Nonlinearity **17** (2004) 925–952 PII: S0951-7715(04)63954-5

Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media: II. The nonlinear theory

J L Bona[1], **M Chen**[2] **and J-C Saut**[3]

[1] Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, IL 60607, USA
[2] Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
[3] UMR de Mathématiques, Université de Paris-Sud, Bâtiment 425, 91405 Orsay, France

Received 22 May 2003, in final form 5 December 2003
Published 27 February 2004
Online at stacks.iop.org/Non/17/925 (DOI: 10.1088/0951-7715/17/3/010)

Recommended by F Merle

Abstract
In part I of this work (Bona J L, Chen M and Saut J-C 2002 Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media I: Derivation and the linear theory *J. Nonlinear Sci.* **12** 283–318), a four-parameter family of Boussinesq systems was derived to describe the propagation of surface water waves. Similar systems are expected to arise in other physical settings where the dominant aspects of propagation are a balance between the nonlinear effects of convection and the linear effects of frequency dispersion. In addition to deriving these systems, we determined in part I exactly which of them are linearly well posed in various natural function classes. It was argued that linear well-posedness is a natural necessary requirement for the possible physical relevance of the model in question.

In this paper, it is shown that the first-order correct models that are linearly well posed are in fact locally nonlinearly well posed. Moreover, in certain specific cases, global well-posedness is established for physically relevant initial data.

In part I, higher-order correct models were also derived. A preliminary analysis of a promising subclass of these models shows them to be well posed.

Mathematics Subject Classification: 35B35, 35Q53, 76B05, 76B15

1. Introduction

This paper is a continuation of our earlier work [11], where we put forward a formal derivation of the four-parameter system

$$\eta_t + u_x + (u\eta)_x + au_{xxx} - b\eta_{xxt} = 0, \qquad u_t + \eta_x + uu_x + c\eta_{xxx} - du_{xxt} = 0 \tag{1.1}$$

to approximate the motion of small-amplitude long waves on the surface of an ideal fluid under the force of gravity and in situations where the motion is sensibly two dimensional. Here, the independent variable, x, is proportional to distance in the direction of propagation while t is proportional to elapsed time. The dependent variables η and u have the following physical interpretation. The quantity $\eta(x,t)+h_0$ corresponds to the total depth of the liquid at the point x at time t, where h_0 is the undisturbed water depth (in the non-dimensionalization appearing here and in part I, h_0 becomes equal to 1). The variable $u(x,t)$ represents the horizontal velocity at the point $(x,y)=(x,\theta h_0)$ (y is the vertical coordinate, with $y=0$ corresponding to the channel bottom or sea bed) at time t. Thus u is the horizontal velocity field at the height θh_0, where θ is a fixed constant in the interval [0, 1]. The constants a, b, c, d satisfy the consistency conditions

$$a+b=\tfrac{1}{2}(\theta^2-\tfrac{1}{3}), \qquad c+d=\tfrac{1}{2}(1-\theta^2)\geqslant 0. \tag{1.2}$$

Higher-order systems that are formally accurate through quadratic order in the small-amplitude and long-wavelength parameters were also derived in our earlier paper [11], henceforth referred to as part I. Equations such as our *abcd*-systems in which surface tension effects are approximately accounted for have recently been derived and analysed in [5].

While having a collection of models to describe a particular wave regime might in principle be useful, the plethora of models represented in (1.1), all of which are formally equivalent, may at first sight be a little daunting. Moreover, the models appearing in (1.1) are all obtained, one from the other, by various linear approximations of the dependent variables. If nonlinear changes are countenanced, larger classes of formally correct approximations emerge (see [6, 12]).

To begin sorting out the prospects of the various representations of dispersion, we provided in part I an analysis of (1.1) linearized around the rest state. It transpired that only a subclass of the systems in (1.1)–(1.2) is linearly well posed for the initial-value problem,

$$\eta(x,0)=\varphi(x), \qquad u(x,0)=\psi(x), \tag{1.3}$$

for $x\in\mathbb{R}$. These are the systems whose dispersive coefficients fall into one of the following three categories:

(C1) $b\geqslant 0,\quad d\geqslant 0,\quad a\leqslant 0,\quad c\leqslant 0,$
(C2) $b\geqslant 0,\quad d\geqslant 0,\quad a=c>0,$
(C3) $b=d<0,\quad a=c>0.$

As was clear in our analysis of the associated linear problems in part I, these are composite classes and their mathematical properties depend very much upon which of the various coefficients a, b, c, d is non-zero. In part I, the analysis was broken into various subclasses to be reviewed presently. The same scheme of analysis will be apparent here as we sometimes rely upon the existing linear theory.

The presumption expressed in part I was that if a particular model fails to be linearly well posed, then it is unlikely to be useful in practical situations. The obverse of this latter presumption is that those that are linearly well posed are indeed potentially suitable for use in practical situations. There is an obvious gap in this heuristic reasoning. It could be that a particular model is linearly well posed but not nonlinearly so.

Even if a particular member of the class of linearly well-posed problems is locally well posed in time, there is a more subtle point that could render it unsuitable as a model, which we now explain. In the scaling that renders the equations in the form (1.1), one does not see the Boussinesq regime of small amplitude and long wavelength. In this scaling, these presumptions reside in the initial data (1.3), which should be of the form

$$\varphi(x)=\epsilon\Phi(\epsilon^{1/2}x), \qquad \psi(x)=\epsilon\Psi(\epsilon^{1/2}x) \tag{1.4}$$

to correspond to the wave regime assumed in the formal derivation of the models (1.1) from the full, two-dimensional Euler equations. Here ϵ is a small positive parameter that characterizes both the small-amplitude and long-wavelength assumptions. More precisely, in part I, we took ϵ to be a/h_0 where h_0 is the undisturbed depth and a is the maximum wave amplitude obtaining in the flow. The representative wavelength in the motion is restricted by the second small parameter $(h_0/\lambda)^2$. It is related to ϵ via the formula

$$\epsilon = S\frac{h_0^2}{\lambda^2},$$

where the non-dimensional constant S is usually called the Stokes number. The Boussinesq regime is characterized by ϵ being small and by S being of order 1. The order 1 constant S is incorporated into the order 1 functions Φ and Ψ, thus disappearing from sight. In versions (1.1), (1.4) of the Boussinesq systems, there are three important timescales, namely $T_0 = \epsilon^{-1/2}$, $T_1 = \epsilon^{-3/2}$ and $T_2 = \epsilon^{-5/2}$.

On the timescale T_0, nonlinear and dispersive effects as encoded in (1.1) in the nonlinear terms and the terms carrying the coefficients a, b, c, d, respectively, formally make only an order ϵ relative difference to the solution of the linear wave equation

$$\eta_t + u_x = 0, \qquad u_t + \eta_x = 0.$$

This is the regime of elementary engineering hydraulics. On the timescale T_1, nonlinear and dispersive effects can make an order 1 relative contribution to the wave motion. This is the Boussinesq timescale. Finally, on the timescale T_2, the effects ignored in the derivation of (1.1) may make an order 1 relative difference to the waves. On this timescale, the model may be unreliable as a depiction of reality and the higher-order models might well be needed. The point of this small diversion is that local well-posedness of one or another of the models in (1.1) is certainly a necessary condition for its practical utility. However, to be truly useful, one also requires that the local theory persist at least up to the timescale T_1, and preferably to the timescale T_2, for initial disturbances of the form (1.4) when Ψ and Φ are smooth, order 1 functions. Indeed, to obviate these considerations, members of the class (1.1) having globally defined solutions corresponding to data of the form (1.4), where ϵ need not be microscopically small, may be preferable. While this point is important, we will not enter into its mathematical elaboration in this paper. However, an extended discussion of this issue will be provided in the forthcoming papers [2, 12].

Thus, our purpose here is simply to demonstrate that for the pure initial-value problem (1.1)–(1.3), all the models satisfying (C1) and (C2) are at least locally nonlinearly well posed. (The case (C3), which indeed is linearly well posed as shown in part I, presents the singular operators $I + \lambda\partial_x^2$, where $\lambda = a$ or $\lambda = -b$ is positive. As such an equation is unlikely to be implemented as a practically useful description, we leave this case aside.) As mentioned above, the question of the timescales over which smooth solutions exist is left for future work. In a way, our results are bad news as they mean no further reduction in the wealth of models is effected by this requirement. On the other hand, questions of imposition of lateral boundary conditions and the like, which are not dealt with in this study, may have the salutary effect of limiting the possibilities. Indeed, for certain well-posed models, more boundary conditions are needed to specify a solution than can be conveniently determined by laboratory or field measurements (see, e.g. [29, 30]).

The plan of this paper is as follows. In section 2, three of the most prominent subclasses, namely BBM-type, KdV-type and what we call weakly dispersive systems are studied. The BBM-type falls to a contraction-mapping argument, while the KdV-type and the weakly dispersive systems are analysed by realizing them as systems coupled only through nonlinear terms. In section 3, the remainder of the systems are analysed using energy methods.

Hamiltonian structure is brought to the fore in section 4 and its consequences for global well-posedness addressed. Section 5 features a preliminary sortie into theory for the higher-order accurate systems developed also in part I. This paper concludes with a brief retrospective and a discussion of future lines of research.

2. Local well-posedness of the full nonlinear systems

In this and the next section, we show that the linearly well posed, first-order correct systems in (C1) and (C2) are also locally well posed when nonlinear effects are included. The systems under consideration are (1.1) with the auxiliary initial conditions (1.3), where a, b, c, d satisfy (1.2) and one of the linear well-posedness conditions (C1) and (C2). Three important particular classes of systems will be considered in this section and the remaining systems will be considered in the next section.

Currently standard notation will be used. The $L_p = L_p(\mathbb{R})$-norm will be denoted as $|\cdot|_p$ for $1 \leqslant p \leqslant \infty$, and the $H^s = H^s(\mathbb{R})$-norm will be denoted by $\|\cdot\|_s$. Note the L_2-norm of a function f is written both as $|f|_2$ and $\|f\|_0$. The product space $X \times X$ will sometimes be abbreviated as X^2, and a function $\boldsymbol{f} = (f_1, f_2)$ in X^2 carries the norm

$$\|\boldsymbol{f}\|_X \equiv \|\boldsymbol{f}\|_{X^2} \equiv (\|f_1\|_X^2 + \|f_2\|_X^2)^{1/2}.$$

We will use C to denote various constants whose value may change with each appearance.

2.1. Purely BBM-type Boussinesq systems

In this subsection, consideration is given to the cases where

$$a = c = 0, \qquad b > 0, \qquad d > 0.$$

In these cases, the initial-value problem can be written, as in Benjamin *et al* [8] and Bona and Chen [10], in the form

$$\eta_t + (I - b\partial_x^2)^{-1}\partial_x(u + u\eta) = 0, \qquad u_t + (I - d\partial_x^2)^{-1}\partial_x(\eta + \tfrac{1}{2}u^2) = 0. \tag{2.1}$$

The following local well-posedness result improves upon the one developed in [10].

Theorem 2.1. *Let $s \geqslant 0$ be given. For any $(\varphi, \psi) \in H^s(\mathbb{R})^2$, there exists $T > 0$ and a unique solution (η, u) of (2.1)–(1.3) that satisfies $(\eta, u) \in C(0, T; H^s(\mathbb{R}))^2$. Additionally, when $s > \frac{1}{2}$, then $(\partial_t^k \eta, \partial_t^k u) \in C(0, T; H^{s+1}(\mathbb{R}))^2$ for $k = 1, 2, \ldots$. Moreover, the correspondence $(\varphi, \psi) \mapsto (\eta, u)$ is locally Lipschitz continuous.*

Proof. For $s \geqslant 0$, the notation X_T^s will serve as an abbreviation for the product space $C(0, T; H_s(\mathbb{R}))^2$. We will continue to state theorems without this abbreviation, however, for the convenience of the casual reader.

Consider first the case with $s = 0$. Theorem 2.1 is proved using a contraction-mapping argument in a closed ball in the space X_T^0, where $T > 0$ will be determined presently.

Fix $(\varphi, \psi) \in L_2(\mathbb{R})^2$. Let $(\tilde{\eta}, \tilde{u}) \in X_T^0$ be given and let the operator $\mathcal{T} = (\mathcal{T}_1, \mathcal{T}_2)$ be defined by

$$\begin{aligned} \eta(x,t) = \mathcal{T}_1(\tilde{\eta}, \tilde{u}) &= \varphi(x) - \int_0^t [(I - b\partial_x^2)^{-1}\partial_x(\tilde{u} + \tilde{u}\tilde{\eta})](x,s)\,\mathrm{d}s, \\ u(x,t) = \mathcal{T}_2(\tilde{\eta}, \tilde{u}) &= \psi(x) - \int_0^t \left[(I - d\partial_x^2)^{-1}\partial_x\left(\tilde{\eta} + \frac{\tilde{u}^2}{2}\right)\right](x,s)\,\mathrm{d}s. \end{aligned} \tag{2.2}$$

The crucial quantity to evaluate is the L_2-norm of the integrands appearing in the definition of $\mathcal{T}$. First note that there are constants C_1, C_2 and C_3 depending only on $b > 0$ such that for any $f \in L_1(\mathbb{R})$

$$\|(I - b\partial_x^2)^{-1} f_x\|_0 \leqslant C_1\|f\|_{-1} \leqslant C_2|f|_1 \tag{2.3}$$

and for any $f \in L_2(\mathbb{R})$

$$\|(I - b\partial_x^2)^{-1} f_x\|_0 \leqslant C_1\|f\|_{-1} \leqslant C_3|f|_2. \tag{2.4}$$

Similar inequalities hold for the integrands appearing in $\mathcal{T}_2$, though the constants naturally depend on d. It thus follows that there is a constant C_4 depending only on b and d such that

$$\begin{aligned}\sup_{0\leqslant t\leqslant T} \|\mathcal{T}(\eta_1, u_1) - \mathcal{T}(\eta_2, u_2)\|_0 &\leqslant TC_4(1 + \|(\eta_1, u_1)\|_{X_T^0} + \|(\eta_2, u_2)\|_{X_T^0})\|(\eta_1, u_1) - (\eta_2, u_2)\|_{X_T^0} \\ &\leqslant TC_4(1 + 2R)\|(\eta_1, u_1) - (\eta_2, u_2)\|_{X_T^0},\end{aligned}$$

provided that (η_1, u_1) and (η_2, u_2) are selected from the closed ball $\bar{B}_R$ of radius R centred at $\mathbf{0}$ in X_T^0. Choosing $R = 2\|(\varphi, \psi)\|_0$ and $T = 1/(2C_4(1 + 2R))$, it is clear that

$$\|\mathcal{T}(\eta_1, u_1) - \mathcal{T}(\eta_2, u_2)\|_{X_T^0} \leqslant \tfrac{1}{2}\|(\eta_1, u_1) - (\eta_2, u_2)\|_{X_T^0}$$

and

$$\|\mathcal{T}(\eta_1, u_1)\|_{X_T^0} = \|\mathcal{T}(\eta_1, u_1) - \mathcal{T}(0, 0) + (\varphi, \psi)\|_{X_T^0} \leqslant TC_4(1 + 2R)R + \|(\varphi, \psi)\|_0 \leqslant R.$$

Therefore $\mathcal{T}$ is a contraction mapping from $\bar{B}_R$ to $\bar{B}_R$ in X_T^0, and the theorem follows. It is worth noting that T depends inversely on $\|(\varphi, \psi)\|_0$, but it does not approach infinity as the data approach 0, at least according to these estimates.

