




PAPER • OPEN ACCESS

Simple two-layer dispersive models in the Hamiltonian reduction formalism

To cite this article: R Camassa *et al* 2023 *Nonlinearity* **36** 4523

View the [article online](#) for updates and enhancements.

Simple two-layer dispersive models in the Hamiltonian reduction formalism

R Camassa¹ , G Falqui^{2,4,5} , G Ortenzi^{2,4,6,*} ,
M Pedroni^{3,4}  and T T Vu Ho^{1,2,4} 

¹ University of North Carolina at Chapel Hill, Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, Chapel Hill, NC 27599, United States of America

² Department of Mathematics and Applications, University of Milano-Bicocca, Via Roberto Cozzi 55, I-20125 Milano, Italy

³ Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università di Bergamo, Viale Marconi 5, I-24044 Dalmine (BG), Italy

⁴ INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

⁵ SISSA, via Bonomea 265, 34136 Trieste, Italy

⁶ Dipartimento di Matematica “Giuseppe Peano” Università di Torino, Via Carlo Alberto 10, I-10123 Torino (TO), Italy

E-mail: giovanni.ortenzi@unimib.it

Received 28 October 2022; revised 16 June 2023

Accepted for publication 3 July 2023

Published 14 July 2023

Recommended by Dr Karima Khusnutdinova



CrossMark

Abstract

A Hamiltonian reduction approach is defined, studied, and finally used to derive asymptotic models of internal wave propagation in density stratified fluids in two-dimensional domains. Beginning with the general Hamiltonian formalism of Benjamin (1986 *J. Fluid Mech.* **165** 445–74) for an ideal, stably stratified Euler fluid, the corresponding structure is systematically reduced to the setup of two homogeneous fluids under gravity, separated by an interface and confined between two infinite horizontal plates. A long-wave, small-amplitude asymptotics is then used to obtain a simplified model that encapsulates most of the known properties of the dynamics of such systems, such as bidirectional wave propagation and maximal amplitude travelling waves in the form of fronts. Further reductions, and in particular devising an asymptotic extension of Dirac's theory of Hamiltonian constraints, lead to the completely integrable

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

evolution equations previously considered in the literature for limiting forms of the dynamics of stratified fluids. To assess the performance of the asymptotic models, special solutions are studied and compared with those of the parent equations

Keywords: stratified fluids, Hamiltonian PDEs, Hamiltonian reductions, dispersive internal wave models, travelling wave solutions

Mathematics Subject Classification numbers: 37K05, 35Q31, 35Q35, 76B70

(Some figures may appear in colour only in the online journal)

1. Introduction

Density stratification in incompressible fluids is an important aspect of theoretical fluid dynamics, and is an inherent component of a wide variety of phenomena related to geophysical application. Displacement of fluid parcels from their neutral buoyancy position within a density stratified flow can result in internal wave motion, whose governing equations are not in general amenable to analytical methods for their solutions. Simplified models able to isolate key mechanism of the dynamics that can be studied in detail, even in their one space-dimensional limit, are therefore valuable and over the years many have been proposed in the literature. A (very) partial list includes [2, 15, 17, 18, 23, 39, 40] among many others.

The main focus of our work is the study of an ideal stratified fluid from a Hamiltonian viewpoint. The governing equations in the absence of viscosity and diffusivity of the stratifying agent are the Euler equations augmented by density advection, and we consider the simplified case consisting of two homogeneous density layers in two spatial dimensions, in the absence of surface tension and confined by two rigid, horizontal, infinite plates. Hamiltonian aspects of such models with an emphasis on the two-layer case have been considered, notably in [2, 17, 18]. Our approach is an alternative to the one used in [17] for their two-layer case, and similarly combines asymptotic expansions with the Hamiltonian structure of the original Euler equations. However, our starting point is the general density-stratified Hamiltonian of Benjamin [1], and does not make use of the generalization to the two-layer case, as in [17], of Zakharov's Hamiltonian structure [41] for free surface water waves. In our approach, once the Hamiltonian reduction [9] is applied to the two-layer case, we consider different balances between nonlinearity and dispersion, which allows us to retain different asymptotic orders in the ensuing models. In particular, in this paper we focus on a model for interfacial waves propagation between two-homogeneous density fluids for which nonlinearity is stronger than dispersion. This model, which we shall refer to as the *ABC system*, consists of a three-parameter pair of coupled evolution equations that generalize to bidirectional propagation the well-known KdV-mKdV or Gardner equation for unidirectional motion, and it reduces to it (together with its Hamiltonian structure) through a systematic application of Dirac's Hamiltonian theory of constraints. In the weakly nonlinear (WNL) regime, for which a precise nonlinearity and dispersion balance is enforced, the model reduces to the well known integrable cases of Kaup–Boussinesq [5, 26, 27, 37].

