcal{X}_T^\mu$, it follows that the right-hand side of (14) is differentiable with respect to t with values in $\mathcal{C}_b^\mu(\mathbb{R})^3$, and, moreover,

$$W_{tt} = \Big(K_x \star G'(W) + K_y \star H'(W)\Big) W_t\,. \tag{19}$$

It further follows that the right-hand side of (19) is differentiable with respect to the temporal variable t. The obvious inductive step then establishes that W is t –differentiable to all orders[1].

(iii) The derivation of (16) from (8) in the context of solutions in $\mathcal{X}_T$ was shown above. The other way around is also clear from the properties of the kernel.

(iv) Because of (iii) and the fact that, whatever the value of $W(t_0)$, the analogue of $\mathcal{A}$ starting at t_0, namely

$$\mathcal{A}_{t_0}(W)(t) = W(t_0) + \int_{t_0}^{t} \Big(K_x \star G(W(s)) + K_y \star H(W(s)) \Big)\, ds\,,$$

is a contraction on a ball about the origin in $\mathcal{C}([t_0,t_1];\mathcal{C}_b(\mathbb{R}^2))^3$ for $t_1 > t_0$ close enough to t_0, it follows that solutions are, locally in time, the fixed points of a contraction mapping. It is therefore immediate that they are locally unique in $\mathcal{X}_T$ and hence unique as long as they exist.

(v) Since the operator $\mathcal{A}$ or the operators $\mathcal{A}_t$ depend continuously on the initial value, it follows that the solution map $W(0) \mapsto W$ from $\mathcal{C}_b(\mathbb{R}^2)^3$ to $\mathcal{X}_T$ is Lipschitz continuous for T small. Indeed, let $W(0)$ and $V(0)$ be two initial data and let W and V be the associated solutions in $\mathcal{X}_T$. Take T small enough so that, say, $\mathcal{A} = \mathcal{A}^{W(0)}$ (with an obvious notation) is a contraction on the ball B_R around the origin in $\mathcal{X}_T$, where R is large enough that both W and V, restricted to $[0,T]$, lie in B_R. Again because of (iii), $V = \mathcal{A}^{V(0)}(V)$, though $\mathcal{A}^{V(0)}$ need not be contractive on B_R. Then, notice that

$$\begin{aligned}
\|W - V\|_{\mathcal{X}_T} &= \|\mathcal{A}^{W(0)}W - \mathcal{A}^{V(0)}V\|_{\mathcal{X}_T} \\
&\leq \|\mathcal{A}^{W(0)}W - \mathcal{A}^{W(0)}V\|_{\mathcal{X}_T} + \|\mathcal{A}^{W(0)}V - \mathcal{A}^{V(0)}V\|_{\mathcal{X}_T} \\
&\leq \Theta\|W - V\|_{\mathcal{X}_T} + \|W(0) - V(0)\|_{\mathcal{C}_b},
\end{aligned}$$

where Θ is the contraction constant for $\mathcal{A}^{W(0)}$ on B_R. As the constant Θ is less than 1, the result follows.

This argument is local in t and can be applied near any point t_0 in the intersection of the existence intervals for the two solutions in question. If we fix attention on a particular solution U in $\mathcal{X}_T$, then it is easy to infer from the foregoing that all initial data near enough to $U(0)$ generate solutions whose time-interval of existence is at least $[0,T]$. $\square$

There are many corollaries and extensions of Theorem 3.1, some of which are enunciated below.

The first point to notice is that the uniqueness statement takes place in a pretty broad function class. This will be important presently when boundary behavior is discussed.

Another interesting point is that the same arguments will work for a wide variety of spatial function classes, which is to say, there is nothing exceptional about the choice of $\mathcal{C}_b(\mathbb{R}^2)$. Indeed, let $Z = Z(\mathbb{R}^2)$ be any Sobolev space of measurable real-valued functions defined on $\mathbb{R}^2$ that has the following properties:

1. Z is a Banach algebra, which means that whenever $f, g \in Z$, then $fg \in Z$, and there is a universal constant c_1 such that

$$\|fg\|_Z \leq c_1 \|f\|_Z \, \|g\|_Z\,;$$

[1] In fact, W is real analytic in t, see *e.g.* the argument in [4], Section 3.

2. convolutions with K_x and K_y define bounded linear operators on Z, so there is a constant c_2 such that

$$\|K_x \star f\|_Z \leq c_2\|f\|_Z \quad \text{and} \quad \|K_y \star f\|_Z \leq c_2\|f\|_Z\,,$$

for all $f \in Z$.

It is well known that Sobolev spaces $W_p^k(\mathbb{R}^2)$ satisfy Property 1 above if and only if they are embedded in $L_\infty(\mathbb{R}^2)$ (in which case, if $p < \infty$ or $k \geq 1$, they are subspaces of $\mathcal{C}_b(\mathbb{R}^2)$). Since the kernels K_x and K_y are L_1-functions, convolutions with them clearly map these spaces boundedly into themselves. For such spaces Z, most of Theorem 3.1 is still valid if $\mathcal{X}_T$ is replaced by $\mathcal{C}([0,T];Z)^3$ throughout. Thus, the local existence result via a contraction-mapping argument applied to the operator $\mathcal{A}$ follows from the same type of estimate as displayed in (17) except the constant C will now also depend on c_1 and c_2. Part (iii) follows since solutions in $\mathcal{C}([0,T];Z)^3$ lie in $\mathcal{C}([0,T];L_\infty(\mathbb{R}^2))^3$. Uniqueness in Part (iv) holds since it already holds in $\mathcal{C}([0,T];L_\infty(\mathbb{R}^2))^3$ and the argument that this unique solution is given locally as the fixed point of a contraction mapping is the same. This in turn implies local Lipschitz continuity of the solution map.

The outcome of this discussion is the following result, stated somewhat informally for brevity.

Theorem 3.2. *Let Z be a Banach space which is continuously embedded in $L_\infty(\mathbb{R}^2)$ and satisfies Properties 1 and 2 above. Then the results (i), (iii), (iv) and (v) of Theorem 3.1 hold when $\mathcal{X}_T$ is replaced throughout by $\mathcal{C}([0,T];Z)^3$.*

Remark 1. Condition 1 above can be replaced by the bilinear estimates

$$\|K_x \star (fg)\|_Z \leq c_1\|f\|_Z\|g\|_Z \quad \text{and} \quad \|K_y \star (fg)\|_Z \leq c_1\|f\|_Z\|g\|_Z$$

and the same conclusions drawn. Notice that Conditions 1 and 2 certainly imply these bilinear inequalities, but that the bilinear estimates are strictly weaker than the conjunction of the two conditions. This observation was used to good effect in the recent sharp well-posedness theory in [14] and it could be used to establish local well posedness in weaker spaces in the present context. Working this out would take us a bit out of our way, so this refinement is passed over here.

Corollary 1. *The initial-value problem (8) is locally well posed in $\mathcal{C}_b^k(\mathbb{R}^2)^3$ and $\mathcal{C}_0^k(\mathbb{R}^2)^3$, for $k = 0,1,2,\cdots$, in $\mathcal{C}_b^{k+\mu}(\mathbb{R}^2)^3$, $\mathcal{C}_0^{k+\mu}(\mathbb{R}^2)^3$, for $k = 0,1,2,\cdots$ and $0 \leq \mu < 1$, and in $W_p^k(\mathbb{R}^2)^3$ provided that $pk > 2$. Moreover the results (i), (iii), (iv) and (v) hold if $\mathcal{X}_T$ is replaced by $\mathcal{C}([0,T];\mathcal{C}_0^k(\mathbb{R}^2))^3$, or by similar spaces of time continuous vector-valued functions with values in one of the aforementioned spaces.*

For data in $\mathcal{C}_b^k(\mathbb{R}^2)^3$ or $\mathcal{C}_0^k(\mathbb{R}^2)^3$, the associated solution W has the property that $W_t \in \mathcal{C}^l([0,T],\mathcal{C}_b^{k+\mu}(\mathbb{R}^2))^3$ (respectively $\mathcal{C}^l([0,T],\mathcal{C}_0^{k+\mu}(\mathbb{R}^2))^3$) for any integer $l \geq 0$ and any μ with $0 \leq \mu < 1$. For data in $W_p^k(\mathbb{R}^2)^3$ with $kp > 2$, or in $\mathcal{C}_b^{k+\mu}(\mathbb{R}^2)^3$ for $0 < \mu < 1$, the solution satisfies $W_t \in \mathcal{C}^l([0,T],W_p^{k+1}(\mathbb{R}^2))^3$ or, respectively, $W_t \in \mathcal{C}^l([0,T];\mathcal{C}_b^{k+1+\mu}(\mathbb{R}^2))^3$, for any $l \geq 0$.

Remark 2. It is worth mention that if T_0 is the maximal existence time for a solution corresponding to some initial data $(\eta_0,u_0,v_0) \in \mathcal{C}_b^\mu(\mathbb{R}^2)^3$, for $\mu > 0$, and if it happens that (η_0,u_0,v_0) actually lies in $\mathcal{C}_b^{k+\mu}(\mathbb{R}^2)^3$ for some $k \geq 1$, then the solution emanating from this data lies in $\mathcal{C}([0,T_1];\mathcal{C}_b^{k+\mu}(\mathbb{R}^2))^3$ for any $T_1 < T_0$. This follows from a straightforward bootstrap argument based on the integral equation

(16), using the fact that convolution with K_x and K_y maps $\mathcal{C}_b^{j+\mu}(\mathbb{R}^2)$ continuously into $\mathcal{C}_b^{j+1+\mu}(\mathbb{R}^2)$ for any $j \geq 0$. This property is related to what T. Kato termed "propagation of regularity" (see Kato [19] and the more recent, related article [13]).

4. **Spatial asymptotics of solutions.** The initial values posited in Section 3 are not necessarily required to have any discernible structure in the far field where $|x| + |y|$ is large. In modeling real waves, it is often the case that natural far-field boundary behavior presents itself.