The Lipschitz continuity of the correspondence between the initial data and the associated solution follows readily because the solution is given as the fixed point of a contraction mapping $\mathcal{T} = \mathcal{T}_{(\varphi,\psi)}$ that depends smoothly on (φ, ψ). More precisely, let

$$(u, v) = \mathcal{T}_{(\varphi,\psi)}(u, v) \qquad \text{and} \qquad (\tilde{u}, \tilde{v}) = \mathcal{T}_{(\tilde{\phi},\tilde{\psi})}(\tilde{u}, \tilde{v})$$

be two solutions of (2.2) in X_T^0 corresponding to initial data (φ, ψ) and $(\tilde{\phi}, \tilde{\psi})$, respectively. Suppose both solutions lie in the closed ball $\bar{B}_R$ of radius R about the origin in X_T^0, where both $\mathcal{T}_{(\varphi,\psi)}$ and $\mathcal{T}_{(\tilde{\phi},\tilde{\psi})}$ are contractions, say with Lipschitz constant Θ and $\tilde{\Theta}$ in $(0, 1)$. The triangle inequality then implies that

$$\begin{aligned}\|(u, v) - (\tilde{u}, \tilde{v})\|_{X_T^0} &= \|\mathcal{T}_{(\varphi,\psi)}(u, v) - \mathcal{T}_{(\tilde{\phi},\tilde{\psi})}(\tilde{u}, \tilde{v})\|_{X_T^0} \\ &\leqslant \|\mathcal{T}_{(\varphi,\psi)}(u, v) - \mathcal{T}_{(\varphi,\psi)}(\tilde{u}, \tilde{v})\|_{X_T^0} + \|\mathcal{T}_{(\varphi,\psi)}(\tilde{u}, \tilde{v}) - \mathcal{T}_{(\tilde{\phi},\tilde{\psi})}(\tilde{u}, \tilde{v})\|_{X_T^0} \\ &\leqslant \Theta\|(u, v) - (\tilde{u}, \tilde{v})\|_{X_T^0} + \|(\varphi, \psi) - (\tilde{\phi}, \tilde{\psi})\|_{L_2(\mathbb{R})^2},\end{aligned}$$

from which Lipschitz continuity follows with a Lipschitz constant at most $1/(1 - \Theta)$.

The general case $s > 0$ can be proved by following the same line of argument using the space X_T^s instead of X_T^0 and the fact that for $s \geqslant -1$,

$$\|fg\|_s \leqslant C\|f\|_{s+1}\|g\|_{s+1}. \tag{2.5}$$

From this inequality follows an appropriate version of (2.3), (2.4), and the contraction-mapping principle may then be successfully implemented in X_T^s for suitably small values of T.

If $s > \frac{1}{2}$, solutions of the integral equation (2.2) lying in X_T^s for some $T > 0$ are automatically C^∞-functions of t. Indeed, $u, \eta \in C_b(\mathbb{R} \times [0, T])$ since $s > \frac{1}{2}$. An exercise in Fourier analysis reveals that for any $q > 0$ the operator $(1 - q\partial_x^2)^{-1}\partial_x$ is realized as a convolution, namely

$$(1 - q\partial_x^2)^{-1}\partial_x w = \int_{-\infty}^{\infty} K_q(x - y)w(y)\,\mathrm{d}y,$$

where

$$K_q(z) = \tfrac{1}{2}\sqrt{q}\,\mathrm{sgn}(z)\,\mathrm{e}^{-|z|/\sqrt{q}}.$$

As this kernel lies in $L_1(\mathbb{R})$, the fundamental theorem of calculus applied to (2.2) shows that η and u are differentiable with respect to t and

$$\eta_t(x,t) = -(I - b\partial_x^2)^{-1}\partial_x(u + u\eta) = -\int_{-\infty}^{\infty} K_b(x-y)(u(y,t) + u(y,t)\eta(y,t))\,\mathrm{d}y,$$

$$u_t(x,t) = -(I - d\partial_x^2)^{-1}\partial_x\left(\eta + \frac{u^2}{2}\right) = -\int_{-\infty}^{\infty} K_d(x-y)\left(\eta(y,t) + \frac{1}{2}u^2(y,t)\right)\mathrm{d}y.$$

Observe that these formulae imply that η_t and u_t lie in $C(0,T;H^{s+1}(\mathbb{R}))$, and so taking a temporal derivative yields one derivative more of spatial smoothness. Once η and u are seen to lie in $C^1(0,T;H^s(\mathbb{R}))$, one sees immediately that because K_b and $K_d \in L_1(\mathbb{R})$, the right-hand sides of the last pair of equations is differentiable with respect to t and their derivatives are given by differentiation under the integrand. Thus, η_t and u_t are differentiable with respect to t and

$$\eta_{tt} = -(I - b\partial_x^2)^{-1}\partial_x(u_t + u_t\eta + u\eta_t), \qquad u_{tt} = -(I - d\partial_x^2)^{-1}\partial_x(\eta_t + uu_t).$$

A straightforward induction then shows that

$$\partial_t^k\eta = -(I - b\partial_x^2)^{-1}\partial_t^{k-1}\partial_x(u + u\eta), \qquad \partial_t^k u = -(I - d\partial_x^2)^{-1}\partial_t^{k-1}\partial_x\left(\eta + \frac{u^2}{2}\right),$$

formulae that are easily justified by (2.3), (2.4), elementary facts about $H^s(\mathbb{R})$ for $s > \frac{1}{2}$ and induction.

To finish this subsection, we indicate the proof of formula (2.5) for the sake of completeness.

Lemma 2.2. *There are absolute constants C such that the following inequalities hold:*

(i) $\|fg\|_s \leqslant C\|f\|_s\|g\|_s$, *if* $s > \frac{1}{2}$,
(ii) $\|fg\|_s \leqslant C\|f\|_0\|g\|_0$, *if* $-1 \leqslant s < -\frac{1}{2}$,
(iii) $\|fg\|_s \leqslant C\|f\|_{(2s+1)/4}\|g\|_{(2s+1)/4}$, *if* $-\frac{1}{2} \leqslant s < 0$,
(iv) $\|fg\|_s \leqslant C\|f\|_{s+1}\|g\|_{s+1}$, *if* $0 \leqslant s \leqslant \frac{1}{2}$.

Proof. Part (i) is classical since $H^s(\mathbb{R}^n)$ is a multiplicative algebra for $s > n/2$ and here, $n = 1$.

Part (ii) can be proved by duality. Since

$$\|fg\|_s = \sup_{\|\varphi\|_{-s}\leqslant 1}\left|\int_{-\infty}^{\infty} fg\varphi\,\mathrm{d}x\right| \tag{2.6}$$

and $H^{-s}(\mathbb{R})$ is embedded in $L_\infty(\mathbb{R})$ for $-1 \leqslant s < -\frac{1}{2}$, one has

$$\left|\int_{-\infty}^{\infty} fg\varphi\,\mathrm{d}x\right| \leqslant |\varphi|_\infty\|f\|_0\|g\|_0 \leqslant C\|\varphi\|_{-s}\|f\|_0\|g\|_0 \leqslant C\|f\|_0\|g\|_0.$$

Part (iii) is proved by a similar argument. Note that $H^{-s}(\mathbb{R}) \subseteq L_{2/(2s+1)}(\mathbb{R})$ with continuous embedding for $-\frac{1}{2} < s < 0$ and $H^{1/2}(\mathbb{R})$ is embedded in $L_q(\mathbb{R})$ for any q in the range $2 \leqslant q < \infty$. The Hölder inequality implies

$$\left|\int_{-\infty}^{\infty} fg\varphi\,\mathrm{d}x\right| \leqslant |fg|_{2/(1-2s)}|\varphi|_{2/(1+2s)} \leqslant C|f|_{4/(1-2s)}|g|_{4/(1-2s)}\|\varphi\|_{-s}$$

and (iii) follows since $H^{(2s+1)/4}(\mathbb{R}) \subseteq L_{4/(1-2s)}(\mathbb{R})$ continuously.

For part (iv), use the classical estimates (see, e.g. [26])

$$\begin{aligned}\|fg\|_s &= \|J^s(fg)\|_0 \leqslant C(\|J^s f\|_0 |g|_\infty + \|J^s g\|_0 |f|_\infty)\\ &\leqslant C(\|f\|_s\|g\|_{1/2+\epsilon} + \|g\|_s\|f\|_{1/2+\epsilon}) \leqslant C\|f\|_{s+1}\|g\|_{s+1},\end{aligned}$$

where $J = (I - \partial_x^2)^{1/2}$ and ϵ is any positive number. Of course C depends on ϵ which is chosen to be $1/2$ here.

Remark 2.3. Further consideration of the solution mapping $(\varphi, \psi) \mapsto (u, v)$ shows this correspondence to be analytic as a map from $H^s(\mathbb{R})^2$ to $C(0, T; H^s(\mathbb{R}))^2$.

2.2. Weakly dispersive systems

We call the systems in (1.1) weakly dispersive when

$$b > 0 \qquad \text{and} \qquad d > 0.$$

The purely BBM-type Boussinesq systems studied in section 2.1 are special cases wherein $a = c = 0$.

Taking the Fourier transform with respect to x, the weakly dispersive systems may be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\hat{\eta}\\ \hat{u}\end{pmatrix} + \mathrm{i}k\mathcal{A}(k)\begin{pmatrix}\hat{\eta}\\ \hat{u}\end{pmatrix} = -\mathrm{i}k\begin{pmatrix}\dfrac{1}{(1+bk^2)}\widehat{u\eta}\\ \dfrac{1}{2}\dfrac{1}{(1+dk^2)}\hat{u}^2\end{pmatrix}, \tag{2.7}$$

where

$$\mathcal{A}(k) = \begin{pmatrix}0 & \omega_1(k)\\ \omega_2(k) & 0\end{pmatrix}$$

and

$$\omega_1(k) = \frac{1 - ak^2}{1 + bk^2}, \qquad \omega_2(k) = \frac{1 - ck^2}{1 + dk^2}.$$

It is convenient to decouple the linear part of the system as in part I; hence, consider the change of variables

$$\eta = \mathcal{H}(v + w) \qquad \text{and} \qquad u = v - w, \tag{2.8}$$

where $\mathcal{H}$ is the Fourier multiplier defined by

$$\widehat{\mathcal{H}g}(k) = h(k)\hat{g}(k) \qquad \text{with } h(k) = \left(\frac{\omega_1(k)}{\omega_2(k)}\right)^{1/2}. \tag{2.9}$$

In terms of these new variables, equation (2.7) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\hat{v}\\ \hat{w}\end{pmatrix} + \mathrm{i}k\begin{pmatrix}\sigma(k) & 0\\ 0 & -\sigma(k)\end{pmatrix}\begin{pmatrix}\hat{v}\\ \hat{w}\end{pmatrix} = -\mathrm{i}kP^{-1}\begin{pmatrix}\dfrac{1}{(1+bk^2)}\widehat{u\eta}\\ \dfrac{1}{2}\dfrac{1}{(1+dk^2)}\hat{u}^2\end{pmatrix}, \tag{2.10}$$

where

$$\sigma(k) = (\omega_1(k)\omega_2(k))^{1/2} \qquad \text{and} \qquad P^{-1} = \frac{1}{2}\begin{pmatrix}\left(\dfrac{\omega_2(k)}{\omega_1(k)}\right)^{1/2} & 1\\ \left(\dfrac{\omega_2(k)}{\omega_1(k)}\right)^{1/2} & -1\end{pmatrix}.$$

Once the form (2.10) is appreciated, it is natural to study these systems according to the order of the pseudo-differential operator $\mathcal{H}$.

Case I: $\mathcal{H}$ has order 0. The requirement of linear well-posedness put forward in part I as essential for physical relevance implies conditions (1.2) and (C1) or (C2) have to be satisfied. The operator $\mathcal{H}$ will have order 0 exactly when

$$\begin{aligned}&a < 0, \quad c < 0, \quad b > 0, \quad d > 0 \quad \text{(the 'generic' case),} \qquad \text{or}\\ &a = c > 0, \quad b > 0, \quad d > 0.\end{aligned} \tag{2.11}$$

The following result, which was established in part I, lemma 5.6, will find use in demonstrating the well-posedness of (1.1)–(2.11).