This paper is organized as follows. Section 2 is concerned with the details of the derivation of the model equations. Specifically, after a brief review of the fundamental governing equations for ideal, density stratified, incompressible fluids in the section's introduction and in § 2.1, we present the elements of Hamiltonian reduction to two-layer flows in § 2.2 and § 2.3. We then proceed to define and apply our asymptotic assumptions in § 2.4–2.6 to derive the limiting form of the Hamiltonian equations of motion. Section 3 studies the structure of our

main model, while the following section, § 4, considers notable reductions that yield known integrable systems for WNL dynamics. Finally, section 5 considers special solutions that serve to illustrate the models’ main features and drawbacks, as well as propose asymptotic equivalences to remedy the latter; lastly, section 6 discusses future directions of investigation and concludes the paper.

2. Density stratified Euler fluids

We consider a perfect, incompressible and variable density fluid confined between two horizontal infinite plates. Thus, the fluid occupies the two-dimensional (2D) domain $(x, z) \in \mathbb{R} \times (-h_2, h_1)$, with $h_1 + h_2 \equiv h$ the distance between bottom and upper boundary. Such a fluid is governed by the incompressible Euler equations for the velocity field $\mathbf{u} = (u, w)$ and non-constant density $\rho(x, z, t)$, in the presence of gravity $-\mathbf{g}\mathbf{k}$,

$$\frac{D\rho}{Dt} = 0, \quad \nabla \cdot \mathbf{u} = 0, \quad \frac{D(\rho\mathbf{u})}{Dt} + \nabla p + \rho\mathbf{g}\mathbf{k} = 0 \tag{2.1}$$

with boundary conditions

$$\mathbf{u}(x = \pm\infty, z, t) = \mathbf{0}, \quad \text{and } w(x, -h_2, t) = w(x, h_1, t) = 0, \quad x \in \mathbb{R}, \quad z \in (-h_2, h_1), \quad t \in \mathbb{R}^+, \tag{2.2}$$

where $z = -h_2$ and $z = h_1$ are the locations of the bottom and top confining plates, respectively. As usual, $D/Dt = \partial/\partial t + \mathbf{u} \cdot \nabla$ is the material derivative.

2.1. The 2D Benjamin model for heterogeneous fluids in a channel

The above system was given a Hamiltonian structure in [1] with basic, locally measurable variables, i.e. the density ρ and the ‘weighted vorticity’ ς defined by

$$\varsigma = \nabla \times (\rho\mathbf{u}) = (\rho w)_x - (\rho u)_z. \tag{2.3}$$

From (2.1), the equations of motion for these two fields are

$$\begin{aligned} \rho_t + u\rho_x + w\rho_z &= 0 \\ \varsigma_t + u\varsigma_x + w\varsigma_z + \rho_x(gz - \frac{1}{2}(u^2 + w^2))_z + \frac{1}{2}\rho_z(u^2 + w^2)_x &= 0. \end{aligned} \tag{2.4}$$

These can be written in the form

$$\rho_t = - \left[\rho, \frac{\delta\mathcal{H}}{\delta\varsigma} \right], \quad \varsigma_t = - \left[\rho, \frac{\delta\mathcal{H}}{\delta\rho} \right] - \left[\varsigma, \frac{\delta\mathcal{H}}{\delta\varsigma} \right], \tag{2.5}$$

where, by definition, $[A, B] = A_x B_z - A_z B_x$, and the functional

$$\mathcal{H} = \int_{\mathcal{D}} \rho \left(\frac{1}{2} |\mathbf{u}|^2 + gz \right) dx dz = \int_{\mathcal{D}} \rho \left(\frac{1}{2} |\nabla\Psi|^2 + gz \right) dx dz \tag{2.6}$$

is simply given by the sum of the kinetic and potential energy, \mathcal{D} being the fluid domain $\mathbb{R} \times (-h_2, h_1)$. The streamfunction Ψ is here used as a placeholder for the map between the weighted vorticity ς and \mathbf{u} defined by $\varsigma = (\rho u)_z - (\rho w)_x \equiv -(\rho\Psi_z)_z - (\rho\Psi_x)_x$. As shown in [1], equation (2.5) are a Hamiltonian system with respect to a linear Hamiltonian structure, that is, they can be written as

$$\rho_t = \{\rho, \mathcal{H}\}, \quad \varsigma_t = \{\varsigma, \mathcal{H}\}$$

for the Poisson brackets defined by the Hamiltonian operator

$$J_B = - \begin{pmatrix} 0 & \rho_x \partial_z - \rho_z \partial_x \\ \rho_x \partial_z - \rho_z \partial_x & \varsigma_x \partial_z - \varsigma_z \partial_x \end{pmatrix}. \tag{2.7}$$