4.1. **Dirichlet conditions.** Suppose to be given initial data (η_0, u_0, v_0) for the system (8), each element of which lies in $\mathcal{C}_b(\mathbb{R}^2)$ as in Section 3. From the just developed theory, there is a solution (η, u, v) of (8) corresponding to this initial data. Suppose, in addition, that the data is such that it becomes one-dimensional in the transverse direction, which is to say

$$\eta_0^{\pm}(x) = \lim_{y\to\pm\infty} \eta_0(x,y), \quad u_0^{\pm}(x) = \lim_{y\to\pm\infty} u_0(x,y), \quad \lim_{y\to\pm\infty} v_0(x,y) = 0, \tag{20}$$

where $\eta_0^{\pm}$, $u_0^{\pm}$ lie in $\mathcal{C}_b(\mathbb{R})$. (We do not yet assume about η_0, u_0, v_0, $\eta_0^{\pm}$ and $u_0^{\pm}$ anything other than continuity and boundedness as $x \to \pm\infty$.)

Intuitively, it is expected that the wave motion emanating from such data will be two-dimensional as $y \to \pm\infty$ since it begins that way. If this is presumed to be the case in a strong sense, that includes $y-$derivatives tending to 0 as $y \to \pm\infty$, then a formal appraisal of the Boussinesq system (8) reveals that the third equation is satisfied identically and the first two equations simplify to

$$\begin{cases} \eta_t + u_x + (\eta u)_x - \eta_{xxt} = 0, \\ u_t + \eta_x + uu_x - u_{xxt} = 0, \end{cases} \tag{21}$$

as $y \to \pm\infty$. The theory in [6] assures that if the reduced system (21) is posed with initial data

$$\eta(x,0) = \eta_0^{\pm}(x) \quad \text{and} \quad u(x,0) = u_0^{\pm}(x),$$

then there are unique solutions $(\eta^+(x,t), u^+(x,t))$ and $(\eta^-(x,t), u^-(x,t))$ defined on some non-trivial time interval $[0,T]$, which have regularity properties corresponding to the regularity of the data, just as in the theorems for the two-dimensional case developed in Section 3.

Define the auxiliary functions

$$\begin{aligned} N^1(x,y,t) &= \tfrac{1}{2}\Big(\eta^+(x,t) + \eta^-(x,t)\Big) + \tfrac{1}{2}\Big(\eta^+(x,t) - \eta^-(x,t)\Big)\tanh(y), \\ U^1(x,y,t) &= \tfrac{1}{2}\Big(u^+(x,t) + u^-(x,t)\Big) + \tfrac{1}{2}\Big(u^+(x,t) - u^-(x,t)\Big)\tanh(y). \end{aligned} \tag{22}$$

Of course, there is nothing significant about the use of the hyperbolic tangent in the above formulas. Any smooth function $f = f(z)$ such that

$$f(z) - \operatorname{sgn}(z) \in L_2(\mathbb{R}), \quad f' \in H^{k-1}(\mathbb{R}),$$

for k large enough would suffice for the theory to follow. Note that

$$\lim_{y\to\pm\infty} N^1(x,y,t) = \eta^{\pm}(x,t) \;\text{ and }\; \lim_{y\to\pm\infty} U^1(x,y,t) = u^{\pm}(x,t). \tag{23}$$

Let $N = \eta - N^1$, $U = u - U^1$, and $V = v$. The system (8) is equivalent to the system

$$\left\{\begin{array}{lcl} N_t + U_x + V_y + (NU)_x + (NU^1)_x & & \\ \qquad +(N^1U)_x + (NV)_y + (N^1V)_y - N_{xxt} - N_{yyt} & = & G^1(N^1,U^1), \\ \\ \qquad U_t + N_x + UU_x + (UU^1)_x + VV_x - U_{xxt} - U_{yyt} & = & G^2(N^1,U^1), \\ \\ \qquad V_t + N_y + UU_y + (UU^1)_y + VV_y - V_{xxt} - V_{yyt} & = & G^3(N^1,U^1), \end{array}\right. \tag{24}$$

for (N, U, V), with the initial conditions,

$$N(\cdot,0) = \eta_0(\cdot) - N^1(\cdot,0),\ \ U(\cdot,0) = u_0(\cdot) - U^1(\cdot,0),\ \ V(\cdot,0) = v_0(\cdot). \tag{25}$$

In more detail, the right-hand sides in the system (24) are

$$\left\{\begin{array}{lcl} G^1(N^1,U^1) & = & -N^1_t - U^1_x - (N^1U^1)_x + N^1_{xxt} + N^1_{yyt}, \\ \\ G^2(N^1,U^1) & = & -U^1_t - (N^1)_x - U^1U^1_x + U_{xxt} + U^1_{yyt}, \\ \\ G^3(N^1,U^1) & = & -N^1_y - U^1(U^1)_y, \end{array}\right.$$

and direct calculation of these reveals they may be simplified to

$$\left\{\begin{array}{l} G^1(N^1,U^1) = \Big(\rho(y) - \rho(y)^2\Big)\Big((\eta^+ - \eta^-)(u^+ - u^-)\Big)_x + (\eta^+ - \eta^-)_t\rho''(y), \\ \\ G^2(N^1,U^1) = \frac{1}{2}\Big(\rho(y) - \rho(y)^2\Big)\Big((u^+ - u^-)^2\Big)_x + (u^+ - u^-)_t\rho''(y), \\ \\ G^3(N^1,U^1) = \rho'(y)\Big[(\eta^+ - \eta^-) + (u^+ - u^-)\Big(\rho(y)u^+ + (1-\rho(y))u^-\Big)\Big], \end{array}\right.$$

using the equations (21) satisfied by (η^+, u^+) and (η^-, u^-), where

$$\rho(y) = \frac{1 + \tanh(y)}{2}\,.$$

Define the vector-valued functions $W = (N, U, V)$ and $G = (G^1, G^2, G^3)$. In terms of W, the system (24)–(25) has the compact form

$$(I - \Delta)W_t = F(W, N^1, U^1) + G(N^1, U^1). \tag{26}$$

Here $F = F(W, N^1, U^1) = (F^1, F^2, F^3)$ and

$$\left\{\begin{array}{l} F^1(W,N^1,U^1) = \partial_x\Big(U + NU + NU^1 + N^1U\Big) + \partial_y\Big(V + NV + N^1V\Big), \\ \\ F^2(W,N^1,U^1) = \partial_x\Big(N + \frac{1}{2}U^2 + UU^1 + \frac{1}{2}V^2\Big), \\ \\ F^3(W,N^1,U^1) = \partial_y\Big(N + \frac{1}{2}U^2 + UU^1 + \frac{1}{2}V^2\Big). \end{array}\right.$$

Notice that the initial data $N(x,y,0)$, $U(x,y,0)$ and $V(x,y,0)$ all have the property that they converge to 0 as $y \to \pm\infty$. This is also true of $G^j(N^1,U^1)$, $j = 1,2,3$. Indeed, not only do the G^j go to 0 as $y \to \pm\infty$, but they do so uniformly for $(x,t) \in \mathbb{R} \times [0,T]$, where $T > 0$ is a joint existence time for the two pairs (η^+, u^+) and (η^-, u^-).

Just as in Section 3, equation (26) may be recast as an integral equation, *viz.*

$$
\begin{aligned}
W(t) = W(0) + \int_0^1 \Big(K_x \star H(W, N^1, U^1) + K_y \star J(W, N^1, U^1)\Big) ds& \\
+ \int_0^1 K \star G(N^1, U^1)\, ds& \\
= \mathcal{B}(W)(t)&
\end{aligned}
\tag{27}
$$

with

$$
H = \begin{pmatrix} U + NU + NU^1 + N^1 \\ N + \frac{1}{2}U^2 + UU^1 + \frac{1}{2}V^2 \\ 0 \end{pmatrix} \text{ and } J = \begin{pmatrix} V + NV + N^1V \\ 0 \\ N + \frac{1}{2}U^2 + UU^1 + \frac{1}{2}V^2 \end{pmatrix}.
$$

It is straightforward, using the contraction-mapping principle, to ascertain that this integral equation has a solution on some time interval $[0, T]$, $T > 0$, and that this solution provides a solution of the initial-value problem (24). What is particularly notable is that the solution necessarily has zero lateral boundary values, *viz.*

$$
\lim_{y\to\pm\infty} \begin{pmatrix} N(x,y,t) \\ U(x,y,t) \\ V(x,y,t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.
$$

This follows because

$$
\mathcal{M} = \{f \in \mathcal{C}_b(\mathbb{R}^2),\ f(x,y) \to 0 \,\text{as}\, y \to \pm\infty\}
$$

is a closed subspace of $\mathcal{C}_b(\mathbb{R}^2)$. Indeed, notice that $\mathcal{B}$ maps $\mathcal{C}([0,T];\mathcal{M})^3$ into itself since $W(0)$ and G lie there, and U^1 and N^1 are bounded. The contraction-mapping principle applied to a suitable ball B_R around the zero-function in $\mathcal{C}([0,T];\mathcal{C}_b(\mathbb{R}^2))^3$ assures that for any starting value W_1 in B_R, the sequence of iterates $W_{j+1} = \mathcal{B}(W_j)$ converges to the unique fixed point $W_\infty = \mathcal{B}(W_\infty)$. If W_1 is chosen in $\mathcal{C}([0,T];\mathcal{M})^3$, then W_j is in $\mathcal{C}([0,T];\mathcal{M})^3$ for all j and as $\mathcal{M}^3$ is closed in $\mathcal{C}_b(\mathbb{R}^2)^3$, it is inferred that W_∞ lies in $\mathcal{C}([0,T];\mathcal{M})^3$.

If we define $(\eta, u, v) = W + (N^1, U^1, 0)$, then this triple solves (8). By uniqueness, it *must* be *the* solution of (8) with initial data (η_0, u_0, v_0), which demonstrates that necessarily,

$$
\begin{gathered}
\lim_{y\to\pm\infty} \eta(x,y,t) = \eta^\pm(x,t), \quad \lim_{y\to\pm\infty} u(x,y,t) = u^\pm(x,t), \\
\lim_{y\to\pm\infty} v(x,y,t) = 0,
\end{gathered}
\tag{28}
$$

for $0 \le t \le T$. Moreover, if the boundary conditions in (20) imposed upon the initial data are assumed uniformly in x, then the limits in (28) are uniform for $(x,t) \in \mathbb{R} \times [0,T]$. These conclusions are summarized in the following theorem.