Lemma 2.4. *If $\mathcal{H}$ has order 0, then it is a bounded mapping of $L_p(\mathbb{R})$ for $1 \leqslant p \leqslant \infty$. The operator $\mathcal{H}$ is also a bounded mapping with bounded inverse of $H^s(\mathbb{R})$ onto itself for $s \geqslant 0$.*

Theorem 2.5. *Assume that (2.11) holds. Let $s \geqslant 0$ and $(\varphi, \psi) \in H^s(\mathbb{R})^2$. Then there exist $T > 0$ and a unique solution pair (η, u) in $C(0, T; H^s(\mathbb{R}))^2$ of (1.1)–(1.3). Additionally, $(\eta_t, u_t) \in C(0, T; H^{s-1}(\mathbb{R}))^2$. Moreover, the correspondence associating initial data to the solution is locally Lipschitz continuous.*

Proof. Consider the equivalent system (2.10) completed with initial values v_0, w_0, where

$$v_0 = \frac{\mathcal{H}^{-1}(\varphi) + \psi}{2} \in H^s(\mathbb{R}) \qquad \text{and} \qquad w_0 = \frac{\mathcal{H}^{-1}(\varphi) - \psi}{2} \in H^s(\mathbb{R}).$$

Taking the inverse Fourier transform, (2.10) is written in the form

$$\frac{\partial}{\partial t}\begin{pmatrix} v \\ w \end{pmatrix} + \mathcal{B}\begin{pmatrix} v \\ w \end{pmatrix} = \mathcal{F}\begin{pmatrix} v \\ w \end{pmatrix}, \tag{2.12}$$

where $\mathcal{B}$ is the skew-adjoint operator with symbol

$$\mathrm{i}k\begin{pmatrix} \sigma(k) & 0 \\ 0 & -\sigma(k) \end{pmatrix}$$

and

$$\mathcal{F}\begin{pmatrix} v \\ w \end{pmatrix} = -\mathcal{P}^{-1}\begin{pmatrix} (I - b\partial_x^2)^{-1}\partial_x[(v - w)\mathcal{H}(v + w)] \\ (I - d\partial_x^2)^{-1}(v - w)\partial_x(v - w) \end{pmatrix}.$$

The operator $\mathcal{B}$ is sometimes called the dispersive operator and $\mathcal{P}^{-1}$ is the pseudo-differential operator having symbol P^{-1}. Denoting by $S(t)$ the group generated by $\mathcal{B}$, it is clear that $S(t)$ is a unitary group on $H^s(\mathbb{R})^2$ for any $s \in \mathbb{R}$.

By Duhamel's formula, (2.12) is equivalent to

$$\begin{pmatrix} v \\ w \end{pmatrix} = S(t)\begin{pmatrix} v_0 \\ w_0 \end{pmatrix} + \int_0^t S(t - s)\mathcal{F}\begin{pmatrix} v \\ w \end{pmatrix} \mathrm{d}s.$$

Using (2.4), (2.5) and the boundedness of $\mathcal{H}$ and $\mathcal{P}^{-1}$, one easily determines that

$$\mathcal{F}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and that there is a constant C for which

$$\left\| \mathcal{F}\begin{pmatrix} f_1 \\ g_1 \end{pmatrix} - \mathcal{F}\begin{pmatrix} f_2 \\ g_2 \end{pmatrix} \right\|_{H^s(\mathbb{R})^2} \leqslant CR\|(f_1, g_1) - (f_2, g_2)\|_{H^s(\mathbb{R})^2}$$

whenever (f_1, g_1) and (f_2, g_2) are selected from the closed ball $\bar{B}_R$ of radius R centred at $\mathbf{0}$ in X_T^s. For fixed $(v_0, w_0) \in H^s(\mathbb{R})^2$, one can prove on the basis of the last inequality, just as in section 2.1, that the mapping $(\tilde{v}, \tilde{w}) \mapsto (v, w)$, where

$$\begin{pmatrix} v \\ w \end{pmatrix} = S(t) \begin{pmatrix} v_0 \\ w_0 \end{pmatrix} + \int_0^t S(t-s)\mathcal{F} \begin{pmatrix} \tilde{v} \\ \tilde{w} \end{pmatrix} \mathrm{d}s$$

is a contraction of $\bar{B}_R$ into itself for R sufficiently large and T sufficiently small. Indeed, we may choose T to be proportional to a quantity of the form $1/(1+\|(\varphi, \psi)\|_{H^s(\mathbb{R})^2})$. Theorem 2.5 follows immediately by using (2.8) and lemma 2.4.

Case II: $\mathcal{H}$ has order -1. This corresponds to the cases where

$$b > 0, \qquad d > 0, \qquad c < 0, \qquad a = 0.$$

Theorem 2.6. *Let $(\varphi, \psi) \in H^{s+1} \times H^s$, $s \geqslant 0$. Then there exist $T > 0$ and a unique solution (η, u) in $C(0, T; H^{s+1}(\mathbb{R})) \times C(0, T; H^s(\mathbb{R}))$ of (1.1)–(1.3). Moreover, $(\eta_t, u_t) \in C(0, T; H^{s+1}(\mathbb{R})) \times C(0, T; H^s(\mathbb{R}))$. The correspondence between initial values and solutions is locally Lipschitz.*

Proof. The assumptions imply that $(v_0, w_0) \in H^s(\mathbb{R})^2$. Using the same argument as for case I, one can prove that there exists $T > 0$ and a unique solution (v, w) that satisfies $(v, w) \in X_T^s$. The first part of the theorem follows after the change of variable $\eta = \mathcal{H}(v + w)$ and $u = v - w$.

Note $\eta_t = (I - b\partial_x^2)^{-1}(-u_x - (u\eta)_x)$. We show $\eta_t \in H^s$ by arguing as follows:

- when $s > \frac{1}{2}$, u and $(u\eta) \in H^s$, and so u_x and $(u\eta)_x \in H^{s-1}$. Therefore $\eta_t \in H^{s+1}$;
- when $0 \leqslant s \leqslant \frac{1}{2}$, the u_x term can be treated as before. Since $\eta \in H^{s+1} \subset H^1$, one observes that $u \to u\eta$ maps L_2 into L_2 and H^1 into H^1. By interpolation, $u \to u\eta$ maps H^s into H^s. Thus $u\eta \in H^s$, and so $(u\eta)_x \in H^{s-1}$. Therefore $\eta_t \in H^{s+1}$.

Similarly, note $u_t = -(I - d\partial_x^2)^{-1}(\eta_x + \eta_{xxx} + (u^2/2)_x)$. Since $\eta_x + \eta_{xxx} \in H^{s-2}$, so $(I - d\partial_x^2)^{-1}(\eta_x + \eta_{xxx}) \in H^s$. For the last term, we again separate the cases:

- when $s > \frac{1}{2}$, $u^2 \in H^s$, and so $(u^2)_x \in H^{s-1}$; therefore $(I - d\partial_x^2)^{-1}(u^2)_x \in H^{s+1} \subset H^s$;
- when $0 \leqslant s \leqslant \frac{1}{2}$, by lemma 2.2 (ii)

$$\|u^2\|_{s-1} \leqslant C\|u\|_0^2 \leqslant C\|u\|_s^2.$$

Therefore,

$$\|(I - d\partial_x^2)^{-1}\partial_x(u^2)\|_s \leqslant C\|u\|_s^2.$$

It then follows $u_t \in H^s$. The proof of the theorem is thus complete.

Remark 2.7. This is the system derived by Bona and Smith [18] when $b = d = -c = \frac{1}{3}$.

Case III: $\mathcal{H}$ has order 1. This corresponds to the cases where

$$b > 0, \qquad d > 0, \qquad c = 0, \qquad a < 0.$$

Theorem 2.8. *Let $(\varphi, \psi) \in H^s \times H^{s+1}$, where $s \geqslant 0$. Then there exist $T > 0$ and a unique solution (η, u) in $C(0, T; H^s(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$ of (1.1)–(1.3). Moreover, $(\eta_t, u_t) \in C(0, T; H^s(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$. The correspondence between initial data and the associated solution is locally Lipschitz continuous.*

Proof. The proof mirrors that of theorem 2.6 and is therefore omitted.

2.3. Purely KdV-type Boussinesq systems

In this section, the systems in (1.1) with

$$b = d = 0, \qquad a \neq 0, \qquad c \neq 0$$

are considered. In view of the constraints (1.2), (C1) and (C2), the only admissible case is when $a = c > 0$, which means $\theta^2 = \frac{2}{3}$ and $a = c = \frac{1}{6}$.

By making a simple additional scaling, one may assume that $a = c = 1$ and thus the system under consideration takes the tidy form

$$\eta_t + u_x + (u\eta)_x + u_{xxx} = 0, \qquad u_t + \eta_x + uu_x + \eta_{xxx} = 0. \tag{2.13}$$

Introduce again v and w by $\eta = v + w$ and $u = v - w$ to obtain the equivalent system,

$$\begin{aligned} &v_t + v_x + v_{xxx} + \tfrac{1}{2}[(v-w)(v+w)]_x + \tfrac{1}{2}(v-w)(v-w)_x = 0, \\ &w_t - w_x - w_{xxx} + \tfrac{1}{2}[(v-w)(v+w)]_x - \tfrac{1}{2}(v-w)(v-w)_x = 0. \end{aligned} \tag{2.14}$$

This is a system of two linear KdV-equations coupled through nonlinear terms. One can apply the theory developed in Kenig *et al* [32, 33] for the scalar KdV. In fact, the function spaces used in their works associated with the group $\mathrm{e}^{t\partial_x^3}$ are exactly the same as those associated with the groups $\mathrm{e}^{t(\partial_x+\partial_x^3)}$ and $\mathrm{e}^{-t(\partial_x+\partial_x^3)}$.

Theorem 2.9. *Let $(\varphi, \psi) \in H^s \times H^s$, $s > \frac{3}{4}$. Then there exist a value $T > 0$ and a unique solution (η, v) of (2.13) lying in $C(0, T; H^s(\mathbb{R}))^2$. Moreover, $(\eta_t, u_t) \in C(0, T; H^{s-3}(\mathbb{R}))^2$. The correspondence between initial data and the associated solution is analytic.*

Remark 2.10. One can imagine strategies to obtain local well-posedness in H^s for $s \geqslant 0$, for instance. For example, one might try to apply Bourgain's methods [21]. Note however, the arguments in the last quoted reference would need modification. In fact, the Bourgain spaces associated with $\partial_t + \partial_x + \partial_x^3$ and with $\partial_t - \partial_x - \partial_x^3$ are not the same, and different bilinear estimates must be established (see appendix A in Saut and Tzvetkov [41] for such a modification in a simpler context).

3. The other admissible Boussinesq systems

In this section, we investigate the admissible cases not covered in section 2. They are classified according to the order of the pseudo-differential operator $\mathcal{H}$ defined in section 2.

3.1. The Boussinesq system when $\mathcal{H}$ has order 2

This subsection is concerned with cases where

$$c = 0, \qquad b = 0, \qquad a < 0, \qquad d > 0.$$

The system then has the form

$$\begin{aligned} &\eta_t + u_x + (u\eta)_x + au_{xxx} = 0, \qquad u_t + \eta_x + uu_x - du_{xxt} = 0, \\ &\eta(x, 0) = \varphi(x), \qquad u(x, 0) = \psi(x). \end{aligned} \tag{3.1}$$

The next result provides a useful local well-posedness result for such systems.

Theorem 3.1. *Assume $(\varphi, \psi) \in H^s(\mathbb{R}) \times H^{s+2}(\mathbb{R})$ where $s \geqslant 1$. Then there exist $T > 0$ and a unique solution (η, u) of (3.1) that is such that $(\eta, u) \in C(0, T; H^s(\mathbb{R})) \times C(0, T; H^{s+2}(\mathbb{R}))$. Moreover, $(\eta_t, u_t) \in C(0, T; H^{s-1}(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$. The solution depends continuously upon perturbations of the initial data in the relevant function classes.*

Proof. (i) *Uniqueness.* A uniqueness result will be established in the larger class $C(0,T;H^1(\mathbb{R})) \times C(0,T;H^2(\mathbb{R}))$. Note that then $(\eta_t,u_t) \in C(0,T;H^{-1}(\mathbb{R})) \times C(0,T;H^2(\mathbb{R}))$. Let (η_1,u_1) and (η_2,u_2) be two solutions in this class that achieve the same value at $t=0$ and let $\eta=\eta_1-\eta_2$ and $u=u_1-u_2$. Then (η,u) satisfies the system

$$\eta_t + u_x + (u_1\eta_1)_x - (u_2\eta_2)_x + au_{xxx} = 0, \qquad u_t + \eta_x + u_1(u_1)_x - u_2(u_2)_x - du_{xxt} = 0 \tag{3.2}$$

with zero initial data. We take the H^1–H^{-1} duality of η with the first equation in (3.2) and take the L_2 scalar product of the second equation with u. By adding the two resulting equations together, one obtains after suitable integrations by parts the equation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}[\eta^2+u^2+du_x^2]\,\mathrm{d}x = -\int_{-\infty}^{\infty}[(u_1\eta_1-u_2\eta_2)_x\eta + (u_1(u_1)_x - u_2(u_2)_x)]\,\mathrm{d}x.$$

Observe that

$$\int_{-\infty}^{\infty}(u_1\eta_1-u_2\eta_2)_x\eta\,\mathrm{d}x = \int_{-\infty}^{\infty}(u\eta_1+u_2\eta)_x\eta\,\mathrm{d}x = \int_{-\infty}^{\infty}\left(u_x\eta\eta_1 + u\eta(\eta_1)_x + \frac{1}{2}(u_2)_x\eta^2\right)\mathrm{d}x,$$

so that

$$\begin{aligned}\left|\int_{-\infty}^{\infty}(u_1\eta_1-u_2\eta_2)_x\eta\,\mathrm{d}x\right| &\leqslant |\eta_1|_\infty\|u_x\|_0\|\eta\|_0 + |u|_\infty\|\eta\|_0\|(\eta_1)_x\|_0 + \frac{1}{2}|(u_2)_x|_\infty\|\eta\|_0^2\\ &\leqslant C(\|u_x\|_0\|\eta\|_0 + \|u\|_1\|\eta\|_0 + \|\eta\|_0^2),\end{aligned}$$

where the constant depends on the norm of η_1 in $C(0,T;H^1(\mathbb{R}))$ and the norm of u_1 in $C(0,T;H^2(\mathbb{R}))$. In a similar fashion, one may check that

$$\left|\int_{-\infty}^{\infty}(u_1(u_1)_x - u_2(u_2)_x)u\,\mathrm{d}x\right| \leqslant C(\|u\|_0^2 + \|u\|_0\|u_x\|_0),$$

where the constants again depend on the norms of (η_1,u_1) and (η_2,u_2) in $C(0,T;H^1(\mathbb{R})) \times C(0,T;H^2(\mathbb{R}))$.

Letting $Y(t)=\int_{-\infty}^{\infty}[\eta^2+u^2+(d-a)u_x^2-adu_{xx}^2]\,\mathrm{d}x$, one finds $\mathrm{d}/\mathrm{d}tY(t) \leqslant CY(t)$. Uniqueness then follows from Gronwall's lemma.

Of course, this calculation is formal, but the continuous-dependence result coupled with further regularity of the solution associated with more regular data allows one to justify rigorously the conclusion.