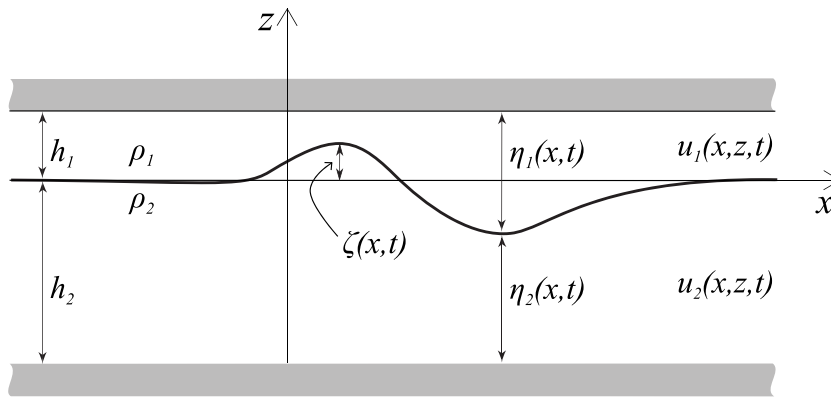


Figure 1. Schematics of the two-layered configuration. The quantity η_2 (resp., η_1) is the total thickness of the lower, heavier (resp., upper, lighter) fluid. The interface ζ is measured from the quiescent state $z = 0$.

2.2. Two-layer case

A simplification of system (2.1) which retains the essential properties of stratification can be obtained by considering a system of two fluids of homogeneous densities $\rho_2 > \rho_1$ in the channel $\mathbb{R} \times (-h_2, h_1)$. The interface between the two homogeneous fluids is described by a smooth function $\zeta = \zeta(x, t)$ (see figure 1). In this case the density and velocity fields can be described as

$$\begin{aligned} \rho(x, z, t) &= \rho_2 + (\rho_1 - \rho_2)\theta(z - \zeta(x, t)) \\ u(x, z, t) &= u_2(x, z, t) + (u_1(x, z, t) - u_2(x, z, t))\theta(z - \zeta(x, t)) \\ w(x, z, t) &= w_2(x, z, t) + (w_1(x, z, t) - w_2(x, z, t))\theta(z - \zeta(x, t)), \end{aligned} \tag{2.8}$$

where θ is the Heaviside function.

A nowadays standard way to reduce the dimensionality of the model is to introduce the layer-averaged velocities as set forth by Wu [38], since in the case of fluids stratified by gravity the vertical direction plays a distinguished role. Let us denote by

$$\bar{u}_1(x, t) = \frac{1}{\eta_1(x, t)} \int_{\zeta}^{h_1} u_1(x, z, t) dz, \quad \bar{u}_2(x, t) = \frac{1}{\eta_2(x, t)} \int_{-h_2}^{\zeta} u_2(x, z, t) dz, \tag{2.9}$$

the layer-averaged velocities, where $\eta_1 = h_1 - \zeta$, $\eta_2 = h_2 + \zeta$ are the thicknesses of the layers. Letting $P(x, t)$ denote the interfacial pressure, the non-homogeneous incompressible Euler equation (2.1) result in the (non-closed) system

$$\begin{aligned} \eta_{it} + (\eta_i \bar{u}_i)_x &= 0, \quad i = 1, 2, \\ \bar{u}_{1t} + \bar{u}_1 \bar{u}_{1x} - g\eta_{1x} + \frac{P_x}{\rho_1} + D_1 &= 0, \\ \bar{u}_{2t} + \bar{u}_2 \bar{u}_{2x} + g\eta_{2x} + \frac{P_x}{\rho_2} + D_2 &= 0. \end{aligned} \tag{2.10}$$

The terms D_1, D_2 at the right hand side of system (2.10) are

$$D_i = \frac{1}{3\eta_i} \partial_x [\eta_i^3 (\bar{u}_{ixt} + \bar{u}_i \bar{u}_{ixx} - (\bar{u}_{ix})^2)] + \dots, \quad i = 1, 2, \tag{2.11}$$

where dots represent terms with nonlocal dependence on the averaged velocities. These terms collect the non-hydrostatic correction to the pressure field, and make the evolution of system (2.10) dispersive. When an asymptotic expansion based on the long-wave assumption $\epsilon \equiv \max[\eta_i/L] \ll 1, i = 1, 2$, is carried out (where L is a typical wavelength), expressions (2.11) explicitly define the leading order dispersive terms in the small parameter ϵ ; truncating at this order makes equation (2.10) local in the layer averaged velocities, resulting in the strongly nonlinear system studied in, e.g. [12].