Theorem 4.1. *Let (η_0, u_0, v_0) be given in $\mathcal{C}_b(\mathbb{R}^2)^3$ and suppose that the boundary behavior (20) holds. Then the solution (η, u, v) of (8) emanating from (η_0, u_0, v_0) satisfies (28) where the lateral boundary values (η^+, u^+) and (η^-, u^-) solve the reduced system (21). Moreover, if the boundary values in (20) are taken on uniformly for $x \in \mathbb{R}$, then the limits in (28) are uniform for (x,t) in $\mathbb{R} \times [0,T]$.*

Remark 3. As no distinction has been made between the independent variables x and y, the same results hold if it is presumed, instead of (20), that $(\eta_0, u_0, v_0) \in$

$\mathcal{C}_b(\mathbb{R}^2)^3$ and

$$\lim_{x\to\pm\infty}\eta_0(x,y)=\eta_1^{\pm}(y),\quad \lim_{x\to\pm\infty}u_0(x,y)=0,\\ \lim_{x\to\pm\infty}v_0(x,y,t)=v^{\pm}(y).\tag{29}$$

In this case, the solutions $\eta_1^{\pm}(y,t)$, $v^{\pm}(y,t)$ of the reduced system

$$\begin{cases}\eta_t+v_y+(\eta v)_y-\eta_{yyt}=0,\\ \quad v_t+\eta_y+vv_y-v_{yyt}=0,\\ \eta(y,0)=\eta_1^{\pm}(y),\quad v(y,0)=v^{\pm}(y),\end{cases}$$

determine the boundary behavior of the solution (η,u,v) of (8) for $x\to\pm\infty$, with the restrictions (29) on the initial data. More precisely, the limits

$$\lim_{x\to\pm\infty}\eta(x,y,t)=\eta_1^{\pm}(y,t),\quad \lim_{x\to\pm\infty}u(x,y,t)=0,$$

$$\lim_{x\to\pm\infty}v(x,y,t)=v^{\pm}(y,t),$$

hold for all t in the relevant temporal existence interval.

Notice in particular the special case where $\eta_1^+(y)\equiv\overline{\eta}$, $\eta_1^-(y)\equiv\underline{\eta}$, $\lim_{x\to\pm\infty}u_0(x,y)\equiv 0$ and $v^{\pm}(y)\equiv 0$ (with the notation introduced in Remark 2). In this case,

$$\eta_1^+(y,t)\equiv\overline{\eta},\quad \eta_1^-(y,t)\equiv\underline{\eta},\quad v^{\pm}(y,t)\equiv 0.\tag{30}$$

If $\overline{\eta}=\underline{\eta}=0$, this corresponds to the situation wherein the initial disturbance is localized in the $x-$direction, for example, the initial configurations described in (2) and (3). If, on the other hand, $\overline{\eta}=0$ and $\underline{\eta}>0$, the data mimics the situation that one obtains in bore propagation where a surge intrudes, from x near $-\infty$, into a quiescent stretch of the fluid (see, *e.g.* [23], [12], [24] for earlier theory in the case where there is no y–variation and [25] for the more general situation). Our theory assures that if the motion begins in such a configuration, it necessarily maintains this asymptotic behavior, which is to say,

$$\lim_{x\to+\infty}\eta(x,y,t)\equiv 0,\quad \lim_{x\to-\infty}\eta(x,y,t)\equiv\underline{\eta}$$

and

$$\lim_{x\to\pm\infty}u(x,y,t)=\lim_{x\to\pm\infty}v(x,y,t)\equiv 0.$$

Of particular interest is the situation where the initial data has bore-like structure as in (30), but also possesses transverse structure as in (20). Such a configuration could be relevant to the modeling of a tsunami approaching a shoreline (but, of course, not to the last stages of run-up and inundation). For such initial data, the compatibility conditions

$$\lim_{x\to-\infty}\lim_{y\to\pm\infty}\eta_0(x,y)=\underline{\eta},\quad \lim_{x\to+\infty}\lim_{y\to\pm\infty}\eta_0(x,y)=\overline{\eta},\tag{31}$$

and

$$\lim_{|(x,y)|\to+\infty}u_0(x,y)=\lim_{|(x,y)|\to+\infty}v_0(x,y)=0,\tag{32}$$

must be imposed.

With these compatibility conditions, it is straightforward to ascertain that the auxiliary functions N^1 and U^1 introduced in (22) also have the property that

$$\lim_{x\to+\infty}N^1(x,y,t)\equiv\overline{\eta},\quad \lim_{x\to-\infty}N^1(x,y,t)\equiv\underline{\eta},$$

and similarly

$$\lim_{x\to\pm\infty} U^1(x,y,t) \equiv 0.$$

In consequence, the variables (N, U, V) are such that

$$\lim_{|(x,y)|\to+\infty} N(x,y,0) = \lim_{|(x,y)|\to+\infty} U(x,y,0) = \lim_{|(x,y)|\to+\infty} V(x,y,0) \equiv 0.$$

Since $\mathcal{C}_0(\mathbb{R}^2)$ is closed in $\mathcal{C}_b(\mathbb{R}^2)$, our previous argument assures that the solution (N, U, V) of (24) lies in $\mathcal{C}([0,T];\mathcal{C}_0(\mathbb{R}^2))^3$ for some $T > 0$. This in turn has the following consequence.

Theorem 4.2. *Let (η_0, u_0, v_0) be given in $\mathcal{C}_b(\mathbb{R}^2)^3$ and suppose this triple respects the boundary behavior in (20) and (30) and the compatibility conditions (31) and (32) hold. Then, the solution (η, u, v) of (8) that starts at (η_0, u_0, v_0) has the boundary behavior (28) where (η^+, u^+) and (η^-, u^-) solve the reduced system (21).*

4.2. **Neumann conditions.** The argument given in Subsection 4.1 may be adapted to analyze the imposition of Neumann boundary conditions. Suppose the initial data (η_0, u_0, v_0) lies in $\mathcal{C}^1_b(\mathbb{R}^2)^3$ and is such that

$$\lim_{y\to\pm\infty} \partial_y\eta_0(x,y) = \lim_{y\to\pm\infty} u_0(x,y) = 0, \quad \text{and, say,} \quad \lim_{y\to+\pm\infty} v_0(x,y) = 0, \tag{33}$$

for all $x \in \mathbb{R}$. The operator $\mathcal{A}$ associated to this initial data as in (16) maps the subspace

$$\mathcal{N} = \Big\{ W = (\eta,u,v) \in \mathcal{C}([0,T];\mathcal{C}^1_b(\mathbb{R}^2))^3 : \text{for all } x\in\mathbb{R},\ t\in[0,T], \\ \lim_{y\to\pm\infty} \partial_y\eta(x,y,t) = \lim_{y\to\pm\infty} u(x,y,t) = \lim_{y\to+\pm\infty} v(x,y,t) = 0 \Big\} \tag{34}$$

into itself. As $\mathcal{N}$ is closed in $\mathcal{C}^1_b(\mathbb{R}^2)^3$, it follows by a familiar argument that the fixed point $W = (\eta, u, v)$ of $\mathcal{A}$ lies in $\mathcal{C}([0,T];\mathcal{N})$, and this in turn means that the solution of (16) with initial data (η_0, u_0, v_0) maintains zero Neumann conditions for η and u and zero Dirichlet condition on v as $y \to \pm\infty$ throughout its interval of existence. These observations are recorded in the following theorem.

Theorem 4.3. *Let $(\eta_0, u_0, v_0) \in \mathcal{C}^1_b(\mathbb{R}^2)^3$ be given and suppose that the boundary conditions (33) hold. Let (η, u, v) be the solution of (16) guaranteed to exist on some time interval $[0, T]$. Then (η, u, v) lies in the subspace $\mathcal{N}$ defined in (34).*

Remark 4. If the initial data satisfies both (20) and (33), it also follows that the solution belongs to $\mathcal{N}$ and satisfies the dynamic boundary conditions (10). This follows because the auxiliary functions N^1 and U^1 have the properties (23) and, additionally,

$$\lim_{y\to\pm\infty} \partial_y N^1(x,y,t) = \lim_{y\to\pm\infty} U^1(x,y,t) = 0.$$

Thus, for any $T > 0$, the operator $\mathcal{B}$ appearing in (27) maps the subspace

$$\mathcal{P} = \Big\{ f = (f_1,f_2,f_3) \in \mathcal{C}([0,T];\mathcal{C}^1_b(\mathbb{R}^2)\cap\mathcal{C}_0(\mathbb{R}^2))^3 : \text{for all } x\in\mathbb{R},\ t\in[0,T], \\ \partial_y f_i(x,y,t) \to 0 \text{ as } y\to\pm\infty, i=1,2 \Big\}$$

into itself. As $\mathcal{P}$ is closed in $\mathcal{C}^1_b(\mathbb{R}^2)^3$, the solution (N, U, V) of (24) also lies in $\mathcal{P}$. This in turn implies that (η, u, v) satisfies (28) and lies in $\mathcal{N}$.

5. **Long-time existence.** The well-posedness theory put forward above is local in time. An analysis of the contraction-mapping argument in Section 4 and the very closely related argument for local existence theory for the reduced system (21) to be found in [6] reveals that the temporal interval $[0, T]$ of existence is of the form

$$T_0 = \frac{C}{1 + |W(0)|_\infty},$$

where C is a constant independent of the initial data, but which may depend on the function class from which the data is drawn. Thus the existence time afforded by taking recourse to the contraction-mapping principle is of order one, and does not become large even if the data is small.

This is not a satisfactory result if one has in mind the modeling of real water waves. When the initial disturbance respects the assumptions made in the derivation of the system (7), then one hopes for an existence theory at least on what we call the Boussinesq time scale, which is explained now.