(ii) *Existence.* This is carried out in the case $s=1$. The general case is similar. The first step is to derive formally *a priori* estimates on solutions of (3.1). As in part (i), multiply the first equation in (3.1) by η and the second equation by $u+au_{xx}$ to obtain, after integration by parts, the inequality

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}[\eta^2+u^2+(d-a)u_x^2-adu_{xx}^2]\,\mathrm{d}x &= \int_{-\infty}^{\infty}\left[\frac{a}{2}u_x^3-\frac{1}{2}u_x\eta^2\right]\mathrm{d}x\\ &\leqslant \frac{|a|}{2}|u_x|_\infty\|u_x\|_0^2 + \frac{1}{2}|u_x|_\infty\|\eta\|_0^2\\ &\leqslant C_0(\|u_x\|_0^{5/2}\|u_{xx}\|_0^{1/2} + \|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta\|_0^2).\end{aligned} \tag{3.3}$$

Since $a<0$ and $d>0$, this differential inequality yields an *a priori* bound in the space $L_\infty(0,T;L_2(\mathbb{R})) \times L_\infty(0,T;H^2(\mathbb{R}))$, on a time interval $[0,T]$ where T is inversely proportional to $\|\varphi\|_0+\|\psi\|_2$. To obtain an $H^1(\mathbb{R}) \times H^3(\mathbb{R})$ estimate, multiply the first equation of (3.1) by $-\eta_{xx}$, the second by $-au_{xxxx}$, integrate over $\mathbb{R}$, integrate by parts and sum the results to obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta_x^2 - au_{xx}^2 - adu_{xxx}^2)\,\mathrm{d}x &= \int_{-\infty}^{\infty}(u_x\eta_{xx} + (u\eta)_x\eta_{xx} + auu_xu_{xxxx})\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\left(-u_{xx}\eta_x - \frac{3}{2}u_x\eta_x^2 - u_{xx}\eta\eta_x + \frac{5a}{2}u_xu_{xx}^2\right)\mathrm{d}x \\ &\leqslant \frac{5|a|}{2}|u_x|_\infty\|u_{xx}\|_0^2 + \|u_{xx}\|_0\|\eta_x\|_0 + \frac{3}{2}|u_x|_\infty\|\eta_x\|_0^2 + |u_{xx}|_\infty\|\eta\|_0\|\eta_x\|_0 \\ &\leqslant C_1(1 + \|\eta_x\|_0 + \|\eta_x\|_0^2 + \|u_{xxx}\|_0^{1/2}\|\eta_x\|_0). \end{aligned} \tag{3.4}$$

In this instance, the constant depends upon the previously derived $L_\infty(0,T;L_2(\mathbb{R})) \times L_\infty(0,T;H^2(\mathbb{R}))$ bound. Gronwall's lemma thus yields an *a priori* estimate in $L_\infty(0,T;H^1(\mathbb{R})) \times L_\infty(0,T;H^3(\mathbb{R}))$, where the value of T is the same as that appearing in the estimate implied by (3.4) since C_1 is bounded at least on the time interval $[0,T]$.

The next step is to justify the *a priori* estimate. One way to attack this issue is to regularize the initial data φ and ψ and to consider a regularized version of (3.1). This can be accomplished by adding a parabolic term $\epsilon\partial_x^4\eta$, $\epsilon > 0$ or a dispersive regularizing term $-\epsilon\eta_{xxt}$ to the left-hand side of the first equation in (3.1). Existence and uniqueness of a smooth local solution for the regularized system are standard, and one gets a sequence of smooth approximate solutions $\{(\eta_\epsilon, u_\epsilon)\}_{\epsilon>0}$. If the term $-\epsilon\eta_{xxt}$ is used, this sequence of smooth approximate solutions is already obtained in section 2.2. By exactly the same procedure as above, one obtains a local bound on $(\eta_\epsilon, u_\epsilon)$ in the space $L_\infty(0,T;H^1(\mathbb{R})) \times L_\infty(0,T;H^3(\mathbb{R}))$, where $T > 0$ does not depend on ϵ. In a little more detail, if the dispersive regularization

$$\eta_t + u_x + (u\eta)_x + au_{xxx} - \epsilon\eta_{xxt} = 0, \qquad u_t + \eta_x + uu_x - du_{xxt} = 0$$

is favoured, then the analogue of (3.3) is

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}[\eta^2 + \epsilon\eta_x^2 + (d-a)u_x^2 - adu_{xx}^2]\,\mathrm{d}x \leqslant C(\|u_x\|_0^{5/2}\|u_{xx}\|_0^{1/2} + \|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta\|_0^2), \tag{3.5}$$

where the right-hand side is independent of $\epsilon > 0$. Thus, there obtain ϵ-independent bounds on $(\eta_\epsilon, u_\epsilon)$ in $L_\infty(0,T;L_2(\mathbb{R})) \times L_\infty(0,T;H^2(\mathbb{R}))$. Similarly, the analogue of (3.4) yields *a priori* deduced, ϵ-independent bounds on $(\eta_\epsilon, u_\epsilon)$ in the space $L_\infty(0,T;H^1(\mathbb{R})) \times L_\infty(0,T;H^3(\mathbb{R}))$. A standard application of the compactness method (see [35], chapter 1) yields the existence of a solution (η, u) in the space $L_\infty(0,T;H^1(\mathbb{R})) \times L_\infty(0,T;H^3(\mathbb{R}))$. If, instead, one passes to the limit strongly using the Bona–Smith technique [17], one adduces that, in fact, the solution (η, u) lies in $C(0,T;H^1(\mathbb{R})) \times C(0,T;H^3(\mathbb{R}))$. The preceding arguments go over with only mild modification to the case where $(\varphi, \psi) \in H^s(\mathbb{R}) \times H^{s+2}(\mathbb{R})$ for any $s \geqslant 1$. The advertised result, theorem 3.1, is thereby established.

Remark 3.2. The *a priori* estimate (3.3) implies the existence of a weak local solution (η, u) of (3.1) corresponding to initial data $(\varphi, \psi) \in L_2(\mathbb{R}) \times H^2(\mathbb{R})$ such that $(\eta, u) \in L_\infty(0,T;L_2(\mathbb{R})) \times L_\infty(0,T;H^2(\mathbb{R}))$, $(\eta_t, u_t) \in L_\infty(0,T;H^{-1}(\mathbb{R})) \times L_\infty(0,T;H^1(\mathbb{R}))$. We indicate briefly the main ingredient of the proof. As above, we obtain relative to a suitable approximate solution $(\eta_\epsilon, u_\epsilon)$ a local bound in the space $L_\infty(0,T;L_2(\mathbb{R})) \times L_\infty(0,T;H^2(\mathbb{R}))$, where $T > 0$ does not depend on ϵ. Passing to the limit is obvious except possibly for the term $(u_\epsilon\eta_\epsilon)_x$. We may assume, by taking a suitable subsequence, that η_ϵ converges weakly to η in $L_2(0,T;L_2(\mathbb{R}))$ and by compactness that u_ϵ converges strongly to u in $L_2(0,T;L_{2,\mathrm{loc}}(\mathbb{R}))$. We deduce that $u_\epsilon\eta_\epsilon$ converges to $u\eta$ in $D'((0,T)\times\mathbb{R})$ and thus $(u_\epsilon\eta_\epsilon)_x$ converges to $(u\eta)_x$ in $D'((0,T)\times\mathbb{R})$.

3.2. The Boussinesq system when $\mathcal{H}$ has order 1

The only admissible cases are the following:

$$a = c \geqslant 0, \qquad b = 0, \qquad d > 0 \tag{3.6}$$

and

$$a < 0, \qquad c < 0, \qquad b = 0, \qquad d > 0. \tag{3.7}$$

In (3.6), the special case

$$a = c = b = 0, \qquad d = \tfrac{1}{3} \tag{3.8}$$

corresponds to what is usually called the classical Boussinesq equation, attributed to Boussinesq. The initial-value problem for it has been studied by Schonbek [42] and Amick [3]. It deserves remark that this is not in fact the system written by Boussinesq [23]. He instead derived the system

$$\eta_t + q_x + (q\eta)_x = 0, \qquad q_t + \eta_x + qq_x + \eta_{xtt} = 0, \tag{3.9}$$

where $\eta(x, t)$ is as before the deviation of the free surface from its rest position at the point x at time t and $q(x, t)$ is the depth-averaged horizontal velocity at x at time t. The formulation appearing in (3.6) arises from recognizing that to the inherent order of approximation being kept, q coincides with the horizontal velocity u at height $(1/\sqrt{3})h_0$, where h_0 is the total depth, and that, to the relevant order, $\eta_{xtt} = -q_{xxt}$ (see [18]). In any event, here is the outcome of the analyses of Amick and Schonbek.

Theorem 3.3. *Suppose* $(\varphi, \psi) \in H^s(\mathbb{R}) \times H^{s+1}(\mathbb{R})$ *where* $s \geqslant 1$ *and that* $\inf_{x\in\mathbb{R}}\{1+\varphi(x)\} > 0$. *Then there is a unique solution* (η, u) *of the classical Boussinesq system*

$$\eta_t + u_x + (u\eta)_x = 0, \qquad u_t + \eta_x + uu_x - \tfrac{1}{3}u_{xxt} = 0,$$

which for any $T > 0$ *lies in* $C(0, T; H^s(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$ *and which takes on the initial values, namely* $\eta(x, 0) = \varphi(x), u(x, 0) = \psi(x)$ *for* $x \in \mathbb{R}$. *Additionally,* $(\eta_t, u_t) \in C(0, T; H^{s-1}(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$. *Moreover, given* $s \geqslant 1$ *and* $r > 0$, *the solution depends continuously upon the initial data in the class*

$$\mathcal{H}^s_r = \{(\varphi, \psi) \in H^s(\mathbb{R}) \times H^{s+1}(\mathbb{R}) : 1 + \varphi(x) \geqslant r \text{ for all } x\}.$$

Remark 3.4. The result stated above does not appear explicitly in the works of Schonbek [42] and Amick [3]. Their results are stated with data taken from $C_c^\infty(\mathbb{R})$, the infinitely differentiable functions with compact support. Our statement is a reformulation of their results that takes account of data with finite regularity. While the calculations in the papers of Schonbek and Amick are time-consuming to check, they implicitly contain theorem 3.3 for integer values $s = 1, 2, \ldots$. For fractional orders, a little more work is required, the details of which we pass over here (but see the proof of theorem 4.3 in the next section).

The models corresponding to the restrictions (3.7) and (3.6) with $a = c \neq 0$ have the form

$$\eta_t + u_x + (u\eta)_x + au_{xxx} = 0, \qquad u_t + \eta_x + uu_x + c\eta_{xxx} - du_{xxt} = 0. \tag{3.10}$$

Theorem 3.5. *For any* $s \geqslant 1$ *and* $(\varphi, \psi) \in H^s(\mathbb{R}) \times H^{s+1}(\mathbb{R})$, *there exist* $T > 0$ *and a unique solution* $(\eta, u) \in C(0, T; H^s(\mathbb{R})) \times C(0, T; H^{s+1}(\mathbb{R}))$ *of (3.10). Additionally,* $(\eta_t, u_t) \in C(0, T; H^{s-2}(\mathbb{R})) \times C(0, T; H^{s-1}(\mathbb{R}))$. *The mapping of initial data to its associated solution is continuous.*

Proof. (i) *Existence*. Consider the case $s = 1$ in the first instance. We content ourselves with providing the proof of an *a priori* estimate that may be justified just as in section 3.1 by regularizing the first equation in (3.10) with the term $-\epsilon\eta_{xxt}$. As we already have a theory in this case, it is simply a matter of allowing ϵ to tend to zero using the technique in Bona and Smith [17].

Multiply the first equation by $|c|\eta$, the second equation by $|a|u$, sum the results and integrate by parts to obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(|c|\eta^2 + |a|u^2 + |ad|u_x^2)\,\mathrm{d}x + |c|\int_{-\infty}^{\infty}(u\eta)_x\eta\,\mathrm{d}x + \int_{-\infty}^{\infty}(|a| - |c|)\eta u_x\,\mathrm{d}x = 0.$$

Since

$$\left|\int_{-\infty}^{\infty}(u\eta)_x\eta\,\mathrm{d}x\right| = \left|\frac{1}{2}\int_{-\infty}^{\infty}u_x\eta^2\,\mathrm{d}x\right| \leqslant \frac{1}{2}|u_x|_\infty\|\eta\|_0^2 \leqslant C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta\|_0^2$$

and

$$\left|\int_{-\infty}^{\infty}\eta u\,\mathrm{d}x\right| \leqslant C\|\eta\|_0\|u_x\|_0,$$

it follows that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(|c|\eta^2 + |a|u^2 + |ad|u_x^2)\,\mathrm{d}x \leqslant C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta\|_0^2. \tag{3.11}$$

Multiply the first equation in (3.10) by $-|c|\eta_{xx}$ and the second by $-|a|u_{xx}$ and integrate the results with respect to x over $\mathbb{R}$. After some cancellation, there appears the formula

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(|c|\eta_x^2 + |a|u_x^2 + |ad|u_{xx}^2)\,\mathrm{d}x &= \int_{-\infty}^{\infty}|c|(u\eta)_x\eta_{xx}\,\mathrm{d}x + \int_{-\infty}^{\infty}|a|uu_xu_{xx}\,\mathrm{d}x\\ &\quad + \int_{-\infty}^{\infty}(|a| - |c|)u_{xx}\eta_x\,\mathrm{d}x.\end{aligned} \tag{3.12}$$

Since

$$\int_{-\infty}^{\infty}(u\eta)_x\eta_{xx}\,\mathrm{d}x = \int_{-\infty}^{\infty}(u_x\eta\eta_{xx} + u\eta_x\eta_{xx})\,\mathrm{d}x = \int_{-\infty}^{\infty} -u_{xx}\eta\eta_x - \frac{3}{2}u_x\eta_x^2\,\mathrm{d}x,$$

it transpires that

$$\begin{aligned}\left|\int_{-\infty}^{\infty}(u\eta)_x\eta_{xx}\,\mathrm{d}x\right| &\leqslant \|u_{xx}\|_0|\eta|_\infty\|\eta_x\|_0 + \tfrac{3}{2}|u_x|_\infty\|\eta_x\|_0^2\\ &\leqslant C\|u_{xx}\|_0\|\eta\|_0^{1/2}\|\eta_x\|_0^{3/2} + C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta_x\|_0^2.\end{aligned}$$