It is important to notice that the first two equations in (2.10), which have the meaning of mass conservation laws,

$$\eta_{jt} + \partial_x(\eta_j \bar{u}_j) = 0, \tag{2.12}$$

are actually the counterpart of the kinematic boundary conditions at the interface. Denoting (here and in what follows) interface velocities by $\tilde{u}_j(x, t) = u_j(x, \zeta(x, t), t)$ and $\tilde{w}_j(x, t) = w_j(x, \zeta(x, t), t)$, this can be seen from the chain of equalities (with $j = 2$),

$$\begin{aligned} \eta_{2t} + \partial_x \int_{-h_2}^{\zeta} u_2(x, z, t) dz &= \zeta_t + \zeta_x \tilde{u}_2 + \int_{-h_2}^{\zeta} u_{2x}(x, z, t) dz \\ &= (\text{using } u_{2x} + w_{2z} = 0) = \zeta_t + \zeta_x \tilde{u}_2 - \int_{-h_2}^{\zeta} w_{2z}(x, z, t) dz \\ &= \zeta_t + \zeta_x \tilde{u}_2 - w_2 \Big|_{-h_2}^{\zeta} = (\text{by the bottom no flux condition}) \\ &= \zeta_t + \zeta_x \tilde{u}_2 - \tilde{w}_2 = 0, \end{aligned} \tag{2.13}$$

and similarly for the upper fluid when $j = 1$. Equation (2.10) come equipped with two constraints. Namely, we have the obvious geometrical constraint $\eta_1 + \eta_2 = h$ and its consequence obtained by summing the equations in the first line of (2.10),

$$(\eta_1 \bar{u}_1 + \eta_2 \bar{u}_2)_x = 0. \tag{2.14}$$

We remark that under suitable far-field boundary conditions (such as vanishing or identical velocities for $x \rightarrow \pm\infty$) this relation translates into the *dynamical* constraint

$$\eta_1 \bar{u}_1 + \eta_2 \bar{u}_2 = 0. \tag{2.15}$$

2.3. The Hamiltonian reduction process

We now discuss how a simple averaging process can be given a Hamiltonian structure, well suited to the discussion of the constrained equations in which our set of reduced coordinates naturally appears. We follow the setting already introduced in [9] and provide a full geometric description of the reduction process. We begin with definitions (2.8), where we suppress time dependence for ease of notation in what follows. The two momentum components are

$$\begin{aligned} \rho u &= \rho_2 u_2(x, z) + (\rho_1 u_1(x, z) - \rho_2 u_2(x, z))\theta(z - \zeta(x)), \\ \rho w &= \rho_2 w_2(x, z) + (\rho_1 w_1(x, z) - \rho_2 w_2(x, z))\theta(z - \zeta(x)), \end{aligned} \tag{2.16}$$

so that the weighted vorticity (2.3) is

$$\begin{aligned} \varsigma &= \rho_2(w_{2x} - u_{2z}) + (\rho_1(w_{1x} - u_{1z}) - \rho_2(w_{2x} - u_{2z}))\theta(z - \zeta(x)) \\ &\quad - (\rho_1 u_1(x, z) - \rho_2 u_2(x, z) + \zeta_x(\rho_1 w_1(x, z) - \rho_2 w_2(x, z)))\delta(z - \zeta(x)), \end{aligned} \tag{2.17}$$

where $\delta(\cdot)$ is the Dirac delta function.

We assume that the motion in each layer is irrotational, so that we are left with a ‘momentum vortex line’ along the interface, that is,

$$\varsigma = (\rho_2 u_2(x, z) - \rho_1 u_1(x, z) + \zeta_x(\rho_2 w_2(x, z) - \rho_1 w_1(x, z)))\delta(z - \zeta(x)). \quad (2.18)$$

We define a projection map 2D \rightarrow 1D as

$$\zeta(x) = \frac{1}{\rho_\Delta} \int_{-h_2}^{h_1} (\rho(x, z) - \rho_1) dz - h_2, \quad \sigma(x) = \int_{-h_2}^{h_1} \varsigma(x, z) dz, \quad (2.19)$$

where $\rho_\Delta = \rho_2 - \rho_1$. When applied to two-layer configurations, the first of these relations is easily obtained from the first of equation (2.8). Moreover, in the two-layer bulk irrotational case,

$$\sigma(x) = \rho_2 u_2(x, \zeta(x)) - \rho_1 u_1(x, \zeta(x)) + \zeta_x(x)(\rho_2 w_2(x, \zeta(x)) - \rho_1 w_1(x, \zeta(x))), \quad (2.20)$$

i.e. the averaged weighted vorticity σ is the tangential momentum shear at the interface.