Referring back to the system (7) written in variables scaled so that the dependent variables are all of order one, the error terms are of quadratic order, which is to say, of order $\alpha\beta$ and β^2, as mentioned previously. Without fortuitous cancellation, one then expects that the error between the solutions of the model system (7) and the full Euler equations (4) will accumulate like $t(\alpha^2 + \beta^2)$ as the solutions evolve. When t is of order $1/\alpha$, the error will then have grown to $\alpha + \beta S^{-1}$, where S is, as before, the Stokes number; this is an error that is still small compared to the order-one size of the dependent variables. While this description is purely formal, it has been given a rigorous basis in [9]. We term the time scale

$$T_1 \cong \frac{1}{\alpha} \cong \frac{1}{\beta} \tag{35}$$

the *Boussinesq time scale* and put forward the point of view that well-posedness theory should extend at least to this time interval to give the model a chance of having predictive power.

We have not been able to establish well-posedness on the time scale T_1 for data that is merely bounded and continuous. Further regularity is needed in our development. However, this further regularity is entirely consistent with real, non-breaking water waves of small amplitude and long wavelength.

In outline, the present section proceeds as follows. The result in view is first derived for the reduced system (21). This is in fact the key calculation as it informs all the subsequent considerations. We then show how the inequalities implying long-time existence for (21) may be generalized to a system of d coupled equations in one space dimension. This in turn points the way to systems of d equations in more than one spatial dimension, which in turn specializes to give the desired result for the original, water-wave system (7).

5.1. **The one-dimensional case.** Suppose the initial data $(\eta_0(x), u_0(x))$ not only lies in $\mathcal{C}_0(\mathbb{R})^2$, but in fact is drawn from $H^k(\mathbb{R})^2$ where $k \geq 2$. The local theory for the reduced system (21) assures existence, uniqueness and continuous dependence on the data of a solution pair $(\eta, u) \in \mathcal{C}([0, T]; H^k(\mathbb{R}))^2$ for some $T > 0$. If it was known *a priori* that for any time interval $[0, \tilde{T}]$ over which the solution exists, the quantity

$$\sup_{0 \leq t \leq \tilde{T}} ||(\eta, u)||_k \tag{36}$$

is bounded, and this holds independently of $\tilde{T} \le T_1$, where T_1 is the Boussinesq time scale as in (35), then a simple iteration of the contraction-mapping argument would yield the desired result.

To obtain an *a priori* bound on the quantity in (36) when $\tilde{T} = T_1$, energy-type arguments are used. In the calculations to follow, the analysis is made assuming the solution (η, u) is smooth and its components η and u, along with their first few partial derivatives, all vanish as $x \to \pm\infty$. This means that when integration by parts is performed, the boundary terms make no contribution.

One justifies the use of the extra smoothness by approximating (η, u) in the space $\mathcal{C}(0,T;H^k)^2$, for a suitable $k \ge 1$, by a sequence $\{(\eta_j, u_j)\}_{j=1}^{\infty}$ in $\mathcal{C}(0,T;H^{k+5})^2$, say. The pair (η_j, u_j) satisfies the system (21) up to error terms R_j and S_j that tend to zero in $\mathcal{C}(0,T;H^{k-1})$ as $j \to +\infty$. It is straightforward to check that the energy-type calculations to follow are still valid up to error terms that tend to zero as $j \to +\infty$. The solutions of the differential inequalities that result provide *a priori* information that is independent of large values of y, and which is exactly what is obtained *via* formal calculations. This standard argument is sketched briefly in the present subsection, but is not dwelt upon subsequently (*cf.* [11]).

Rewrite the reduced system in the variables corresponding to the full system (7), *viz.*

$$\begin{cases} \eta_t + u_x + \varepsilon(\eta u)_x - \varepsilon\eta_{xxt} = 0, \\ u_t + \eta_x + \varepsilon u u_x - \varepsilon u_{xxt} = 0, \\ \eta(x,0) = \eta_0(x), \quad u(x,0) = u_0(x). \end{cases} \tag{37}$$

Here the Stokes number $S = \frac{\alpha}{\beta}$ has again been set to 1 for simplicity and ε denotes the common value of α and β. In these variables, the Boussinesq time scale T_1 is a quantity of order $\frac{1}{\varepsilon}$.

Multiply the first equation in (37) by η and the second by u and integrate both results over $\mathbb{R}$. After summing the results and performing suitable integrations by parts, it is found that

$$\frac{1}{2}\frac{d}{dt}\int_{-\infty}^{\infty}\Big(\eta^2 + u^2 + \varepsilon u_x^2 + \varepsilon\eta_x^2\Big)dx = -\frac{1}{2}\varepsilon\int_{-\infty}^{\infty}\eta^2 u_x\,dx.$$

Similarly, multiplying the first equation by η_{xx} and the second by u_{xx} and integrating over $\mathbb{R}$ leads to

$$\frac{1}{2}\frac{d}{dt}\int_{-\infty}^{\infty}\Big(\eta_x^2 + u_x^2 + \varepsilon u_{xx}^2 + \varepsilon\eta_{xx}^2\Big)dx = \varepsilon\int_{-\infty}^{\infty}\Big(\eta u_x\eta_{xx} + \eta_x u\eta_{xx} + u u_x u_{xx}\Big)dx.$$

It will be convenient to have at our disposal the function $X : [0,T] \longrightarrow \mathbb{R}_+$ defined to be

$$X(t) = ||u(\cdot,t)||_2 + ||\eta(\cdot,t)||_2.$$

In terms of X, the last two formulas imply the inequalities

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big(u^2 + \eta^2 + \varepsilon u_x^2 + \varepsilon\eta_x^2\Big)dx \le C\varepsilon X^3(t), \tag{38}$$

and

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big(u_x^2 + \eta_x^2 + \varepsilon u_{xx}^2 + \varepsilon\eta_{xx}^2\Big)dx \le C\varepsilon X^3(t), \tag{39}$$

where the two constants denoted C are independent of ε and of $X(t)$.

The crucial step in the analysis presented here is an H^2-estimate, which is a bit more subtle than (38) and (39). Start with the following calculation;

$$\begin{aligned}
&\frac{d}{dt}\int_{-\infty}^{\infty}\Big((1+\gamma\varepsilon\eta)u_{xx}^2+\eta_{xx}^2+\varepsilon u_{xxx}^2+\varepsilon\eta_{xxx}^2\Big)dx\\
&\qquad=\varepsilon\gamma\int_{-\infty}^{\infty}\eta_t u_{xx}^2\,dx+2\varepsilon\gamma\int_{-\infty}^{\infty}\eta u_{xx}u_{xxt}\,dx\\
&\qquad\quad+2\int_{-\infty}^{\infty}\Big(u_{xx}u_{xxt}+\eta_{xx}\eta_{xxt}-\varepsilon u_{xx}u_{xxxxt}-\varepsilon\eta_{xx}\eta_{xxxxt}\Big)dx\\
&=I_1+I_2+I_3,
\end{aligned}\tag{40}$$

where the constant γ will be prescribed presently and several integrations by parts have been made.

Consider first the integral I_1. From the first equation in (37), it is discerned that

$$\eta_t=-(1-\varepsilon\partial_x^2)^{-1}\partial_x(u+\varepsilon\eta u),$$

whence

$$|\eta_t|_\infty\le||\eta_t||_1\le||u_x||_1+\varepsilon||(u\eta)_x||_1\le||u||_2+\varepsilon||u||_2\,||\eta||_2\le X(t)+\varepsilon X^2(t).\tag{41}$$

It follows immediately that

$$I_1\le C\varepsilon\gamma\Big(X^3(t)+\varepsilon X^4(t)\Big).$$

Turning to I_3, notice from (37) again that

$$u_{xxt}-\varepsilon u_{xxxxt}=\partial_x^2(u_t-\varepsilon u_{xxt})=-\partial_x^3\Big(\eta+\frac{1}{2}\varepsilon u^2\Big)$$

and, similarly,

$$\eta_{xxt}-\varepsilon\eta_{xxxxt}=-\partial_x^3(u+\varepsilon u\eta).$$

It thus transpires that

$$\begin{aligned}
\frac{1}{2}I_3 &= -\int_{-\infty}^{\infty}u_{xx}\Big(\eta+\frac{1}{2}\varepsilon u^2\Big)_{xxx}dx-\int_{-\infty}^{\infty}\eta_{xx}\Big(u+\varepsilon u\eta\Big)_{xxx}dx\\
&= -\frac{5}{2}\varepsilon\int_{-\infty}^{\infty}u_x\Big(u_{xx}^2+\eta_{xx}^2\Big)dx-3\varepsilon\int_{-\infty}^{\infty}\eta_x u_{xx}\eta_{xx}dx\\
&\quad+\varepsilon\int_{-\infty}^{\infty}\eta u_{xx}\eta_{xxx}dx\\
&\le 8\varepsilon X^3(t)+\varepsilon\int_{-\infty}^{\infty}\eta u_{xx}\eta_{xxx}dx.
\end{aligned}\tag{42}$$

Naturally, the last integral on the right-hand side of (42) presents a problem.

Attention is now given to I_2. Remark first that

$$\begin{aligned}
\frac{1}{2\gamma\varepsilon}I_2 &= \int_{-\infty}^{\infty}\eta u_{xx}u_{xxt}dx=-\int_{-\infty}^{\infty}\eta u_{xx}\partial_x^2\Big(\eta_x+\varepsilon uu_x-\varepsilon u_{xxt}\Big)dx\\
&= -\int_{-\infty}^{\infty}\eta u_{xx}\eta_{xxx}dx-\varepsilon\int_{-\infty}^{\infty}\eta u_{xx}(uu_x)_{xx}dx+\varepsilon\int_{-\infty}^{\infty}\eta u_{xx}u_{xxxxt}dx\\
&= -\int_{-\infty}^{\infty}\eta u_{xx}\eta_{xxx}dx+J_1+J_2.
\end{aligned}$$

Using Leibniz's rule, the integral J_1 may be expressed as

$$\begin{aligned} J_1 &= -\varepsilon \int_{-\infty}^{\infty} \eta u_{xx}\Big(u u_{xxx} + 3u_x u_{xx}\Big)dx \\ &= \varepsilon \int_{-\infty}^{\infty} \Big(\frac{1}{2}(\eta u)_x - 3\eta u_x\Big) u_{xx}^2 dx \\ &\le 4\varepsilon X^4(t). \end{aligned}$$