Similarly,

$$\left|\int_{-\infty}^{\infty}u_{xx}\eta_x\,\mathrm{d}x\right| \leqslant C\|u_{xx}\|_0\|\eta_x\|_0$$

and

$$\left|\int_{-\infty}^{\infty}uu_xu_{xx}\,\mathrm{d}x\right| = \left|\frac{1}{2}\int_{-\infty}^{\infty}u_x^3\,\mathrm{d}x\right| \leqslant \frac{1}{2}|u_x|_\infty\|u_x\|_0^2 \leqslant C\|u_x\|_0^{5/2}\|u_{xx}\|_0^{1/2}.$$

Combining the preceding ruminations leads to the estimate

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(|c|\eta^2 + |a|u^2 + (|ad| + |a|)u_x^2 + |c|\eta_x^2 + |ad|u_{xx}^2)\,\mathrm{d}x\\ &\quad\leqslant C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta\|_0^2 + C\|u_{xx}\|_0\|\eta\|_0^{1/2}\|\eta_x\|_0^{3/2}\\ &\qquad + C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{1/2}\|\eta_x\|_0^2 + C\|u_x\|_0^{5/2}\|u_{xx}\|_0^{1/2}\\ &\quad\leqslant C(\|u_x\|_0^3 + \|u_{xx}\|_0^3 + \|\eta\|_0^3 + \|\eta_x\|_0^3).\end{aligned}$$

Setting

$$Y(t) = \int_{-\infty}^{\infty} (|c|\eta^2 + |a|u^2 + (|a| + |ad|)u_x^2 + |c|\eta_x^2 + |ad|u_{xx}^2)\,\mathrm{d}x$$

and because a, c and d are not zero, the last inequality implies that $Y'(t) \leqslant CY^{3/2}$, which provides the desired bound in $L_\infty(0, T; H^1) \times L_\infty(0, T; H^2)$ for T sufficiently small. Note that these bounds would be sufficient to infer an existence result by weak-compactness arguments in $L_\infty(0, T; H^1) \times L_\infty(0, T; H^2)$ for suitable values of $T > 0$. Following the line of argument put forward in Bona and Smith [17], existence in the smaller space $C(0, T; H^1) \times C(0, T; H^2)$ may be adduced. Continuous dependence of the solution (η, u) on variations of the initial data within their respective function classes follows because of the strong convergence and well-posedness of the regularized problem.

(ii) *Uniqueness*. Let (η_1, u_1), (η_2, u_2) be two solutions in $C(0, T; H^1) \times C(0, T; H^2)$ with the same initial values and let $\eta = \eta_1 - \eta_2$ and $u = u_1 - u_2$. By calculations similar to those appearing in the derivation of *a priori* bounds, there obtains at once the integral equality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} (|c|\eta^2 + |a|u^2 + |ad|u_x^2)\,\mathrm{d}x + |c|\int_{-\infty}^{\infty} [(u_1\eta)_x + (u\eta_2)_x]\eta\,\mathrm{d}x + \frac{1}{2}\int_{-\infty}^{\infty} |a|(u_1 + u_2)_x u^2\,\mathrm{d}x + \int_{-\infty}^{\infty} (|a| - |c|)\eta u_x\,\mathrm{d}x = 0.$$

The formal integration of η times the first equation has to be understood in the sense of the H^1–H^{-1} duality, and so the integration by parts is justified. Observe that

$$\left|\int_{-\infty}^{\infty} \eta u_x\,\mathrm{d}x\right| \leqslant C\|\eta\|_0\|u_x\|_0, \qquad \left|\int_{-\infty}^{\infty} (u_1\eta)_x\eta\,\mathrm{d}x\right| = \frac{1}{2}\left|\int_{-\infty}^{\infty} (u_1)_x\eta^2\,\mathrm{d}x\right| \leqslant C\|u_1\|_2\|\eta\|_0^2$$

and

$$\left|\int_{-\infty}^{\infty} (u_1 + u_2)_x u^2\,\mathrm{d}x\right| \leqslant C(\|u_1\|_2 + \|u_2\|_2)\|u\|_0^2.$$

On the other hand,

$$\int_{-\infty}^{\infty} (u\eta_2)_x\eta\,\mathrm{d}x = \int_{-\infty}^{\infty} u_x\eta_2\eta\,\mathrm{d}x + \int_{-\infty}^{\infty} u(\eta_2)_x\eta\,\mathrm{d}x,$$

whence

$$\begin{aligned}\left|\int_{-\infty}^{\infty} (u\eta_2)_x\eta\,\mathrm{d}x\right| &\leqslant |\eta_2|_\infty\|u_x\|_0\|\eta\|_0 + \|u\|_0^{1/2}\|u_x\|_0^{1/2}\|(\eta_2)_x\|_0\|\eta\|_0 \\ &\leqslant C(\|\eta\|_0^{3/2}\|\eta_x\|_0^{1/2}\|u_x\|_0) + \|u\|_0^{1/2}\|u_x\|_0^{1/2}\|(\eta_2)_x\|_0\|\eta\|_0).\end{aligned}$$

Similarly, one can show that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} (|c|\eta_x^2 + |a|u_x^2 + |ad|u_{xx}^2)\mathrm{d}x \leqslant C(\|u\|_0^3 + \|u_x\|_0^3 + \|u_{xx}\|_0^3 + \|\eta\|_0^3 + \|\eta_x\|_0^3).$$

Combining the above inequalities and using Gronwall's lemma, uniqueness follows.

Remark 3.6. Systems in (3.6) with $a = c \neq 0$ and in (3.7) are given the appellation Benjamin–Ono-type because they can be reduced to a pair of equations whose linearization uncouples to a pair of linear Benjamin–Ono-type equations. This point is clarified in the next paragraph.

As in section 2.2, and using the change of variables (2.8), the system (3.10) becomes (2.10), where

$$\sigma(k) = \left[\frac{(1 - ak^2)(1 - ck^2)}{(1 + dk^2)}\right]^{1/2}.$$

Its linearization about the rest state is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\hat{v}\\ \hat{w}\end{pmatrix} + \mathrm{i}k\begin{pmatrix}\sigma(k) & 0\\ 0 & -\sigma(k)\end{pmatrix}\begin{pmatrix}\hat{v}\\ \hat{w}\end{pmatrix} = 0.$$

To see that this latter system is a perturbation of a diagonalized Benjamin–Ono operator, write

$$\sigma(k) = \left[\frac{(1-ak^2)(1-ck^2)}{(1+dk^2)}\right]^{1/2} = \left(\frac{ac}{d}\right)^{1/2}|k| + r(k),$$

where

$$r(k) = \frac{d-(ac+ad+dc)k^2}{d[(1-ak^2)(1-ck^2)(1+dk^2)]^{1/2} + (acd)^{1/2}|k|(1+dk^2)}.$$

The symbol of the dispersion operator $\mathcal{B}$ is seen to be

$$\mathrm{i}\begin{pmatrix}\left(\frac{ac}{d}\right)^{1/2} k|k| + kr(k) & 0\\ 0 & -\left(\frac{ac}{d}\right)^{1/2} k|k| - kr(k)\end{pmatrix}.$$

To understand this operator, which indeed represents a perturbation of a pair of uncoupled Benjamin–Ono equations, attention is given to the residual operator R, whose symbol is $\mathrm{i}kr(k)$.

Lemma 3.7. *Let R be the pseudo-differential operator with symbol $\mathrm{i}kr(k)$. The operator R is bounded in L_p for $1 < p < \infty$. But it is not bounded in L_1 or L_∞.*

Proof. Note that

$$kr(k) = -\frac{ac+ad+dc}{2d(acd)^{1/2}}\operatorname{sgn}(k) + s(k),$$

where $s(k)$ is a function that is in $C^1(\mathbb{R}\backslash 0)$ and is a multiplier in L_p, for $1 < p < \infty$, by Mikhlin's theorem (see, e.g. [44]). Moreover $s \in H^1(\mathbb{R})$ and decays like k^{-2} at infinity, and so it defines an L_1- and L_∞-multiplier. The result then follows from the well-known fact that the Hilbert transform is bounded in L_p if and only if $1 < p < \infty$.

The previous considerations establish that the dispersive operator $\mathcal{B}$ has the structure

$$\mathcal{B} = \begin{pmatrix}\left(\frac{ac}{d}\right)^{1/2}\mathcal{J}\partial_x^2 - \frac{ac+ad+bc}{2d(acd)^{1/2}}\mathcal{J} + \mathcal{K} & 0\\ 0 & -\left(\frac{ac}{d}\right)^{1/2}\mathcal{J}\partial_x^2 + \frac{ac+ad+bc}{2d(acd)^{1/2}}\mathcal{J} - \mathcal{K}\end{pmatrix}, \tag{3.13}$$

where $\mathcal{J}$ is the Hilbert transform and $\mathcal{K}$ is a skew-adjoint operator in L_2 that is bounded on all the L_p-spaces for $1 \leqslant p \leqslant \infty$ and is of order -2, which is to say $\mathcal{K} \in \mathcal{L}(H^s(\mathbb{R}), H^{s+2}(\mathbb{R}))$ for $s \geqslant 0$.

Remark 3.8. In view of the results in [39] and [34] for the usual Benjamin–Ono equation, it seems very unlikely that the flow map in theorem 3.5 is more regular than continuous. In fact, it was established in [39] that the flow map for the BO equation cannot be C^2 in any Sobolev space. Very recently, Koch and Tzvetkov [34] have proved that it cannot be even locally uniformly continuous in any Sobolev class.

However, Koch and Tzvetkov did show using means other than the standard ones that the BO equation is locally well posed in $H^s(\mathbb{R})$ for $s > 5/4$. This has been improved by Kenig and Koenig [31] to $s > 9/8$ and more recently by Tao [45] to $H^1(\mathbb{R})$.

3.3. The Boussinesq system when $\mathcal{H}$ has order 0

In this subsection, consideration is given to systems in (1.1) with

$$c = d = 0, \qquad a < 0, \qquad b > 0$$

and those with

$$a = b = 0, \qquad d > 0, \qquad c < 0,$$

respectively.

These two cases can both be studied using the techniques presented below. In consequence, the proof is worked out only for the first case. Thus, attention is given to the Boussinesq system

$$\eta_t + u_x + (u\eta)_x + au_{xxx} - b\eta_{xxt} = 0, \qquad u_t + \eta_x + uu_x = 0. \tag{3.14}$$

First, consider the case $a = -1$ and $b = 1$. Then (3.14) may be given the form

$$\eta_t + u_x + (I - \partial_x^2)^{-1}(u\eta)_x = 0, \qquad u_t + \eta_x + uu_x = 0. \tag{3.15}$$

Assuming there is in hand a smooth solution, multiply the first equation in (3.15) by η and the second by u and integrate over $\mathbb{R}$. Integration by parts yield

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta^2 + u^2)\,\mathrm{d}x = -\int_{-\infty}^{\infty}\eta(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x.$$

Using (2.3), it transpires that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta^2 + u^2)\,\mathrm{d}x \leqslant C\|\eta\|_0^2\|u\|_0. \tag{3.16}$$

Next, multiply the two equations in (3.15) by $-\eta_{xx}$ and $-u_{xx}$, respectively, and integrate as before to get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta_x^2 + u_x^2)\,\mathrm{d}x = \int_{-\infty}^{\infty}\eta_{xx}(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x + \int_{-\infty}^{\infty}uu_xu_{xx}\,\mathrm{d}x.$$

Note first that

$$\left|\int_{-\infty}^{\infty}uu_xu_{xx}\,\mathrm{d}x\right| \leqslant C\|u_x\|_0^{5/2}\|u_{xx}\|_0^{1/2}. \tag{3.17}$$

Because

$$\int_{-\infty}^{\infty}\eta_{xx}(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x = -\int_{-\infty}^{\infty}\eta_x(I - \partial_x^2)^{-1}(u\eta)_{xx}\,\mathrm{d}x,$$

it follows that

$$\left|\int_{-\infty}^{\infty}\eta_{xx}(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x\right| \leqslant \|\eta_x\|_0\|u\eta\|_0 \leqslant C\|u\|_0^{1/2}\|u_x\|_0^{1/2}\|\eta_x\|_0\|\eta\|_0. \tag{3.18}$$

Finally, multiply the equations in (3.15) by η_{xxxx} and u_{xxxx}, respectively, integrate over $\mathbb{R}$ and integrate by parts to get the relations

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(u_{xx}^2 + \eta_{xx}^2)\,\mathrm{d}x = \int_{-\infty}^{\infty}-\eta_{xxxx}(I - \partial_x^2)^{-1}(u\eta)_x - u_{xxxx}uu_x\,\mathrm{d}x.$$

Since

$$\int_{-\infty}^{\infty}\eta_{xxxx}(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x = \int_{-\infty}^{\infty}\eta_{xx}(I - \partial_x^2)^{-1}(u\eta)_{xxx}\,\mathrm{d}x,$$

there obtains

$$\left|\int_{-\infty}^{\infty}\eta_{xxxx}(I - \partial_x^2)^{-1}(u\eta)_x\,\mathrm{d}x\right| \leqslant \|\eta_{xx}\|_0\|(u\eta)_x\|_0 \leqslant \|\eta_{xx}\|_0(|u|_\infty\|\eta_x\|_0 + \|u_x\|_0|\eta|_\infty). \tag{3.19}$$

Similarly, it is adduced that

$$\left|\int_{-\infty}^{\infty} u_{xxxx}(uu_x)\,\mathrm{d}x\right| = \left|\int_{-\infty}^{\infty} u_{xx}(uu_{xxx} + 3u_x u_{xx})\,\mathrm{d}x\right|$$
$$= \left|\frac{5}{2}\int_{-\infty}^{\infty} u_x u_{xx}^2\,\mathrm{d}x\right| \leqslant C\|u_x\|_0^{1/2}\|u_{xx}\|_0^{5/2}.$$

Combining this with (3.16)–(3.19), there is inferred an H^2-bound that is valid at least on some positive time interval.

In the general case, that is (3.14) with $a < 0$, $b > 0$ and $a \neq -b$, observe that

$$\frac{1-ak^2}{1+bk^2} = -\frac{a}{b} + \frac{a+b}{b(1+bk^2)},$$

so that (3.14) is equivalent to

$$\eta_t - \frac{a}{b}u_x + \frac{a+b}{b}(I - b\partial_x^2)^{-1}u_x + (I - b\partial_x^2)^{-1}(u\eta)_x = 0, \qquad u_t + \eta_x + uu_x = 0.$$

The preceding analysis then settles the issue. The conclusions emanating from this discussion are summarized in the following result.