The geometry thus far outlined fits the Hamiltonian reduction scheme devised in [31]. Indeed, such a scheme considers a manifold \mathcal{P} endowed with a Poisson tensor, such as J_B , a submanifold $\mathcal{M} \subset \mathcal{P}$, a distribution D contained in the tangent bundle to \mathcal{P} restricted to \mathcal{M} , $T\mathcal{P}|_{\mathcal{M}}$, and state that a Poisson reduction to \mathcal{M}/Φ , with Φ denoting the intersection $T\mathcal{M} \cap D$, is possible when (some geometrical assumptions on the regularity of D and on its action on \mathcal{M} being granted)

1. J_B is invariant under D .
2. At each point of \mathcal{M} it holds

$$J_B(D^0) \subset T\mathcal{M} + D, \quad (2.21)$$

$D^0 \subset T^*\mathcal{P}|_{\mathcal{M}}$ being the annihilator of D in the cotangent bundle to \mathcal{P} restricted to \mathcal{M} .

In particular (see the example in [31]), in our case we identify the following geometric objects:

1. \mathcal{P} is the configuration space $M^{(2)}$ of the 2D fields, parameterised by $(\rho(x, z), \varsigma(x, z))$, and J_B is the Benjamin Poisson tensor (2.7)

$$J_B = - \begin{pmatrix} 0 & \rho_x \partial_z - \rho_z \partial_x \\ \rho_x \partial_z - \rho_z \partial_x & \varsigma_x \partial_z - \varsigma_z \partial_x \end{pmatrix}. \quad (2.22)$$

2. \mathcal{M} is given by the two-layer configuration space

$$\{\rho(x, z) = \rho_2 - \rho_\Delta \theta(z - \zeta(x)), \varsigma(x, z) = \sigma(x)\delta(z - \zeta(x))\}. \quad (2.23)$$

3. D is the image under J_B of the annihilator $T\mathcal{M}^0$ of the tangent space to \mathcal{M} in $TM^{(2)}|_{\mathcal{M}}$.

To show how our model fits the Marsden–Ratiu scheme we first notice that the $T\mathcal{M}$ can be described as the space of pairs of generalised functions of the form

$$\{\dot{\rho} = \rho_\Delta \dot{\zeta} \delta(z - \zeta), \dot{\zeta} = \sigma \delta(z - \zeta) - \sigma \dot{\zeta}'(z - \zeta)\}, \quad (2.24)$$

δ' being the derivative of the Dirac’s- δ function. Notice the link between the $\delta(z - \zeta)$ -coefficient of $\dot{\rho}$ and the $\delta'(z - \zeta)$ -coefficient of $\dot{\zeta}$. The annihilator $T\mathcal{M}^0$ is readily computed as pairs of smooth functions $(\phi(x, z), \psi(x, z))$ satisfying

$$\psi(x, \zeta) = 0, \quad \rho_\Delta \phi(x, \zeta) + \sigma \psi_z(x, \zeta) = 0. \quad (2.25)$$

Since on \mathcal{M} we have

$$\begin{aligned} \rho_x &= \rho_\Delta \zeta_x \delta(z - \zeta), & \rho_z &= -\rho_\Delta \delta(z - \zeta), \\ \varsigma_x &= \sigma_x \delta(z - \zeta) - \sigma \zeta_x \delta'(z - \zeta), & \varsigma_z &= \sigma \delta'(z - \zeta), \end{aligned} \tag{2.26}$$

the restriction of the Poisson tensor J_B on \mathcal{M} acquires the form

$$J_B|_{\mathcal{M}} = - \begin{pmatrix} 0 & \delta(z - \zeta) \rho_\Delta (\zeta_x \partial_z + \partial_x) \\ \delta(z - \zeta) \rho_\Delta (\zeta_x \partial_z + \partial_x) & \delta(z - \zeta) \sigma_x \partial_z - \delta'(z - \zeta) \sigma (\zeta_x \partial_z + \partial_x) \end{pmatrix}. \tag{2.27}$$

Hence the image of $J_B|_{\mathcal{M}}$ is the space of vectors

$$\begin{pmatrix} \dot{\rho} \\ \dot{\varsigma} \end{pmatrix} = - \begin{pmatrix} \rho_\Delta (\zeta_x \psi_z + \psi_x) \delta(z - \zeta) \\ (\rho_\Delta (\zeta_x \phi_z + \phi_x) + \sigma_x \psi_z) \delta(z - \zeta) - \sigma (\zeta_x \psi_z + \psi_x) \delta'(z - \zeta) \end{pmatrix}. \tag{2.28}$$