Regarding J_2, notice that

$$\begin{aligned} J_2 &= -\varepsilon \int_{-\infty}^{\infty} \Big(\eta u_{xxx} u_{xxxt} + \eta_x u_{xx} u_{xxxt}\Big)dx \\ &= -\frac{\varepsilon}{2}\frac{d}{dt}\int_{-\infty}^{\infty} \eta u_{xxx}^2 dx + \frac{\varepsilon}{2}\int_{-\infty}^{\infty} \eta_t u_{xxx}^2 dx - \varepsilon \int_{-\infty}^{\infty} \eta_x u_{xx} u_{xxxt} dx. \end{aligned} \tag{43}$$

Because of (41), the second term on the right-hand side of (43) may be bounded above thusly;

$$\frac{\varepsilon}{2}\int_{-\infty}^{\infty} \eta_t u_{xxx}^2 dx \le C\varepsilon\Big(X(t) + \varepsilon X^2(t)\Big)\int_{-\infty}^{\infty} u_{xxx}^2 dx.$$

Again using (37), there follows the relation

$$\varepsilon u_{xxxt} = -\varepsilon \partial_x^2 (1 - \varepsilon\partial_x^2)^{-1} \partial_x^2 \Big(\eta + \frac{\varepsilon}{2}u^2\Big),$$

whence

$$|\varepsilon u_{xxxt}|_2 \le C\Big(X(t) + \varepsilon X^2(t)\Big).$$

It follows that

$$\varepsilon \int_{-\infty}^{\infty} \eta_x u_{xx} u_{xxxt} dx \le \varepsilon |\eta_x|_\infty |u_{xx}|_2 \, |u_{xxxt}|_2 \le C\, X^2(t)\Big(X(t) + \varepsilon X^2(t)\Big).$$

Thus, we see that

$$\begin{aligned} J_2 \le -\frac{\varepsilon}{2}\frac{d}{dt}\int_{-\infty}^{\infty} \eta u_{xxx}^2 dx + C\varepsilon\Big(X(t) + \varepsilon X^2(t)\Big)\int_{-\infty}^{\infty} u_{xxx}^2 dx \\ + C X^2(t)\Big(X(t) + \varepsilon X^2(t)\Big). \end{aligned}$$

The estimates of the integrals I_1, I_2 and I_3 may be combined to yield the helpful inequality

$$\begin{aligned} I_1 + I_2 + I_3 \le\ & C\varepsilon\gamma X^3(t) + C\varepsilon^2\gamma X^4(t) \\ & + 2\varepsilon(1-\gamma)\int_{-\infty}^{\infty} \eta u_{xx}\eta_{xxx} dx - \varepsilon^2\gamma\frac{d}{dt}\int_{-\infty}^{\infty} \eta u_{xxx}^2 dx \\ & + C\varepsilon^2\gamma\Big(X(t) + \varepsilon X^2(t)\Big)|u_{xxx}|_2^2. \end{aligned} \tag{44}$$

First, choose $\gamma = 1$ to rid ourselves of the troublesome term and then combine (44) and (40) to deduce that

$$\begin{aligned} &\frac{d}{dt}\int_{-\infty}^{\infty} \Big[(1+\varepsilon\eta)u_{xx}^2 + \eta_{xx}^2 + \varepsilon\Big((1+\varepsilon\eta)u_{xxx}^2 + \eta_{xxx}^2\Big)\Big]dx \\ &\quad \le C\varepsilon X^3(t) + C\varepsilon^2 X^4(t) + C\varepsilon^2\Big(X(t) + \varepsilon X^2(t)\Big)\int_{-\infty}^{\infty} u_{xxx}^2 dx. \end{aligned} \tag{45}$$

We are now in a position to adduce the following interesting result. Define $Y(t)$ by

$$Y(\eta, u) = Y(t) = X(t) + \varepsilon\Big(|\eta_{xxx}(\cdot,t)|_2 + |u_{xxx}(\cdot,t)|_2\Big).$$

Theorem 5.1. *Let $R > 0$ be specified. Then there is an $\varepsilon_0 = \varepsilon_0(R) > 0$ and a constant $C = C(R)$ such that if initial data $(\eta_0, u_0) \in H^3(\mathbb{R}) \times H^3(\mathbb{R})$ has*

$$Y(\eta_0, u_0) \leq R$$

and if $\varepsilon \leq \varepsilon_0$, then the solution (η, u) to system (37) emanating from (η_0, u_0) exists for at least the time interval $\left[0, \frac{C}{\varepsilon}\right]$.

Proof. This result of long-time existence for fixed data (data in a fixed ball in $H^3(\mathbb{R})$) is an elementary consequence of the differential inequalities (38), (39) and (45) together with a Gronwall-type argument.

It remains only to justify the calculations made assuming η and u are smooth for initial data that lies only in H^3. The argument is straightforward and can be found, for example, in a similar context in [11].

In a little more detail, let (η, u) be a solution in $\mathcal{C}(0, T; H^3)^2$ for some $T > 0$. Fix $\varepsilon > 0$. It follows immediately from Theorem 3.1, Part (i), that $(\eta_t, u_t) \in \mathcal{C}(0, T; H^4)^2$. Approximate (η, u) by a sequence $\{(\eta_j, u_j)\}_{j=1}^{\infty}$ in $\mathcal{C}^1(0, T; H^{10})^2$, say, so that

$$||\eta_j - \eta||_{\mathcal{C}(0,T,H^3)} + ||\partial_t \eta_j - \partial_t \eta||_{\mathcal{C}(0,T,H^4)} \to 0$$

and

$$||u_j - u||_{\mathcal{C}(0,T,H^3)} + ||\partial_t u_j - \partial_t u||_{\mathcal{C}(0,T,H^4)} \to 0$$

as $j \to \infty$.

Of course the pairs (η_j, u_j) are not solutions of the system (37), but as they are eventually close to (η, u), it is straightforward to ascertain that

$$\begin{cases} \partial_t \eta_j + \partial_x u_j + \varepsilon \partial_x(\eta_j u_j) - \varepsilon \partial_x^2 \partial_t \eta_j & = R_j, \\ \partial_t u_j + \partial_x \eta_j + \varepsilon u_j \partial_x u_j - \varepsilon \partial_x^2 \partial_t u_j & = S_j, \end{cases}$$

where

$$||R_j||_{\mathcal{C}(0,T;H^2)},\ ||S_j||_{\mathcal{C}(0,T;H^2)} \to 0$$

as $j \to \infty$.

Letting $v = \eta_j$ and $w = u_j$, it is clear that (38) and (39) hold in the revised form

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big(v^2 + w^2 + \varepsilon v_x^2 + \varepsilon w_x^2\Big)dx \leq C\varepsilon X^3(t) + \int_{-\infty}^{\infty}\Big(|v|\,|R_j| + |w|\,|S_j|\Big)dx, \quad (46)$$

and

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{\infty}&\Big(v_x^2 + w_x^2 + \varepsilon v_{xx}^2 + \varepsilon w_{xx}^2\Big)dx \\ &\leq C\varepsilon X^3(t) + \int_{-\infty}^{\infty}\Big(|v|\,|\partial_x^2 R_j| + |w|\,|\partial_x^2 S_j|\Big)dx.\end{aligned} \quad (47)$$

A tedious, but straightforward, calculation reveals that (45) holds up to error terms involving R_j and S_j. Precisely, the differential inequality

$$\frac{d}{dt}\int_{-\infty}^{+\infty}\Big((1+\varepsilon v)w_{xx}^2 + v_{xx}^2 + \varepsilon((1+\varepsilon v)w_{xxx}^2 + v_{xxx}^2)\Big)dx \leq C\varepsilon X_j^3(t) + C\varepsilon^2 X_j^4(t) + \alpha_j$$

emerges, where ε is fixed, but arbitrary in the range $(0, 1]$. The term α_j may be bounded above as follows:

$$|\alpha_j| \leq CX^2(t)\Big[||R_j||_2 + ||S_j||_2\Big]. \quad (48)$$

Because v and w have extra regularity, all the computations leading to (46)–(48) are easily justified.

Since the term in square brackets in (48) tends to zero as $j \to +\infty$, uniformly on $[0, T]$, for j sufficiently large, the conclusions of the theorem hold by appeal to a Gronwall argument. The result follows. □

5.2. **Long-time existence for a general system in one dimension.** An analysis of the essential aspects of the calculations in Section 5.1 points towards a more general result. The elucidation of this result is the provenance of the present subsection.

Consider a system of d coupled partial differential equations in one space dimension of the form

$$\partial_t U + L\partial_x U + \varepsilon M(U)\partial_x U - \varepsilon \partial_x^2 \partial_t U = 0, \tag{49}$$

where $U(x,t) = (u_1(x,t), u_2(x,t), \cdots, u_d(x,t))$ and $x \in \mathbb{R}$. The following hypotheses are made concerning L and M.

1. The $d \times d$ matrix $L = \Big[l_{ij}\Big]$ of real numbers is symmetric,
2. $M(U) = \Big[m_{ij}(U)\Big]$ is a $d \times d$ matrix whose components are smooth functions of $U = (u_1, \cdots, u_d)$ and
3. there is a $d \times d$ symmetric matrix $N(U)$ of smooth functions such that
$$\Big(\mathrm{I}_d + \varepsilon N(U)\Big)\Big(L + \varepsilon M(U)\Big)$$
is symmetric, which is to say, the hyperbolic part of (49) (the system without the dispersive terms $\partial_x^2 \partial_t U$) is symmetrizable. (Here, I_d connotes the $d \times d$ identity matrix in $\mathbb{R}^{d^2}$).

Hypotheses 1 and 2 allow us to formulate a local existence theory for the initial-value problem corresponding to initial data

$$U(x,0) = U_0(x). \tag{50}$$

Indeed, the argument in favor of such a proposition proceeds just as in the case $d = 2$, and does not require the symmetry of L in Hypothesis 1, nor the structure apparent in Hypothesis 3. This result is stated without a detailed proof in the following proposition.

Proposition 1. *The initial-value problem (49)–(50) is locally well posed in $\mathcal{C}_b(\mathbb{R})^d$, and indeed in Z^d, where $Z = Z(\mathbb{R})$ is any Sobolev space of functions $\{f : \mathbb{R} \mapsto \mathbb{R}\}$ embedded in $L_\infty(\mathbb{R})$ and satisfying the properties 1 and 2 of Theorem 3.2.*

Long-time existence in the Sobolev class $Z^d = H^2(\mathbb{R})^d$ is established using energy estimates as in Section 5.1. For the following calculations, the full force of the hypotheses 1, 2 and 3 above are used.