Theorem 3.9. *For any $(\varphi, \psi) \in H^2(\mathbb{R}) \times H^2(\mathbb{R})$, there exist $T > 0$ and an $(\eta, u) \in C(0, T; H^2(\mathbb{R}))^2$ that is the unique solution of (3.14) with initial value (φ, ψ). Additionally, $(\eta_t, u_t) \in C(0, T; H^1(\mathbb{R}))^2$. The correspondence $(\varphi, \psi) \to (\eta, u)$ is continuous from the space $H^2(\mathbb{R}) \times H^2(\mathbb{R})$ to $C(0, T; H^2(\mathbb{R}))^2$.*

Remark 3.10. The same result holds in $H^s(\mathbb{R}) \times H^s(\mathbb{R})$ for any $s \geqslant 2$.

3.4. The Boussinesq system when $\mathcal{H}$ has order −1

This corresponds to the cases where

$$a = c \geqslant 0, \qquad d = 0, \qquad b > 0$$

and so the system has the form

$$\eta_t + u_x + (u\eta)_x + au_{xxx} - b\eta_{xxt} = 0, \qquad u_t + \eta_x + uu_x + a\eta_{xxx} = 0. \tag{3.20}$$

This system may be treated in the same fashion as (3.10), which was analysed in section 3.2 upon noting that the term u_{xxx} or η_{xxx} played no role in the *a priori* estimates.

Theorem 3.11. *Let $(\varphi, \psi) \in H^{s+1}(\mathbb{R}) \times H^s(\mathbb{R})$ where $s \geqslant 1$. Then there exist $T > 0$ and $(\eta, u) \in C(0, T; H^{s+1}(\mathbb{R})) \times C(0, T; H^s(\mathbb{R}))$ that is the unique solution of (3.20) with initial value (φ, ψ). Additionally, $(\eta_t, u_t) \in C(0, T; H^{s-1}(\mathbb{R})) \times C(0, T; H^{s-2}(\mathbb{R}))$. The mapping from initial data to solutions is continuous.*

The overall outcome of the foregoing analysis is what was mentioned in the introduction and stated as a working hypothesis in part I [11]. For a system of the form displayed in (1.1), if the coefficients satisfy one of the conditions (C1), (C2) or (C3), then according to part I, they are linearly well posed in suitable Sobolev classes. This present analysis shows that the nonlinear problems associated with cases corresponding to (C1) and (C2) are all at least locally well posed.

Theorem 3.12. *The systems in (1.1) with a, b, c, d satisfying (1.2) and (C1) or (C2) are all locally well posed in suitable Sobolev classes.*

Remark 3.13. If one considers the dispersionless system

$$\eta_t + u_x + (u\eta)_x = 0, \qquad u_t + \eta_x + uu_x = 0$$

(also known as the Saint Venant system or shallow water theory), one has, in the hyperbolic region $1 + \eta > 0$, local well-posedness in $H^s(\mathbb{R})^2$ for $s > \frac{3}{2}$ (see theorem 2.1 in Majda [37]). This result is better than the one presented here as far as the level of regularity assumed of the initial data is concerned. In fact, the same result is true for the *abcd*-system with dispersion, but we have avoided this development. It is worth noting that when dispersion is present, two salutary effects arise. First, the proof of well-posedness is easier, and, second, the restriction on η is not necessary.

4. Hamiltonian structure and energy estimates

4.1. Hamiltonian structure

Our study of Hamiltonian structure and the resulting global theory commences with a few elementary observations.

Because dissipation is ignored in the derivation of (1.1) and the overlying Euler equations are Hamiltonian, it is expected that some of the systems in (1.1) will likewise possess a Hamiltonian form. One finds readily for the cases wherein $b = d$ that the functional

$$H(\eta, u) = \frac{1}{2}\int_{-\infty}^{\infty} (c\eta_x^2 + au_x^2 - \eta^2 - u^2 - u^2\eta)\,\mathrm{d}x$$

serves as a Hamiltonian. Indeed, if one defines the operator J to be

$$J = \begin{pmatrix} 0 & (I - b\partial_x^2)^{-1}\partial_x \\ (I - d\partial_x^2)^{-1}\partial_x & 0 \end{pmatrix}, \tag{4.1}$$

then the system (1.1) is identical to the system

$$\partial_t \begin{pmatrix} \eta \\ u \end{pmatrix} = J \,\mathrm{grad}\, H \begin{pmatrix} \eta \\ u \end{pmatrix}, \tag{4.2}$$

where grad H stands for the gradient or Euler derivative, computed with respect to the $L_2 \times L_2$-inner product, of the functional H. When $b = d$, the operator J is skew-adjoint, and then (4.2) comprises a Hamiltonian system. In this case, it follows at once that the Hamiltonian is a constant of the motion, which is to say that

$$H(\eta(\cdot, t), u(\cdot, t)) = H(\varphi, \psi),$$

provided the solution pair (η, u) is sufficiently smooth and vanishes appropriately at infinity.

Remark 4.1. It is worth noticing that for any values of b and d,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} (c\eta_x^2 + au_x^2 - \eta^2 - u^2 - u^2\eta)\,\mathrm{d}x = 2(b - d)\int_{-\infty}^{+\infty} \eta_t u_{xt}\,\mathrm{d}x.$$

This follows immediately from (4.1) upon writing

$$J = S + (d - b)\partial_x J^2 \mathcal{G},$$

where

$$S = \begin{pmatrix} 0 & (I - b\partial_x^2)^{-1}\partial_x \\ (I - b\partial_x^2)^{-1}\partial_x & 0 \end{pmatrix}$$

is skew-symmetric and

$$\mathcal{G} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Thus, the above Hamiltonian structure does not persist to cases where $b \neq d$. Of course, there could be an entirely different Hamiltonian structure in this case, though we conjecture this prospect not to be the case.

Another obvious invariant of the system (1.1) occurring when $b = d$ is the impulse functional I given by

$$I(\eta, u) = \int_{-\infty}^{\infty} \eta u + b\eta_x u_x \, \mathrm{d}x$$

(see [7] for a wide-ranging account of impulse and its implications for models of physical systems). The functionals H and I, together with the invariants

$$\int_{-\infty}^{\infty} u(x,t)\,\mathrm{d}x \qquad \text{and} \qquad \int_{-\infty}^{\infty} \eta(x,t)\,\mathrm{d}x,$$

the second of which expresses preservation of mass, appear to be the only conservation laws that are spatial integrals of polynomials of the dependent variables and their derivatives for the system (1.1). As none of these is composed only of positive terms, they do not on their own provide the *a priori* information one needs to conclude the global existence of solutions to the initial-value problem. However, as is demonstrated next, a time-dependent relationship can be coupled with the invariance of the Hamiltonian to give suitable information leading to a global existence theory.

4.2. Global solutions

It will be assumed throughout this section that

$$b = d > 0, \qquad a \leqslant 0, \qquad c < 0. \tag{4.3}$$

Attention is first given to the non-degenerate cases

$$b = d > 0, \qquad a < 0, \qquad c < 0. \tag{4.4}$$

Theorem 4.2. *Assume that (4.4) holds. Let $s \geqslant 1$ and suppose $(\varphi, \psi) \in H^s(\mathbb{R}) \times H^s(\mathbb{R})$ is such that*

$$\inf_{x\in\mathbb{R}}\{1 + \varphi(x)\} > 0 \qquad \text{and} \qquad |H(\varphi,\psi)| < |c|^{1/2}. \tag{4.5}$$

Then the solution (η, u) of the system (1.1) is global, which means it can be extended to $C(\mathbb{R}_+; H^s(\mathbb{R})) \times C(\mathbb{R}_+; H^s(\mathbb{R}))$. Moreover, the H^1-norm of both η and u is uniformly bounded in t.

Proof. The argument is an adaptation of one made by Bona and Smith [18], and consequently, it is only sketched. Because of (4.5) and the fact that the solution evolves continuously in time with values in $H^1(\mathbb{R})$, and so continuously in time with values in $C_b(\mathbb{R})$, there is a $t_0 > 0$ such that the local solution found in theorem 2.5 satisfies $1 + \eta(x,t) > 0$ for all $x \in \mathbb{R}$ and $0 \leqslant t \leqslant t_0$. For $t \in [0, t_0]$ and $x \in \mathbb{R}$, one has

$$\begin{aligned}\eta^2(x,t) &\leqslant \int_{-\infty}^{+\infty} |\eta\eta_x|\,\mathrm{d}x = \frac{1}{\sqrt{|c|}}\int_{-\infty}^{+\infty} \sqrt{|c|}|\eta||\eta_x|\,\mathrm{d}x \leqslant \frac{1}{2\sqrt{|c|}}\int_{-\infty}^{+\infty}(\eta^2 + |c|\eta_x^2)\,\mathrm{d}x\\ &\leqslant \frac{1}{\sqrt{|c|}}|H(\eta,u)| = \frac{1}{\sqrt{|c|}}|H(\varphi,\psi)| \equiv \alpha^2.\end{aligned}$$

Assuming (4.5), it follows that $\alpha^2 < 1$, whence

$$\sup_{x\in\mathbb{R}} |\eta(x,t)| \leqslant \alpha < 1, \qquad \text{for } 0 \leqslant t \leqslant t_0. \tag{4.6}$$

Thus it must be the case that $1 + \eta(x,t) \geqslant 1 - \alpha > 0$, for all $x \in \mathbb{R}$. Moreover, as long as the solution continues to exist in $H^1(\mathbb{R})$, this positive lower bound on $1 + \eta(x,t)$ continues to hold, independent of $t \geqslant 0$, by simply reapplying the above argument.

On the other hand, because of (4.6), it transpires that

$$\|u\|_1^2 + \|\eta\|_1^2 \leqslant 2\beta |H(\eta, u)| = 2\beta |H(\varphi, \psi)|,$$

where $\beta = \max\{(1-\alpha)^{-1}, |c|^{-1}, |a|^{-1}\}$. This in turn provides a uniform bound in $H^1(\mathbb{R}) \times H^1(\mathbb{R})$, and one simply iterates the contraction-mapping argument applied earlier to this system to obtain global well-posedness in $H^1(\mathbb{R})$.

Once the $H^1(\mathbb{R})$-norm of η and u are known to be uniformly bounded, it follows readily that for any $s \geqslant 1$, the $H^s(\mathbb{R})$-norm must remain bounded on bounded time intervals provided φ, ψ lie in $H^s(\mathbb{R})$. To see this, write (1.1) in the form

$$\eta_t + (I - b\partial_x^2)^{-1}\partial_x(u + u\eta + au_{xx}) = 0, \qquad u_t + (I - b\partial_x^2)^{-1}\partial_x\left(\eta + \frac{u^2}{2} + c\eta_{xx}\right) = 0 \tag{4.7}$$

and apply the operator $J^s = (I - \partial_x^2)^{s/2}$ to both equations in (4.7). Then multiply the first equation by $-cJ^s\eta$ and the second equation by $-aJ^s u$. Upon integrating over $\mathbb{R}$, integrating by parts and summing the results, there appears

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(|c||J^s\eta|^2 + |a||J^s u|^2)\,\mathrm{d}x \\ \leqslant \left|\int_{-\infty}^{\infty}((I - b\partial_x^2)^{-1}(|a|J^s u_x)J^s\eta + (I - b\partial_x^2)^{-1}(|c|J^s\eta_x)J^s u)\,\mathrm{d}x\right| \\ +|c|\int_{-\infty}^{\infty}|J^s(u\eta)(I - b\partial_x^2)^{-1}\eta_x| + \frac{|a|}{2}|J^s(u^2)(I - b\partial_x^2)^{-1}u_x|\,\mathrm{d}x.\end{aligned} \tag{4.8}$$

For any $s \geqslant 0$, there is a constant C depending only on s such that

$$|J^s(fg)|_2 \leqslant C(|f|_\infty \|J^s g\|_0 + |g|_\infty \|J^s f\|_0) \tag{4.9}$$

for $f, g \in H^s \cap L_\infty$ (see, e.g. [26]). Taking into account the uniform H^1-bound on η and u and the embedding of $H^1(\mathbb{R})$ into $L_\infty(\mathbb{R})$, it is deduced from (4.8) and the Gronwall lemma that $\|\eta\|_s$ and $\|u\|_s$ are bounded on any time interval $[0, T]$ by a constant $C = C(T, \|\varphi\|_s, \|\psi\|_s)$. This completes the proof of theorem 4.2.

Attention is now given to the degenerate version,

$$b = d > 0, \qquad a = 0, \qquad c < 0, \tag{4.10}$$

of the case considered above. The system (1.1) in this case resembles a Bona–Smith system, at least in its dispersive structure. The following result obtains in this case.

Theorem 4.3. *Assume that (4.10) holds. For* $s \geqslant 1$*, let* $(\varphi, \psi) \in H^{s+1}(\mathbb{R}) \times H^s(\mathbb{R})$ *be such that*

$$\inf_{x\in\mathbb{R}}\{1 + \varphi(x)\} > 0 \qquad \text{and} \qquad |H(\varphi, \psi)| < |c|^{1/2}. \tag{4.11}$$

Then the local solution (η, u) *of (1.1) may be extended to a global solution in* $C(\mathbb{R}_+; H^{s+1}(\mathbb{R})) \times C(\mathbb{R}_+; H^s(\mathbb{R}))$*. In consequence, the problem (1.1) is globally well posed in* $H^{s+1}(\mathbb{R}) \times H^s(\mathbb{R})$*, and, moreover, the* $H^1(\mathbb{R}) \times L_2(\mathbb{R})$*-norm of* (η, u) *is uniformly bounded in* t*.*

Proof. Following the line of the first argument in the proof of theorem 4.2 for the non-degenerate case (4.4), one finds only that for all $t \geqslant 0$ for which the solution exists,

$$\|\eta\|_1^2 + \|u\|_0^2 \leqslant 2\beta|H(\eta, u)| = 2\beta|H(\varphi, \psi)|, \tag{4.12}$$

where $\beta = \max\{(1-\alpha)^{-1}, |c|^{-1}\}$, and that

$$1 + \eta(x,t) \geqslant 1 - \alpha$$

for such t and all $x \in \mathbb{R}$, with

$$\alpha^2 = \frac{1}{\sqrt{|c|}}|H(\varphi, \psi)|$$

as before.