This expression can be used to show that $D = J_B(T\mathcal{M}^0)$ reduces to the null vector. In fact, let us consider (2.28) with (ϕ, ψ) in $T\mathcal{M}^0$. Thanks to the δ -function factor, the first component of (2.28) can be written as $\dot{\rho} = -\rho_\Delta (\zeta_x \psi_z(x, \zeta) + \psi_x(x, \zeta)) \delta(z - \zeta)$, and the coefficient of the δ vanishes being the total x -derivative of the first of (2.25). By using the generalised function identity $f(y)\delta'(y) = f(0)\delta'(y) - f'(0)\delta(y)$ the second component of (2.28) can be written as

$$\begin{aligned} \dot{\varsigma} &= \sigma (\zeta_x \psi_z(x, \zeta) + \psi_x(x, \zeta)) \delta'(z - \zeta) \\ &\quad - \left(\rho_\Delta (\zeta_x \phi_z(x, \zeta) + \phi_x(x, \zeta)) + \sigma_x \psi_z(x, \zeta) + \sigma \zeta_x \psi_{zz}(x, \zeta) + \sigma \psi_{xz}(x, \zeta) \right) \delta(z - \zeta), \end{aligned} \tag{2.29}$$

which vanishes as well thanks to (2.25). The vanishing of D confirms that the reduced manifold is isomorphic to the submanifold \mathcal{M} , which guarantees the invariance of J_B . As for the characteristic condition for reduction, i.e. equation (2.21), this follows explicitly from (2.28) which displays how the image $J_B|_{\mathcal{M}}$ is contained in $T\mathcal{M}$ as determined in equation (2.24).

We can now compute the expression of the reduced Poisson tensor as follows. We consider the pull-back to $M^{(2)}$ of a generic 1-form $(\mu_\zeta(x), \mu_\sigma(x))$ on the manifold $M^{(1)}$, parameterised by $(\zeta(x), \sigma(x))$, under the map (2.19), given by

$$\begin{pmatrix} 1 \\ \rho_\Delta \mu_\zeta(x), \mu_\sigma(x) \end{pmatrix}. \tag{2.30}$$

Applying the Poisson tensor (2.7) evaluated on \mathcal{M} to this covector, we obtain

$$\begin{pmatrix} \dot{\rho}(x, z) \\ \dot{\varsigma}(x, z) \end{pmatrix} = - \begin{pmatrix} \rho_\Delta \delta(z - \zeta(x)) (\mu_\sigma(x))_x \\ \delta(z - \zeta(x)) (\mu_\zeta(x))_x - \sigma(x) \delta'(z - \zeta(x)) (\mu_\sigma(x))_x \end{pmatrix}. \tag{2.31}$$

Pushing this vector to $M^{(1)}$ via the tangent map to (2.19),

$$\dot{\zeta} = \frac{1}{\rho_\Delta} \int_{-h_2}^{h_1} \dot{\rho}(x, z) dz, \quad \dot{\sigma} = \int_{-h_2}^{h_1} \dot{\varsigma}(x, z) dz,$$

yields the vector

$$(\dot{\zeta}, \dot{\sigma}) = (-\partial_x \mu_\sigma, -\partial_x \mu_\zeta),$$

owing to the fact that $\int_{-h_2}^{h_1} \delta'(z - \zeta(x)) dz = 0$ (we work under the assumption that the fluid interface never touches the boundary, i.e. the strict inequalities $-h_2 < \zeta < h_1$ hold).

Hence, the expression of the reduction of the Benjamin Poisson tensor J_B on the manifold $M^{(1)}$ is given in the coordinates $(\zeta(x), \sigma(x))$ by the constant tensor

$$J_{\text{red}} = - \begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix}. \tag{2.32}$$

This structure coincides with the one introduced in [2] by a direct inspection of the Hamiltonian formulation of two-layer models. We stress that within our setting the above Poisson tensor is obtained by the process of Hamiltonian reduction from the Lie-Poisson structure of the general heterogeneous incompressible Euler 2D fluids of [1]. Moreover, by means of our choice of reducing map (2.19), we directly obtain a set of coordinates (ζ, σ) that can be called *Darboux* coordinates, since they are the analog of the coordinates (u, v) for the nonlinear wave equation in 1 + 1 dimensions $u_{tt} = F''(u)u_{xx}$ derived from the Hamiltonian functional $\mathcal{H} = \frac{1}{2} \int_{\mathbb{R}} (u_t^2 + F(u)) dx$ by means of the Poisson structure (2.32).

2.4. The evolution variables and the Hamiltonian

The basic feature of the Hamiltonian reduction process is that with this approach the natural dependent variables are the displacement from the equilibrium position ζ and the tangential interface momentum shear

$$\begin{aligned} \sigma(x) &= \rho_2 u_2(x, \zeta(x)) - \rho_1 u_1(x, \zeta(x)) + \zeta_x(x)(\rho_2 w_2(x, \zeta(x)) - \rho_1 w_1(x, \zeta(x))) \\ &\equiv \rho_2 \tilde{u}_2(x) - \rho_1 \tilde{u}_1(x) + \zeta_x(x)(\rho_2 \tilde{w}_2(x) - \rho_1 \tilde{w}_1(x)) \end{aligned} \tag{2.33}$$

(we recall and use hereafter that a tilde over a quantity stands for its evaluation at the interface, e.g. $\tilde{u}_1(x, t) = u_1(x, \zeta(x, t), t)$ etc).