We begin with an $L_2(\mathbb{R})$–estimate, *viz.*

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big(|U|^2 + \varepsilon|\partial_x U|^2\Big)dx = -2\int_{-\infty}^{\infty}\Big(U \cdot L\partial_x U + \varepsilon U \cdot M(U)\partial_x U\Big)dx.$$

The first term on the right-hand side vanishes on account of the symmetry of L. The second term is bounded in a simple way, leading to

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{\infty}\Big(|U|^2 + \varepsilon|\partial_x U|^2\Big)dx &\leq \varepsilon|M(U)|_\infty|\partial_x U|_2|U|_2 \\ &\leq C\varepsilon||U||_1^2 F_0(||U||_1),\end{aligned} \tag{51}$$

where F_0 depends on the growth of the matrix elements m_{ij} of M.

For an $H^1(\mathbb{R})-$estimate, a similar computation yields

$$\begin{aligned}\frac{d}{dt}\int_{-\infty}^{\infty}\Big(|\partial_x U|^2+\varepsilon|\partial_x^2 U|^2\Big)dx &= 2\int_{-\infty}^{\infty}\Big(\partial_x U\cdot\partial_t\partial_x U-\varepsilon\partial_x U\cdot\partial_x^3 U\partial_t U\Big)dx\\ &= -2\varepsilon\int_{-\infty}^{\infty}\partial_x U\cdot\Big[L\partial_x^2 U+\varepsilon\partial_x\Big(M(U)\partial_x U\Big)\Big]dx.\end{aligned}$$

The first term on the right-hand side of the preceding equation vanishes because of the symmetry of the constant matrix L, and a bound for the second term gives immediately the inequality

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big(|\partial_x U|^2+\varepsilon|\partial_x^2 U|^2\Big)dx\le C\varepsilon\Big(F_0(X)X^2+F_1(X)X^3\Big),$$

where $F_1(X)$ depends on the gradients ∇m_{ij}, $1\le i,j\le d$, of the matrix entries of M and on $X(t)=||U(\cdot,t)||_2$.

So far no useful bound is revealed. The crucial step is, as before, the estimate in $H^2(\mathbb{R})$. Use the symmetry of the matrix $N(U)$ and calculate as follows;

$$\begin{aligned}&\frac{d}{dt}\int_{-\infty}^{\infty}\Big[\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x^2 U\cdot\partial_x^2 U+\varepsilon|\partial_x^3 U|^2\Big]dx\\ &\qquad=2\int_{-\infty}^{\infty}\Big[\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x^2\partial_t U\cdot\partial_x^2 U+\varepsilon\partial_x^3 U\cdot\partial_x^3\partial_t U\Big]dx\\ &\qquad\quad+\varepsilon\int_{-\infty}^{\infty}\partial_t N(U)\partial_x^2 U\cdot\partial_x^2 U\,dx.\end{aligned}\tag{52}$$

Using equation (49), and after an integration by parts, the first integral on the right-hand side of (52) is seen to equal

$$\begin{aligned}&-2\int_{-\infty}^{\infty}\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\Big\{\partial_x^2\Big[\Big(L+\varepsilon M(U)\Big)\partial_x U\Big]\Big\}\cdot\partial_x^2 U\,dx\\ &\qquad\qquad+2\varepsilon^2\int_{-\infty}^{\infty}N(U)\partial_x^4\partial_t U\cdot\partial_x^2 U\,dx=I_1+I_2.\end{aligned}$$

Thus, it follows that

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big[\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x^2 U\cdot\partial_x^2 U+\varepsilon|\partial_x^3 U|^2\Big]dx=I_1+I_2+I_3,$$

where I_3 has the obvious definition. Attention is now given to determining effective bounds on the three integrals I_1, I_2 and I_3.

Notice that, from equation (49),

$$\partial_t U=-(1-\varepsilon\partial_x^2)^{-1}\Big[L\partial_x U+\varepsilon M(U)\partial_x U\Big],$$

and, consequently,

$$\begin{aligned}|\partial_t U|_\infty &\le CX\Big(1+F_0(X)\Big)\\ &= F_2(X),\end{aligned}\tag{53}$$

where C is independent of ε, and F_0 is defined in (51). It follows that

$$\begin{aligned}|I_3| &\le C\varepsilon F_3(X)F_2(X)|\partial_x^2 U|_2^2\\ &= C\varepsilon X^2F_4(X),\end{aligned}\tag{54}$$

with $F_3(X)$ depending on the growth of the elements ∇n_{ij}, $1\le i,j\le d$, of the matrix entries of N.

To bound I_1, write it out as follows;

$$\begin{aligned} I_1 &= -2\int_{-\infty}^{\infty}\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\Big(L+\varepsilon M(U)\Big)\partial_x^3 U\cdot\partial_x^2 U dx \\ &\quad -4\varepsilon\int_{-\infty}^{\infty}\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x M(U)\partial_x^2 U\cdot\partial_x^2 U dx \\ &\quad -2\varepsilon\int_{-\infty}^{\infty}\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x^2 M(U)\partial_x U\cdot\partial_x^2 U dx. \end{aligned}$$

The lowest-order term $\int_{-\infty}^{\infty} L\partial_x^3 U\cdot\partial_x^2 U dx$ vanishes, since L is a constant, symmetric matrix. In consequence, it transpires that

$$\begin{aligned} I_1 = &- 2\varepsilon\int_{-\infty}^{\infty}\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\Big[\partial_x^2 M(U)\partial_x U+2\partial_x M(U)\partial_x^2 U\Big]\cdot\partial_x^2 U dx \\ &- 2\varepsilon\int_{-\infty}^{\infty}\Big[M(U)+N(U)\Big(L+\varepsilon M(U)\Big)\Big]\partial_x^3 U\cdot\partial_x^2 U dx, \end{aligned}$$

whose absolute value is clearly bounded above by a quantity of the form

$$\varepsilon X^2 F_4(X),$$

since $\varepsilon M(U)+\varepsilon N(U)\Big(L+\varepsilon M(U)\Big)=(\mathrm{I}_d+\varepsilon N(U))(L+\varepsilon M(U))$ and L is symmetric.

To estimate I_2, similar calculations are effective, *viz.*

$$\begin{aligned} \frac{1}{2\varepsilon^2}I_2 &= \int_{-\infty}^{\infty} N(U)\partial_x^4\partial_t U\cdot\partial_x^2 U dx \\ &= -\int_{-\infty}^{\infty} N(U)\partial_t\partial_x^3 U\cdot\partial_x^3 U dx-\int_{-\infty}^{\infty}\partial_x N(U)\partial_x^3\partial_t U\cdot\partial_x^2 U dx \\ &= -\frac{d}{dt}\int_{-\infty}^{\infty} N(U)\partial_x^3 U\cdot\partial_x^3 U dx+\int_{-\infty}^{\infty}\partial_t N(U)\partial_x^3 U\cdot\partial_x^3 U dx \\ &\quad +\int_{-\infty}^{\infty} N(U)\partial_x^3 U\cdot\partial_x^3\partial_t U dx \\ &= -\frac{d}{dt}\int_{-\infty}^{\infty} N(U)\partial_x^3 U\cdot\partial_x^3 U dx+J_1+J_2. \end{aligned}$$

The argument continues by obtaining bounds on J_1 and J_2. For J_1, use again the bound on $|\partial_t N(U)|_\infty$ appearing in (53) to obtain

$$|J_1|\le CF_3(X)|\partial_x^3 U|^2. \tag{55}$$

For J_2, use equation (49) to write

$$\varepsilon\partial_x^3\partial_t U=-\varepsilon\partial_x^2(\mathrm{I}_d-\varepsilon\partial_x^2)^{-1}\partial_x\Big((L+\varepsilon M(U))\partial_x U\Big),$$

so that

$$\varepsilon|\partial_x^3\partial_t U|_2\le CF_5(X).$$

It follows that

$$J_2\le CF_6(X). \tag{56}$$

Collecting estimates (52) to (56), there obtains the inequality

$$\frac{d}{dt}\int_{-\infty}^{\infty}\Big[\Big(\mathrm{I}_d+\varepsilon N(U)\Big)\partial_x^2 U\cdot\partial_x^2 U+\varepsilon|\partial_x^3 U|^2+\varepsilon^2 N(U)\partial_x^3 U\cdot\partial_x^3 U\Big]dx\le C\varepsilon F_7(X),$$

where C is a constant independent of ε and F_7 is an increasing function, at least quadratic in X, which depends on the growth of the m_{ij} and n_{ij}, $1 \leq i, j \leq d$, and their gradients. Well-posedness on the Boussinesq time-scale now follows. A formal statement of the result appears in the next subsection.

5.3. **Long-time existence for a general system in n dimensions.** Consider a system of d coupled partial differential equations in n spatial dimensions of the form

$$\begin{cases} \partial_t U + \displaystyle\sum_{i=1}^{n} \Big(L_i + \varepsilon M_i(U) \Big) \partial_{x_i} U - \varepsilon \Delta \partial_t U = 0, \\ U(x,0) = U_0(x), \end{cases} \tag{57}$$

where $U(x,t) = (u_1(x,t), u_2(x,t), \cdots, u_d(x,t))$, $\Delta = \displaystyle\sum_{i=1}^{n} \partial_{x_i}^2$ and $x \in \mathbb{R}^n$.

The following hypotheses are made concerning L_i and M_i, for $i = 1, \cdots, n$.

1. The $d \times d$ matrices L_i of real numbers are symmetric,
2. the $M_i(U)$ are $d \times d$ matrices of smooth functions of $U = (u_1, \cdots, u_d)$ and
3. there is a $d \times d$ symmetric matrix $N(U)$ of smooth functions such that, for all $i = 1, 2, \cdots, n$,
$$\Big(\mathrm{I}_d + \varepsilon N(U) \Big) \Big(L_i + \varepsilon M_i(U) \Big)$$
is symmetric, which is to say, the hyperbolic part of (57) (the system without the dispersive terms $\Delta \partial_t U$) is symmetrizable.