The further bounds appearing in the proof of theorem 4.2 are obtained differently in this case. First, bounds in $H^2(\mathbb{R}) \times H^1(\mathbb{R})$ are obtained. To this end, multiply the first equation in (4.7) by η_{xxxx} and the second by u_{xx}. After summing these two equations, integrating by parts and performing standard estimates, we arrive at the following differential inequality:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta_{xx}^2 + u_x^2)\,\mathrm{d}x \leqslant C_1|\eta_{xx}|_2|u_x|_2 + C_2|u_x|_2 + C_3|u_x|_2^{3/2}$$
$$+C_4|\eta_{xx}|_2(|u_x|_2|\eta|_\infty + |u|_\infty|\eta_x|_2), \tag{4.13}$$

where C_1 and C_4 depend only on b and c, while C_2 and C_3 depend also on the previously obtained bounds on $\|\eta\|_1$ and $|u|_2$. Using the above-mentioned bounds again to further estimate the last term on the right-hand side of (4.13) gives a differential inequality to which a Gronwall-type lemma applies and yields the property that (η, u) is bounded in $C(0, T; H^2(\mathbb{R})) \times C(0, T; H^1(\mathbb{R}))$ for any time interval $[0, T]$ over which the solution exists. As this equation has a local well-posedness theory in $H^{s+1}(\mathbb{R}) \times H^s(\mathbb{R})$ for any $s \geqslant 1$, we thus conclude global existence at least in $H^2(\mathbb{R}) \times H^1(\mathbb{R})$.

Because of the local well-posedness theory, the theorem will be concluded as soon as it is ascertained that for any $r \geqslant 1$, $\|\eta(\cdot,t)\|_{r+1} + \|u(\cdot,t)\|_r$ are *a priori* bounded on any finite range $0 \leqslant t \leqslant T$. Such bounds are obtained by an energy-type argument as follows. Let $r \geqslant 1$ and suppose $(\varphi, \psi) \in H^{r+1}(\mathbb{R}) \times H^r(\mathbb{R})$. If (η, u) is the solution of (1.1)–(4.10) emanating from initial data (φ, ψ), then from our local well-posedness theory together with the $H^2 \times H^1$ theory established above, it is adduced that (η, u) exists globally in time and that $\|\eta(\cdot,t)\|_2 + \|u(\cdot,t)\|_1$ is bounded on bounded time intervals. Moreover, for small values of t, $\|\eta(\cdot,t)\|_{r+1}$ and $\|u(\cdot,t)\|_r$ are bounded. With r as above, apply J^{r+1} to the first equation in (4.7), multiply by $J^{r+1}\eta$ and integrate over $\mathbb{R}$. Apply J^r to the second equation in (4.7), multiply by $J^r u$ and integrate. (Recall that because of the continuous dependence, which is part of the import of well-posedness, we may perform these operations and those to follow with impunity as issues of regularity and decay at infinity may be handled by regularizing the data, working with the much smoother solutions and then passing to the limit in the resulting inequalities that do not depend upon the extra regularity.) Adding the formulae derived thereby yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}\{(J^{r+1}\eta)^2 + (J^r u)^2\}\,\mathrm{d}x = -\int_{-\infty}^{\infty} J^{r+1}\eta(I - b\partial_x^2)^{-1}\partial_x J^{r+1}(u + u\eta)\,\mathrm{d}x$$
$$-\int_{-\infty}^{\infty} J^r u(I - b\partial_x^2)^{-1}\partial_x J^r\left(\eta + \frac{1}{2}u^2 + c\eta_{xx}\right)\mathrm{d}x. \tag{4.14}$$

The Cauchy–Schwarz inequality suffices to estimate all the quadratic terms on the right-hand side of (4.14). For the cubic terms, (4.9) comes to our rescue; there obtains

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}\{(J^{r+1}\eta)^2 + (J^r u)^2\}\,\mathrm{d}x \leqslant D_1|J^{r+1}\eta|_2|J^r u|_2 + D_2|J^{r-1}u|_2|J^r\eta|_2$$
$$+D_3|J^{r+1}\eta|_2(|u|_\infty|J^r\eta|_2 + |\eta|_\infty|J^r u|_2) + D_4|J^r u|_2|u|_\infty|J^{r-1}u|_2. \tag{4.15}$$

The constant D_1 depends on b and c, and D_2 depends on b, while D_3 and D_4 depend on b and r. Because $|u|_\infty \leqslant \|u\|_1$ is bounded on bounded time intervals and $|\eta|_\infty \leqslant \|\eta\|_1$ is bounded independent of t,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}\{(J^{r+1}\eta)^2 + (J^r u)^2\}\,\mathrm{d}x \leqslant D_5|J^{r+1}\eta|_2|J^r u|_2,$$

where D_5 depends on b, c, r and T via the bound on $\|u\|_{C(0,T;H^1(\mathbb{R}))}$. From this, the desired estimate follows on account of Gronwall's lemma. The proof of the theorem is thereby concluded.

Remark 4.4. One might wonder whether or not the higher Sobolev norms of the global solution obtained in theorem 4.2 are uniformly bounded in t. This question is open (as is the corresponding one for the solution of the BBM equation), but one can easily prove the following weaker result; for $t \geqslant 0$,

$$\|\partial_x^k\eta(\cdot,t)\|_2 + \|\partial_x^k u(\cdot,t)\|_2 \leqslant C(1+t^{k-1}), \qquad k = 2, 3, \ldots. \tag{4.16}$$

Consider first the case $k = 2$. Write (1.1) as

$$\eta_t + (I - b\partial_x^2)^{-1}(u_x + (u\eta)_x + au_{xxx}) = 0, \qquad u_t + (I - b\partial_x^2)^{-1}(\eta_x + uu_x + c\eta_{xxx}) = 0. \tag{4.17}$$

Differentiate (4.17) twice with respect to x and take the L_2-scalar product with $(-c\eta_{xx}, -au_{xx})$ to obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(-a\eta_{xx}^2 - cu_{xx}^2)\,\mathrm{d}x = c\int_{-\infty}^{\infty}(I - b\partial_x^2)^{-1}(\partial_x(u_x + (u\eta)_x))\eta_{xx}\,\mathrm{d}x$$
$$-a\int_{-\infty}^{\infty}(I - b\partial_x^2)^{-1}(\partial_x(\eta_x + uu_x))u_{xx}\,\mathrm{d}x. \tag{4.18}$$

Because of the uniform bounds on $\|u(\cdot,t)\|_1$ and $\|\eta(\cdot,t)\|_1$, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(\eta_{xx}^2 + u_{xx}^2)\,\mathrm{d}x \leqslant C(\|\eta_{xx}(\cdot,t)\|_0 + \|u_{xx}(\cdot,t)\|_0)$$

which yields (4.16) for $k = 2$.

The general case follows easily by induction and using

$$(I - b\partial_x^2)^{-1}\partial_x^k = \frac{1}{b}\partial_x^{k-2} + \mathcal{R}_k,$$

where $\mathcal{R}_k$ is a pseudo-differential operator of order $k - 4$.

Similar results were obtained by Staffilani [43] and Bourgain [22] for the generalized KdV-equation and other nonlinear dispersive equations.

Similar polynomial bounds do not seem to be available in the context of theorem 4.3.

5. Local well-posedness of a class of higher-order Boussinesq systems

Higher-order systems in the form

$$\eta_t - b\eta_{xxt} + b_2\eta_{xxxxt} = -u_x - (\eta u)_x - au_{xxx} + b(\eta u)_{xxx} - (a + b - \tfrac{1}{3})(\eta u_{xx})_x - a_2 u_{xxxxx},$$
$$u_t - du_{xxt} + d_2 u_{xxxxt} = -\eta_x - c\eta_{xxx} - uu_x - c(uu_x)_{xx} - (\eta\eta_{xx})_x + (c + d - 1)u_x u_{xx}$$
$$+(c+d)uu_{xxx} - c_2\eta_{xxxxx} \tag{5.1}$$

were also derived in part I. These systems are formally second-order approximations of the full, two-dimensional Euler equations. The constants $a, b, c, d, a_2, b_2, c_2, d_2$ satisfy (1.2) and

$$a_2 - b_2 = -\tfrac{1}{2}(\theta^2 - \tfrac{1}{3})b + \tfrac{5}{24}(\theta^2 - \tfrac{1}{5})^2, \qquad c_2 - d_2 = \tfrac{1}{2}(1-\theta^2)c + \tfrac{5}{24}(1-\theta^2)(\theta^2 - \tfrac{1}{5}), \tag{5.2}$$

where, as before, $\theta \in [0, 1]$.

In this section, attention is restricted to a particular subclass of higher-order regularized long-wave systems that have

$$\begin{array}{llll} b \geqslant 0, & b_2 > 0, & a < 0, & a_2 = 0, \\ d \geqslant 0, & d_2 > 0, & c < 0, & c_2 = 0. \end{array} \tag{5.3}$$

Our goal is to study the well-posedness of these systems. The systems satisfying (5.3) are obtained by choosing

$$\begin{array}{lll} \tfrac{1}{3} < \theta^2 < 1, & a_2 = 0, & c_2 = 0, \\ c < -\dfrac{5}{12}\left(\theta^2 - \dfrac{1}{5}\right), & b \geqslant \dfrac{1}{2}\left(\theta^2 - \dfrac{1}{3}\right), & b > \dfrac{5}{12}\dfrac{(\theta^2 - 1/5)^2}{(\theta^2 - 1/3)} \end{array} \tag{5.4}$$

and then evaluating a, d, b_2, d_2 from (1.2) and (5.2). These were called higher-order regularized long-wave systems or higher-order BBM-systems in part I.

With arguments similar to those put forward in section 2.1 for pure BBM-Boussinesq systems, one may infer easily the local well-posedness for these higher-order systems. To effect such a theory, first convert (5.1) with the values (5.3)–(5.4) of the coefficients into a coupled system of integral equations. For the first equation, invert the operator $(1 - b\partial_x^2 + b_2\partial_x^4)$ subject to zero boundary conditions at $\pm\infty$, perform formal integrations by parts and integration over $[0, t]$ to obtain

$$\eta(x,t) = \varphi(x) + \int_0^t \{K_1 * [u + \eta u - \alpha_1 \eta_x u_x] + M_1 * [-a u_x + b(\eta u)_x - \alpha_1 \eta u_x]\}\, \mathrm{d}\tau, \tag{5.5}$$

where $*$ denotes convolution over $\mathbb{R}$, $\alpha_1 = a + b - \frac{1}{3}$,

$$\hat{K}_1(k) = \frac{\mathrm{i}k}{1 + bk^2 + b_2 k^4} \qquad \text{and} \qquad \hat{M}_1(k) = \frac{-k^2}{1 + bk^2 + b_2k^4}.$$

Similarly, for the second equation in (5.1), one obtains by the same set of operations the formula

$$u(x,t) = \psi(x) + \int_0^t \left\{ K_2 * \left[\eta + \frac{1}{2}u^2 - (\eta_x)^2 + \left(c + d + \frac{1}{2}\right)(u_x)^2\right] \right.$$
$$\left. + M_2 * [-c\eta_x - \eta\eta_x + d u u_x] \right\} \mathrm{d}\tau, \tag{5.6}$$

where

$$\hat{K}_2 = \frac{\mathrm{i}k}{1 + dk^2 + d_2k^4} \qquad \text{and} \qquad \hat{M}_2 = \frac{-k^2}{1 + dk^2 + d_2k^4}.$$

An application of the contraction-mapping principle to a suitable ball around the origin in $C(0, T; H^s(\mathbb{R}))^2$ for appropriate choices of T yields the following result.

Theorem 5.1. *For any $(\varphi, \psi) \in H^s(\mathbb{R})^2$, $s \geqslant 1$, there exist $T > 0$ and a unique solution (η, u) of (5.5) and (5.6), both of whose components η and u lie in $C(0, T; H^s(\mathbb{R}))$. The mapping that associates to initial data the corresponding solution is Lipschitz continuous. The temporal derivatives η_t and u_t both lie in $C(0, T; H^{s+1}(\mathbb{R}))$.*

The first step is to understand the way in which convolution with K_i and M_i $(i = 1, 2)$ maps various function classes.

Lemma 5.2. *Let $s \geqslant 0$ be given. There are constants C_j, $j = 1, \ldots, 6$, depending only upon s such that, for $i = 1, 2$,*

(i) if $f \in H^s(\mathbb{R})$, then $\|K_i * f\|_{s+3} \leqslant C_1\|f\|_s$,
(ii) if $f \in H^s(\mathbb{R})$, then $\|M_i * f\|_{s+2} \leqslant C_2\|f\|_s$,
(iii) if $f, g \in H^s(\mathbb{R})$, then $\|K_i * (fg)\|_{s+2} \leqslant C_3\|f\|_s\|g\|_s$,
(iv) while if $f, g \in H^s(\mathbb{R})$, then $\|M_i * (fg)\|_{s+1} \leqslant C_4\|f\|_s\|g\|_s$.

If, on the other hand, $s \geqslant 1$, then

(v) for $f, g \in H^s(\mathbb{R})$, $\|K_i * (f'g')\|_{s+1} \leqslant C_5\|f\|_s\|g\|_s$,
(vi) and for $f, g \in H^s(\mathbb{R})$, $\|M_i * (fg')\|_s \leqslant C_6\|f\|_s\|g\|_s$.

Proof. These inequalities follow straightforwardly by computing in Fourier transformed variables and making use of the elementary facts that the Fourier transform of a product is a convolution and, conversely, the Fourier transform of a convolution is the product of the associated Fourier transforms. Inequalities (i) and (ii) are sharp, but (iii) and (iv) can be improved in that stronger norms can appear on the left-hand sides without disturbing their validity. Inequalities (v) and (vi) follow from (iii) and (iv), respectively. Inequality (vi) can obviously be improved. The various sharper results that are available are not needed here.