In this respect, the approach we pursue here differs from the Green–Naghdi setting of, e.g. [12], which considers *layer averaged* velocities (following the seminal paper [38]). Specifically, here we shall use and adapt to our case the setting discussed in [40] (see also [2, 41]) where the equations for internal wave motion are written using two sets of coordinates:

- i) the *boundary velocity basis*, in which ζ is complemented by $u_{01}(x, t) = u_1(x, h_1, t)$ in the upper layer and by $u_{02}(x, t) = u_2(x, h_2, t)$ in the lower layer.
- ii) the *interface velocity basis*, where we use the variables entering the Hamiltonian reduction process, that is $\tilde{u}_j(x, t) = u_j(x, \zeta(x, t), t)$.

While for some aspects of the theory it is advantageous to use layer-mean velocities (see [11, 12, 40]), as mentioned above these are not the ones most naturally suggested by our Hamiltonian reduction procedure, and therefore we choose to express energy, the mass conservation as well as the ensuing dynamical constraint in terms of interface variables.

Following [37, 40], we use the assumed bulk irrotationality of the fluid flow to introduce the bulk velocity potentials $\varphi_j(x, z)$, which we Taylor expand with respect to the vertical variable z . By the vanishing of the vertical velocity at the physical boundaries $z = h_1$, and $z = -h_2$ we obtain the Taylor expansions

$$\varphi_j(x, z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} H_j(z)^{2n} \partial_x^{2n} \varphi_{0j}(x) \tag{2.34}$$

where

$$H_1(z) = z - h_1, \quad H_2(z) = z + h_2, \tag{2.35}$$

and $\varphi_{01}(x) = \varphi_1(x, h_1)$, $\varphi_{02} = \varphi_2(x, -h_2)$ are the values of the potential at the rigid lids.

The horizontal velocities are then given by

$$u_j = \partial_x \varphi_j(x, z) = \sum_{j=0}^{\infty} \frac{(-1)^n}{(2n)!} H_j(z)^{2n} \partial_x^{2n} \partial_x \varphi_{0j}(x) = \sum_{j=0}^{\infty} \frac{(-1)^n}{(2n)!} H_j(z)^{2n} \partial_x^{2n} u_{0j}(x), \tag{2.36}$$

$u_{0j}(x)$ being the horizontal velocities at $z = h_1$ (for $j = 1$) and at $z = -h_2$ (for $j = 2$).

Likewise, the vertical velocities are given by

$$w_j(x, z) = \partial_z \varphi_j(x, z) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!} H_j(z)^{2n+1} \partial_x^{2n+1} u_{0j}(x). \tag{2.37}$$

Notice that the boundary conditions $w_1(x, h_1) = w_2(x, -h_2) = 0$ are satisfied.

Since

$$H_1(\zeta) = -\eta_1, \quad H_2(\zeta) = \eta_2, \quad \text{i.e., } H_j(\zeta) = (-1)^j \eta_j, \quad j = 1, 2, \tag{2.38}$$

where $\eta_1(x) = h_1 - \zeta(x)$ (resp. $\eta_2(x) = h_2 + \zeta(x)$) is the thickness of the upper (resp. lower) layer, the interface velocities can be directly obtained by formulas (2.36) and (2.37) as

$$\tilde{u}_j = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \eta_j^{2n} \partial_x^{2n} u_{0j}(x), \quad \tilde{w}_j = (-1)^{j-1} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \eta_j^{2n+1} \partial_x^{2n+1} u_{0j}(x). \tag{2.39}$$

For later use, we express (from the same formulas) the layer-mean horizontal velocities in terms of the fluid thicknesses and the (respective) boundary velocities as

$$\begin{aligned} \bar{u}_1(x) &\equiv \frac{1}{\eta_1} \int_{\zeta}^{h_1} u_1(x, z) \, dz = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \eta_1(x)^{2n} \partial_x^{2n} u_{01}(x) \\ \bar{u}_2(x) &\equiv \frac{1}{\eta_2} \int_{-h_2}^{\zeta} u_2(x, z) \, dz = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \eta_2(x)^{2n} \partial_x^{2n} u_{02}(x). \end{aligned} \tag{2.40}$$

2.5. Rescaling the spatial independent variables: the ϵ -expansion and the mass conservation laws

To make the formal Taylor series (2.39) and (2.40) effective in the construction of asymptotic models for interfacial wave motion we have to rescale variables (see, e.g. [37, 40]). In particular, we set

$$x = Lx^*, \quad z = hz^*, \tag{2.41}$$

where L is a typical horizontal scale (say, a typical wavelength) and h is the total height of the vertical channel. As usual, we assume that the ratio $\epsilon = h/L$ be the small dispersion parameter of the theory. Indeed, by using these scalings, we can turn the Taylor series (2.36) and (2.37) as well as (2.39) and (2.40) into asymptotic series in the small parameter ϵ .