Theorem 5.2. *Let there be given an integer s such that $s > \frac{n}{2} + 1$ where $n \geq 1$ is an integer. For any initial data $U_0 \in H^s(\mathbb{R}^n)^d$, there exists a $T_0 > 0$, independent of ε, such that if the Hypotheses 1, 2 and 3 (in Subsection 5.2 for $n = 1$ and the present subsection for $n > 1$) hold, then Problem (49) for $n = 1$ and Problem (57) for $n > 1$ have a unique solution U^ε in $\mathcal{C}\Big([0, \frac{T_0}{\varepsilon}]; H^s(\mathbb{R}^n)^d\Big)$ corresponding to U_0. The solution mapping $U_0 \mapsto U$ is uniformly Lipschitz on bounded subsets of $H^s(\mathbb{R}^n)^d$ and, moreover, there exists a constant $C_0 > 0$ such that*

$$||U^\varepsilon||_{L_\infty\left([0, \frac{T_0}{\varepsilon}]; H^s(\mathbb{R}^n)^d\right)} \leq C_0.$$

Proof. The result for $n = 1$ was established in the last subsection.

The proof for higher spatial dimension follows the same lines as for the one-dimensional case as is now demonstrated. Exactly as in (51), there obtains the $L_2(\mathbb{R}^n)$–estimate,

$$\frac{d}{dt} \int_{\mathbb{R}^n} \Big(|U|^2 + \varepsilon |\nabla U|^2 \Big) dx \quad \leq \quad C\varepsilon ||U||_1^2 F(|U|_\infty),$$

where F depends on the growth of the matrix elements m_{ij} of M. Because $s > \frac{n}{2}$, the H^s–norm can be used to bound the L_∞–norm and so

$$\frac{d}{dt} \int_{\mathbb{R}^n} \Big(|U|^2 + \varepsilon |\nabla U|^2 \Big) dx \quad \leq \quad \varepsilon F_0(||U||_s),$$

where $F_0(r)$ has the form $Cr^2 F(r)$.

Fix j with $1 \le j \le n$ and let ∂ temporarily denote ∂_{x_j}. Use equation (57) and the symmetry of the matrix $N(U)$ to calculate as follows; for any k with $2 \le k \le s$,

$$\frac{d}{dt}\int_{\mathbb{R}^n}\Big((\mathrm{I}_d+\varepsilon N(U))\partial^k U\cdot\partial^k U+\varepsilon\partial^k\nabla U\,:\,\partial^k\nabla U\Big)dx$$

$$=2\int_{\mathbb{R}^n}\Big((\mathrm{I}_d+\varepsilon N(U))\partial^k\partial_t U\cdot\partial^k U+\varepsilon\partial^k\partial_t\nabla U:\partial^k\nabla U\Big)dx$$
$$+\varepsilon\int_{\mathbb{R}^n}N'(U)(\partial_t U)\partial^k U\cdot\partial^k U\,dx$$

$$=2\int_{\mathbb{R}^n}\Big((\mathrm{I}_d+\varepsilon N(U))\partial^k\partial_t U\cdot\partial^k U-\varepsilon\partial^k\Delta\partial_t U\cdot\partial^k U\Big)dx$$
$$+\int_{\mathbb{R}^n}N'(U)(\partial_t U)\partial^k U\cdot\partial^k U\,dx$$

$$=2\int_{\mathbb{R}^n}(\mathrm{I}_d+\varepsilon N(U))\Big\{-\partial^k\sum_{i=1}^n(L_i+\varepsilon M_i(U))\partial_{x_i}U\Big\}\cdot\partial^k U\,dx$$
$$+2\varepsilon^2\int_{\mathbb{R}^n}N(U)\partial^k\Delta\partial_t U\cdot\partial^k U\,dx+\int_{\mathbb{R}^n}N'(U)(\partial_t U)\partial^k U\cdot\partial^k U\,dx$$

$$=I_1+I_2+I_3.$$

To bound I_3, write equation (57) in the form

$$\partial_t U=-(\mathrm{I}_d-\varepsilon\Delta)^{-1}\sum_{i=1}^n\Big(L_i+\varepsilon M_i(U)\Big)\partial_{x_i}U, \tag{58}$$

from which it follows immediately that

$$||\partial_t U||_{s-1}\le C||U||_s,\qquad |\partial_t U|_\infty\le C||U||_s.$$

In consequence, there obtains

$$I_3\le\varepsilon F_1(||U||_s)$$

for some function F_1 depending only on the growth of the n_{ij} and m_{ij}, and of the gradients of the m_{ij}.

To estimate I_1, write

$$I_1=-2\int_{\mathbb{R}^n}\partial^k\Big(\sum_{i=1}^n L_i\partial_{x_i}U\Big)\cdot\partial^k U\,dx-2\int_{\mathbb{R}^n}\varepsilon\Big(\partial^k\Big(\sum_{i=1}^n M_i(U)\partial_{x_i}U\Big)\cdot\partial^k U$$
$$+N(U)\Big(\partial^k\sum_{i=1}^n(L_i+\varepsilon M_i(U))\partial_{x_i}U\Big)\cdot\partial^k U\Big)dx.$$

Because of the symmetry of the constant matrices L_i, $1\le i\le n$, the first integral on the right-hand side vanishes. In the second integral, the most troublesome terms are those where $k+1$ derivatives are taken of the vector U, *viz.*

$$-2\int_{\mathbb{R}^n}\varepsilon\Big(M_i(U)\partial^k\partial_{x_i}U\cdot\partial^k U+N(U)(L_i+\varepsilon M_i(U))\partial^k\partial_{x_i}U\cdot\partial^k U\Big)dx$$

$$=-2\int_{\mathbb{R}^n}\varepsilon\Big(M_i(U)+N(U)(L_i+\varepsilon M_i(U))\Big)\partial^k\partial_{x_i}U\cdot\partial^k U\,dx,$$

for $i = 1, \cdots, n$. Because of the symmetry of the operators in this integral, it follows immediately that

$$|I_1| \le \varepsilon F_2(||U||_s).$$

For I_2, we proceed thusly;

$$\begin{aligned}
I_2 &= 2\varepsilon^2 \int_{\mathbb{R}^n} N(U)\partial_t \partial^k \Delta U \cdot \partial^k U \, dx \\
&= -2\varepsilon^2 \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N'(U)(\partial_{x_\ell} U)\partial_t \partial^k \partial_{x_\ell} U \cdot \partial^k U \, dx \\
&\quad -2\varepsilon^2 \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N(U)\partial_t \partial^k \partial_{x_\ell} U \cdot \partial^k \partial_{x_\ell} U \, dx \\
&= -2\varepsilon^2 \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N'(U)(\partial_{x_\ell} U)\partial_t \partial^k \partial_{x_\ell} U \cdot \partial^k U \, dx \\
&\quad +2\varepsilon^2 \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N(U)\partial^k \partial_{x_\ell} U \cdot \partial_t \partial^k \partial_{x_\ell} U \, dx \\
&\quad -\varepsilon^2 \frac{d}{dt} \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N(U)\partial^k \partial_{x_\ell} U \cdot \partial^k \partial_{x_\ell} U \, dx \\
&\quad +\varepsilon^2 \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N'(U)(\partial_t U)\partial^k \partial_{x_\ell} U \cdot \partial^k \partial_{x_\ell} U \, dx \\
&= J_1 - \varepsilon^2 \frac{d}{dt} \sum_{\ell=1}^{n} \int_{\mathbb{R}^n} N(U)\partial^k \partial_{x_\ell} U \cdot \partial^k \partial_{x_\ell} U \, dx + J_2.
\end{aligned}$$

Equation (58) implies that

$$\partial^k \partial_{x_\ell} \partial_t U = -\partial^k (\mathrm{I}_d - \varepsilon\Delta)^{-1} \partial_{x_\ell} \sum_{i=1}^{n} \Big(L_i + \varepsilon M_i(U) \Big) \partial_{x_i} U,$$

for $\ell = 1, \cdots, n$. Thus, there obtains the inequality

$$|\varepsilon \partial_t \partial^k \partial_{x_\ell} U|_2 \le CF(||U||_s),$$

for $\ell = 1, \cdots, n$, which shows that

$$J_1 \le \varepsilon C \Big(|N'(U)|_\infty \, |\partial_t U|_\infty + |N(U)|_\infty \Big) F(||U||_s).$$

For J_2, it immediately follows that

$$J_2 \le \varepsilon^2 |N'(U)|_\infty \, |\partial_t U|_\infty \int_{\mathbb{R}^n} \sum_{\ell=1}^{n} \partial^k \partial_{x_\ell} U \cdot \partial^k \partial_{x_\ell} U \, dx.$$

If we define $Y = ||U||_s + \varepsilon ||U||_{s+1}$, then

$$|I_2| \le \varepsilon CF(Y).$$

In consequence, it is seen that

$$\frac{d}{dt}\Big(\sum_{k=[\frac{n}{2}+1]+1}^{s}\int_{\mathbb{R}^n}(\mathrm{I}_d+\varepsilon N(U))\partial^k U\cdot\partial^k U\,dx+\varepsilon\int_{\mathbb{R}^n}|\partial^k\nabla U|^2\,dx$$

$$+\varepsilon^2\int_{\mathbb{R}^n}\sum_{i=1}^{n}N(U)\partial^k\partial_{x_i}U\cdot\partial^k\partial_{x_i}U\,dx\Big)\le C\varepsilon F(Y),$$

where F is a function of Y depending on the growth of the m_{ij} and n_{ij} and their gradients. The conclusions stated in Theorem 5.3 now follow by solving the preceding differential inequality. □

5.4. **Application to water waves.** The preceding theory is now specialized to the water-wave models of interest here. The one-dimensional, reduced system (37) has the form (49) with $U=(\eta,u)$ and

$$L=\begin{bmatrix}0&1\\1&0\end{bmatrix},\qquad M(U)=\begin{bmatrix}u&\eta\\0&u\end{bmatrix}.$$

The matrix

$$N(U)=\begin{bmatrix}0&0\\0&\eta\end{bmatrix}$$

is appropriate for symmetrizing the hyperbolic part of (37), because

$$\Big(\mathrm{I}_2+\varepsilon N(U)\Big)\Big(L+\varepsilon M(U)\Big)=\begin{bmatrix}\varepsilon u&1+\varepsilon\eta\\1+\varepsilon\eta&\varepsilon u(1+\varepsilon\eta)\end{bmatrix}$$

is symmetric. Thus all the hypotheses set forth in Subsection 5.2 are satisfied and long-time well-posedness is established (see [2] for an alternative proof of this fact).