Proof of theorem 5.1. (i) *Uniqueness*. It suffices to prove the uniqueness of the solution when $s = 1$. Suppose that for some $T > 0$, (η_1, u_1) and (η_2, u_2) in $C(0, T; H^1(\mathbb{R}))^2$ are two solution pairs of (5.5) and (5.6) corresponding to the same initial data. Define functions η and u by $(\eta, u) = (\eta_1 - \eta_2, u_1 - u_2)$. Then $\eta(x, 0) = u(x, 0) = 0$ in $H^1(\mathbb{R})$ by assumption and

$$\eta(x, t) = \int_0^t \{K_1 * [u + \eta u_1 + \eta_2 u - \alpha_1 \eta_x (u_1)_x - \alpha_1 (\eta_2)_x u_x]$$
$$+ M_1 * [-a u_x + b(\eta_x u_1 + (\eta_2)_x u + \eta (u_1)_x + \eta_2 u_x) - \alpha_1 \eta (u_1)_x - \alpha_1 \eta_2 u_x]\}\,\mathrm{d}\tau.$$

Using lemma 5.2 systematically, it is straightforward to see that

$$\|\eta\|_1 \leqslant C \int_0^t [\|u\|_0 + \|\eta\|_0 + \|u\|_1 + \|\eta\|_1]\,\mathrm{d}\tau \leqslant C_1 \int_0^t \|u\|_1 + \|\eta\|_1\,\mathrm{d}\tau,$$

where the constant C_1 depends upon $\|u_1\|_{C(0,T;H^1(\mathbb{R}))}$ and $\|\eta_2\|_{C(0,T;H^1(\mathbb{R}))}$. Similarly, one shows that

$$\|u\|_1 \leqslant C_2 \int_0^t \|u\|_1 + \|\eta\|_1\,\mathrm{d}\tau,$$

where C_2 has the same type of dependence on u_1 and η_2 as does C_1. In sum, we have

$$\|u\|_1 + \|\eta\|_1 \leqslant C \int_0^t \|u\|_1 + \|\eta\|_1\,\mathrm{d}\tau$$

for $0 \leqslant t \leqslant T$, where C is bounded at least on $[0, T]$. Gronwall's lemma then implies that

$$\|\eta\|_1 + \|u\|_1 = 0 \qquad \text{for } t \in [0, T].$$

(ii) *Existence*. Fix $s \geqslant 1$. Let X_T^s be as before with the usual norm $\|(\eta, u)\|_{X_T^s} = \|\eta\|_{C(0,T;H^s)} + \|u\|_{C(0,T;H^s)}$. Write the pair of integral equations (5.5) and (5.6) symbolically as $(\eta, u) = \mathcal{A}(\eta, u)$, where $\mathcal{A}$ is a map from X_T^s to itself. We will show $\mathcal{A}$ has a fixed point in X_T^s by suitably choosing T and using the contraction-mapping principle.

Suppose that both (η_1, u_1) and (η_2, u_2) lie in the closed ball B_R of radius R about $\mathbf{0}$ in X^s_T. As in the proof of theorem 2.1, the following inequality obtains:

$$|\mathcal{A}(\eta_1, u_1) - \mathcal{A}(\eta_2, u_2)|$$
$$\leqslant C(1 + \|(\eta_1, u_1)\|_{X^s_T} + \|(\eta_2, u_2)\|_{X^s_T}) \int_0^T (\|\eta_1 - \eta_2\|_s + \|u_1 - u_2\|_s)\,\mathrm{d}\tau$$
$$\leqslant CT(1 + 2R)\|(\eta_1 - \eta_2, u_1 - u_2)\|_{X^s_T} = \Theta\|(\eta_1 - \eta_2, u_1 - u_2)\|_{X^s_T}.$$

It follows from this inequality that for $(\eta, u) \in B_R$,

$$\|\mathcal{A}(\eta, u)\|_{X^s_T} = \|\mathcal{A}(\eta, u) - \mathcal{A}(0, 0) + (\varphi, \psi)\|_{X^s_T} \leqslant \Theta\|(\eta, u)\|_{X^s_T} + \|(\varphi, \psi)\|_{H^s} \leqslant \Theta R + b,$$

where $b = \|(\varphi, \psi)\|_{H^s}$. Thus, if we choose $R = 2b$ and $T = T(b) = 1/(2(1 + 2R)C)$, it is seen that

$$\Theta = \tfrac{1}{2} \qquad \text{and} \qquad \|\mathcal{A}(\eta, u)\|_{X^s_T} \leqslant R.$$

Thus, $\mathcal{A}$ is a contraction on B_R, and the contraction-mapping theorem can be used to establish existence of a solution.

The extra spatial regularity of η_t and u_t follows from lemma 5.2 upon differentiating (5.5) and (5.6) with respect to t.

6. Conclusions

As was the initial goal, we have shown for the class (1.1)–(1.2) of Boussinesq systems that all those that were determined in part I to be linearly well posed are in fact locally nonlinearly well posed in suitable function classes, except the very special cases in (C3). A Hamiltonian subclass (4.3) of the class (1.1)–(1.2) is seen have constituents that are globally well posed in the physically relevant realm of small-amplitude, long-wavelength disturbances. Thus, one has a set of models for the propagation of long-crested waves in the Boussinesq regime (small-amplitude, long waves with Stokes number of order 1) with satisfactory mathematical theories, at least as regards the pure initial-value problems (1.3).

There are several interesting points growing out of this work that are worthy of further study. On the more mathematically motivated side, there is the question, raised already in section 1, of whether or not a particular model has suitably bounded and smooth solutions on the Boussinesq timescale $T_1 = \epsilon^{-3/2}$ and, indeed, on the longer timescale $T_2 = \epsilon^{-5/2}$ for physically relevant initial disturbances. (Recall that $\epsilon = a/h_0$ is the ratio of wave amplitude to water depth.) This leads naturally to the issue of whether the initial-value problem (1.1)–(1.2) under the restrictions (C1) or (C2) are globally well posed in time and to related questions of possible singularity formation in finite time. Numerical simulations not reported here indicate that some of the equations do feature singularity formation in finite time for large initial data, just as happens for KdV-type unidirectional models in the same long-wave regime (see, e.g. [4, 13, 16, 38]). On the other hand, we have found no evidence of singularity formation for physically relevant disturbances. These issues are currently under investigation.

Another, more practical goal would be to develop a theory for boundary-value problems for a range of the systems (1.1). When such models are used to describe real phenomena, non-homogeneous boundary conditions imposed at finite spatial positions often intrude just as they do for unidirectional models (see, e.g. [19, 20, 27, 28] and the references therein). A beginning for this may be found in Bona and Chen [10]. Moreover, numerical simulations are invariably performed on bounded domains, and so a relevant theory is needed in this context.

A range of issues connected with the existence and interaction of solitary waves is also relevant for these system. Preliminary numerical simulations indicate that solitary-wave solutions play the same sort of central role for (1.1) in the evolution of initial disturbances as they do for unidirectional KdV or BBM models (see [1, 14, 15]). Questions of existence are dealt with in Chen [24, 25], but dynamical issues are still under study.

Detailed comparisons of predictions with laboratory and field data are always the final test of a model's relevance. Some evidence in terms of (2.1) is offered in Bona and Chen [10], but further work is needed. The use of these models in a coastal engineering context is also being developed (see, e.g. [9]).

Finally, this theory is limited by the assumption that the wave motion is long crested and so sensibly independent of the independent variable perpendicular to the principle direction of propagation. Models having the same level of approximation, but allowing for significant variation in the other horizontal coordinate direction have appeared in the literature (see, e.g. [36, 40]). A comprehensive appraisal along the line of this study is being mounted [12].

References

[1] Ablowitz M J and Segur H 1981 Solitons and the inverse scattering transform *SIAM Stud. Appl. Math.* vol 4 (Philadelphia, PA: Society for Industrial and Applied Mathematics)

[2] Alazmann A A, Albert J P, Bona J L, Chen M and Wu J 2004 Comparisons between the BBM equation and a Boussinesq system *J. Nonlinear Sci.* submitted

[3] Amick C J 1984 Regularity and uniqueness of solutions to the Boussinesq system of equations *J. Diff. Eqns* **54** 231–47

[4] Angulo J, Bona J L, Linares F and Scialom M 2002 Scaling, stability and singularities for nonlinear, dispersive wave equations: the critical case *Nonlinearity* **15** 759–86

[5] Angulo Pava J 1999 On the Cauchy problem for a Boussinesq-type system *Adv. Diff. Eqns* **4** 457–92

[6] Ben Youssef W and Colin T 2000 Rigorous derivation of Korteweg-de Vries-type systems from a general class of nonlinear hyperbolic systems *Math. Model. Numer. Anal.* **34** 873–911

[7] Benjamin T B 1984 Impulse, flow force and variational principles *IMA J. Appl. Math.* **32** 3–68

[8] Benjamin T B, Bona J L and Mahony J J 1972 Model equations for long waves in nonlinear dispersive systems *Phil. Trans. R. Soc. Lond.* A **272** 47–78

[9] Boczar-Karakiewicz B, Bona J L, Romańczyk W and Thornton E B 2003 Seasonal and interseasonal variability of sand bars at Duck, NC, USA *Eur. J. Mech. B Fluids* submitted

[10] Bona J L and Chen M 1998 A Boussinesq system for two-way propagation of nonlinear dispersive waves *Physica* D **116** 191–224

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

[12] Bona J L, Colin T and Lannes D 2003 Long wave approximations for water waves *Arch. Rat. Mech. Anal.* submitted

[13] Bona J L, Dougalis V A, Karakashian O A and McKinney W R 1995 Conservative, high-order numerical schemes for the generalized Korteweg-de Vries equation *Phil. Trans. R. Soc. Lond.* A **351** 107–64

[14] Bona J L, McKinney W R and Restrepo J M 2000 Stable and unstable solitary-wave solutions of the generalized regularized long-wave equation *J. Nonlinear Sci.* **10** 603–38

[15] Bona J L, Pritchard W G and Scott L R 1980 Solitary-wave interaction *Phys. Fluids* **23** 438–41

[16] Bona J L and Saut J-C 1993 Dispersive blowup of solutions of generalized Korteweg-de Vries equations *J. Diff. Eqns* **103** 3–57

[17] Bona J L and Smith R 1975 The initial-value problem for the Korteweg-de Vries equation *Phil. Trans. R. Soc. Lond.* A **278** 555–601

[18] Bona J L and Smith R 1976 A model for the two-way propagation of water waves in a channel *Math. Proc. Cambridge Phil. Soc.* **79** 167–82

[19] Bona J L, Sun S M and Zhang B-Y 2002 A non-homogeneous boundary-value problem for the Korteweg-de Vries equation in a quarter plane *Trans. Am. Math. Soc.* **354** 427–90 (electronic)

[20] Bona J L, Sun S M and Zhang B-Y 2003 A nonhomogeneous boundary-value problem for the Korteweg-de Vries equation posed on a finite domain *Comm. Partial Diff. Eqns* **28** 1391–436

[21] Bourgain J 1993 Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations *Geom. Funct. Anal.* **3** 107–56, 209–62

[22] Bourgain J 1996 On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE *Int. Math. Res. Not.* 277–304

[23] Boussinesq J V 1872 Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond *J. Math. Pures Appl.* **17** 55–108

[24] Chen M 1998 Exact solutions of various Boussinesq systems *Appl. Math. Lett.* **11** 45–9

[25] Chen M 2000 Solitary-wave and multi-pulsed traveling-wave solutions of Boussinesq systems *Appl. Anal.* **75** 213–40

[26] Coifman R R and Meyer Y 1986 Nonlinear harmonic analysis, operator theory and PDE *Beijing Lectures in Harmonic Analysis (Beijing, 1984)* (Princeton, NJ: Princeton University Press) *Ann. Math. Stud.* **112** 3–45

[27] Colliander J E and Kenig C E 2002 The generalized Korteweg-de Vries equation on the half line *Comm. Partial Diff. Eqns* **27** 2187–266

[28] Faminskii A V 2003 An initial boundary-value problem in a half strip for the KdV equation in fractional-order Sobolev spaces *Preprint*

[29] Fokas A S and Pelloni B 2003 Boundary value problems for Boussinesq type systems *Preprint*

[30] Hammack J L, Henderson D M and Segur H 2003 Deep-water waves with persistent, two-dimensional surface patterns *Preprint*

[31] Kenig C E and Koenig K D 2003 On the local well-posedness of the Benjamin–Ono and modified Benjamin–Ono equations *Preprint*

[32] Kenig C E, Ponce G and Vega L 1991 Well-posedness of the initial value problem for the Kortweg-de Vries equation *J. Am. Math. Soc.* **4** 323–47

[33] Kenig C E, Ponce G and Vega L 1993 Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle *Comm. Pure Appl. Math.* **46** 527–620

[34] Koch H and Tzvetkov N 2003 On the local well-posedness of the Benjamin–Ono equation in $H^s(\mathbb{R})$ *Int. Math. Res. Not.* 1449–64

[35] Lions J-L 1969 *Quelques méthodes de résolution des problèmes aux limites non linéaires* (Paris: Dunod Gauthier-Villars)

[36] Madsen P and Schäffer H A 1998 Higher-order Boussinesq-type equations for surface gravity waves: derivation and analysis *Phil. Trans. R. Soc. Lond.* A **356** 3123–84

[37] Majda A 1984 Compressible fluid flow and systems of conservation laws in several space variables *Applied Mathematical Science* vol 53 (New York: Springer)

[38] Martel Y and Merle F 2002 Blow up in finite time and dynamics of blow up solutions for the L^2-critical generalized KdV equation *J. Am. Math. Soc.* **15** 617–64 (electronic)

[39] Molinet L, Saut J-C and Tzvetkov N 2001 Ill-posedness issues for the Benjamin–Ono and related equations *SIAM J. Math. Anal.* **33** 982–8

[40] Restrepo J M and Bona J L 1995 Three-dimensional model for the formation of longshore sand structures on the continental shelf *Nonlinearity* **8** 781–820

[41] Saut J-C and Tzvetkov N 2000 On a model system for the oblique interaction of internal gravity waves *Math. Model. Numer. Anal.* **34** 501–23 (Special issue for R Temam's 60th birthday)

[42] Schonbek M E 1981 Existence of solutions for the Boussinesq system of equations *J. Diff. Eqns* **42** 325–52

[43] Staffilani G 1997 On the growth of high Sobolev norms of solutions for KdV and Schrödinger equations *Duke Math. J.* **86** 109–42

[44] Stein E M 1970 *Singular integrals and differentiability properties of functions* (Princeton, NJ: Princeton University Press)

[45] Tao T 2003 Global well-posedness of the Benjamin–Ono equation in $H^1(R)$ *Preprint*