The full two-dimensional system has the detailed form

$$\begin{cases}\eta_t+u_x+v_y+\varepsilon\Big((\eta u)_x+(\eta v)_y\Big)-\varepsilon\Big(\eta_{xxt}+\eta_{yyt}\Big)&=&0,\\ u_t+\eta_x+\varepsilon\Big(uu_x+vv_x\Big)-\varepsilon\Big(u_{xxt}+u_{yyt}\Big)&=&0,\\ v_t+\eta_y+\varepsilon\Big(uu_y+vv_y\Big)-\varepsilon\Big(v_{xxt}+v_{yyt}\Big)&=&0.\end{cases}\tag{59}$$

In terms of the notation in Subsection 5.3, $d=3$ and $n=2$ (three equations in two spatial dimensions). We first use the fact that the flow is irrotational, so that $u_y-v_x=0$ at all times. Thus the term vv_x may be replaced by vu_y in the second equation in (59) and, similarly, uu_y by uv_x in the third equation. With these substitutions, the system (59) has the form (57) with $U=(\eta,u,v)$ and

$$L_1=\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix},\qquad L_2=\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix},$$

and

$$M_1(U)=\begin{bmatrix}u&\eta&0\\0&u&0\\0&0&u\end{bmatrix},\quad M_2(U)=\begin{bmatrix}v&0&\eta\\0&v&0\\0&0&v\end{bmatrix}.$$

The matrix

$$N(U)=\begin{bmatrix}0&0&0\\0&\eta&0\\0&0&\eta\end{bmatrix}$$

is appropriate for symmetrizing the hyperbolic part of (59) since, as a matter of fact, the matrices

$$(\mathrm{I}_3 + \varepsilon N(U))(L_1 + \varepsilon M_1(U)) = \begin{bmatrix} \varepsilon u & 1+\varepsilon\eta & 0 \\ 1+\varepsilon\eta & \varepsilon u(1+\varepsilon\eta) & 0 \\ 0 & 0 & \varepsilon u(1+\varepsilon\eta) \end{bmatrix}$$

and

$$(\mathrm{I}_3 + \varepsilon N(U))(L_2 + \varepsilon M_2(U)) = \begin{bmatrix} \varepsilon v & 0 & 1+\varepsilon\eta \\ 0 & \varepsilon v(1+\varepsilon\eta) & 0 \\ 1+\varepsilon\eta & 0 & \varepsilon v(1+\varepsilon\eta) \end{bmatrix}$$

are symmetric. Thus the hypotheses of Theorem 5.2 are verified for the full, two-dimensional system, and long-time well-posedness is concluded.

6. Conclusion.

Considered here was a Boussinesq system of equations that serves as a model for surface water-wave propagation. The theory developed, while not encompassing overturning waves, is fully three-dimensional. Allowance is made for non-localized disturbances, so that line solitary waves and their perturbations are within the scope of our development.

In addition to appropriate local well-posedness results, we have also established long-time existence for disturbances that are spatially localized in one of the horizontal directions, taken here to be the $x-$direction. Such disturbances might be called long-crested. It is our view that to be interesting from the perspective of the potential applications, theory for this kind of model needs to persist at least on what is here termed the Boussinesq time scale. This is the time scale on which nonlinear and dispersive effects can make an order one relative contribution to the wave motion. On time scales which are only of order one, simply using the linear wave equation suffices for approximating solutions of the full water-wave problem.

In a companion paper, we will develop similar long-time existence theory for certain types of disturbances that are not localized in either spatial direction. This will include in particular solutions corresponding to modeling the propagation of bores (see Section 4). The longer-time theory in this case is a bit more subtle because even the slices of the wave profile corresponding to fixed values of y have infinite energy.

Another interesting investigation would be to consider the type of initial disturbances featured here, but posed for others of the *abcd*–systems (5). Results valid on Boussinesq time scales have recently been developed in case the initial data evanesces to the rest state in all directions (Sobolev-class initial data, see [21] and [26]). When non-homogeneous boundary conditions are present, like those in (20) and their subsequent dynamical consequences (9), theory is likely to be more subtle than that developed here since, as soon as terms having three x– or y–derivatives appear in the equations, extra boundary conditions at infinity will need to be prescribed. It is not immediately clear what additional auxiliary specifications would be appropriate in modeling water waves, especially as regards the lateral conditions appearing in (20) at $y = \pm\infty$.

Aknowledgments. JB would like to record his thanks to the Laboratoire d'Analyse et de Mathématiques Appliquées at the Université Paris-Est Créteil for hospitality during the gestation of this work and to the Équipe de Mathématiques Appliquées at the Université Bordeaux 1 for support during the final phases of the project. TC and CG thank the Department of Mathematics, Statistics and Computere Science at the University of Illinois at Chicago for support and hospitality during part of this collaboration.

REFERENCES

[1] M. Abramowitz and I. A. Stegun (ed.), "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables," Dover, New York, 1965.

[2] A. A. Alazman, J. P. Albert, J. L. Bona, M. Chen and J. Wu, *Comparisons between the BBM–equation and a Boussinesq system*, Adv. Differential Equations, **11** (2006), 121–166.

[3] B. Alvarez-Samaniego and D. Lannes, *Large time existence for 3D water-wavesand asymptotics*, Invent. Math, **171** (2008), 485–541.

[4] [10.1098/rsta.1972.0032] T. B. Benjamin, J. L. Bona and J. J. Mahony, *Model equations for long waves in nonlinear, dispersive media*, Philos. Trans. Roy. Soc. London Ser. A, **272** (1972), 47–78.

[5] J. L. Bona and P. J. Bryant, *A mathematical model for long waves generated by wavemakers innon-linear dispersive systems*, Proc. Cambridge Philos. Soc., **73** (1973), 391–405.

[6] J. L. Bona and M. Chen, *A Boussinesq system for two-way propagation of nonlinear dispersive waves*, Phys. D, **116** (1998), 191–224.

[7] J. L. Bona, M. Chen and J.-C. Saut, *Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. I: Derivation and linear theory*, J. Nonlinear Sci., **12** (2002), 283–318.

[8] J. L. Bona, M. Chen and J.-C. Saut, *Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. II: The nonlinear theory*, Nonlinearity, **17** (2004), 925–952.

[9] J. L. Bona, T. Colin and D. Lannes, *Long wave approximation for water waves*, Arch. Ration. Mech. Anal., **178** (2005), 373–410.

[10] J. L. Bona and V. A. Dougalis, *An initial and boundary value problem for amodel equation for the propagation of long waves*, J. Math. Anal. Appl., **75** (1980), 503–522.

[11] J. L. Bona and H. Kalisch, *Models for internal waves in deep water*, Discrete Contin. Dynam. Systems, Series B, **6** (2000), 1–20.

[12] J. L. Bona, S. V. Rajopadhye and M. E. Schonbek, *Models for propagation of bores. I. Two dimensional theory*, Differential Integral Equations, **7** (1994), 699–734.

[13] J. L. Bona, S. M. Sun and B. Y. Zhang, *Conditional and unconditional well-posedness for nonlinear evolution equations*, Adv. Differential Equations, **9** (2004), 241–265.

[14] J. L. Bona and N. Tzvetkov, *Sharp well-posedness results for the BBM equation*, Discrete Contin. Dyn. Systems, Series A, **23** (2009), 1241–1252.

[15] V. A. Dougalis, D. E. Mitsotakis and J.-C. Saut, *On some Boussinesq systems in two space dimensions: theory and numerical analysis*, M2AN Math. Model. Numer. Anal., **41** (2007), 825–854.

[16] V. A. Dougalis, D. E. Mitsotakis and J.-C. Saut, *On initial-boundary value problems for a Boussinesq system of BBM–BBM type in a plane domain*, Discrete Contin. Dynam. Systems, Series A, **23** (2009), 1191–1204.

[17] I. S. Gradshteyn and I. M. Ryzhik, "Table of Integrals, Series, and Products," 7^{th} edition (prepared by Yu. V. Geronimus and M. Yu. Tseytlin, transl. ed. A. Jeffrey andD. Zwillinger), Elsevier/Academic Press, Amsterdam, 2007.

[18] B. B. Kadomtsev and V. I. Petviashvili, *Stability of solitary waves in weakly dispersing media*, Doklady Akademii Nauk SSSR, **192** (1970), 753–756. (Russian) (transl. Soviet Phys. Dokl., **15** (1970), 539–541).

[19] T. Kato, *On Nonlinear Schrödinger Equations. II. H^s – Solutions and unconditional well-posedness*, J. Anal. Math., **67** (1995), 281–306.

[20] D. Lannes and J.-C. Saut, *Weakly transverse Boussinesq systems and the KP approximation*, Nonlinearity, **19** (2006), 2853–2875.

[21] M. Ming, J.-C. Saut and P. Zhang, *Long time existence of solutions to Boussinesq systems*, Submitted.

[22] L. Molinet, J.-C. Saut and N. Tzvetkov, *Remarks on the mass constraint for KP–type equations*, SIAM J. Math. Anal., **39** (2007), 627–641.

[23] D. H. Peregrine, *Calculation of the development of an undular bore*, J. Fluid Mechanics, **25** (1966), 321–330.

[24] S. V. Rajopadhye, *Propagation of bores. II. Three–dimensional theory*, Nonlinear Anal., **27** (1996), 963–986.

[25] S. V. Rajopadhye, *Some models for the propagation of bores*, J. Differential Equations, **217** (2005), 179–203.

[26] J.-C. Saut and L. Xu, *The Cauchy problem on large time for surface waves Boussinesq systems*, J. Math. Pures Appl. **97** (2012), 635–662.

Received July 2011; revised May 2012

E-mail address: bona@math.uic.edu
E-mail address: thierry.colin@math.u-bordeaux1.fr
E-mail address: colette.guillope@u-pec.fr