For the sake of simplicity, hereafter we shall drop asterisks from the formulas. We remark that, unless otherwise explicitly stated, horizontal lengths are scaled by L and vertical lengths by h . Henceforth, we will abuse notation a little and use the order symbol $O(\cdot)$ to denote the magnitude of bounded dimensional quantities whenever this can be done without generating confusion.

For the velocity fields we have

$$\begin{aligned} u_j(x, z) &= \sum_{j=0}^{\infty} \frac{(-1)^n}{(2n)!} \epsilon^{2n} H_j(z)^{2n} \partial_x^{2n} u_{0j}(x), \\ w_j(x, z) &= (-1)^{j-1} \epsilon \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \epsilon^{2n} H_j(z)^{2n+1} \partial_x^{2n+1} u_{0j}(x). \end{aligned} \tag{2.42}$$

It is worth taking into account here and below the expected (Lagrangian) scaling of vertical vs. horizontal velocities $w_j/u_j = O(\epsilon)$. Similarly we have

$$\begin{aligned} \tilde{u}_j &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \epsilon^{2n} \eta_j^{2n} \partial_x^{2n} u_{0j} = u_{0j} - \frac{\epsilon^2}{2} \eta_j^2 u_{0j,xx} + O(\epsilon^4) \\ \tilde{w}_j &= (-1)^{j-1} \epsilon \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \epsilon^{2n} \eta_j^{2n+1} \partial_x^{2n+1} u_{0j} \\ &= \epsilon (-1)^{j-1} \left(\eta_j u_{0j,x} - \frac{\epsilon^2}{6} \eta_j^3 u_{0j,xxx} + O(\epsilon^4) \right) \\ \bar{u}_j &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \epsilon^{2n} \eta_j^{2n} \partial_x^{2n} u_{0j} = u_{0j} - \frac{\epsilon^2}{6} \eta_j^2 u_{0j,xx} + O(\epsilon^4). \end{aligned} \tag{2.43}$$

It should be noticed that, contrary to [37, 40], for the time being we do not rescale the dependent variables u, w ; this will be done at a later stage, when we shall rescale the Hamiltonian variable σ once the constraints mentioned above in section 2 will be taken into account.

At leading order in the expansion with respect to the small dispersion parameter ϵ , we have

$$\bar{u}_j = \tilde{u}_j, \bar{w}_j = \tilde{w}_j \simeq 0, \text{ with } \sigma = \rho_2 \tilde{u}_2 - \rho_1 \tilde{u}_1 = \rho_2 \bar{u}_2 - \rho_1 \bar{u}_1, \tag{2.44}$$

that is, σ reduces to the *horizontal* momentum shear. At this order one can view the motion as satisfying the so-called columnar motion ansatz (see, e.g. [36]). Thus at higher orders the ansatz fails, since we have

$$\sigma = \rho_2 \tilde{u}_2 - \rho_1 \tilde{u}_1 + \epsilon \zeta_x (\rho_2 \tilde{w}_2 - \rho_1 \tilde{w}_1) \tag{2.45}$$

and columnar motion is no longer consistent with (2.42).

For the reader's convenience, we now collect in compact form a few consequences of the expansions (2.43), that can be found in section 13 of [37]. First, from (2.43) notice that inverting

$$\tilde{u}_j = u_{0j} - \frac{\epsilon^2}{2} \eta_j^2 u_{0j,xx} + O(\epsilon^4). \tag{2.46}$$

yields

$$u_{0j} = \tilde{u}_j + \frac{\epsilon^2}{2} \eta_j^2 \tilde{u}_{j,xx} + O(\epsilon^4). \tag{2.47}$$

A straightforward computation shows that

$$\tilde{w}_j = (-1)^{j+1} \epsilon \left(\eta_j \tilde{u}_{j,x} + \frac{\epsilon^2}{3} (\eta_j^3 \tilde{u}_{j,xx})_x + O(\epsilon^4) \right). \tag{2.48}$$

Also, as far as the asymptotic relations between interface and layer-averaged velocities are concerned, we have, again from (2.43),

$$\bar{u}_j = u_{0j} - \frac{\epsilon^2}{6} \eta_j^2 u_{0j,xx} + O(\epsilon^4), \tag{2.49}$$

which yields, thanks to (2.47),

$$\bar{u}_j = \tilde{u}_j + \frac{\epsilon^2}{3} \eta_j^2 \tilde{u}_{j,xx} + O(\epsilon^4). \tag{2.50}$$

The mass conservation laws for the two fluids, expressed without approximation by the pair of equations

$$\eta_{jt} + \partial_x(\eta_j \bar{u}_j) = 0, \quad j = 1, 2, \tag{2.51}$$

are translated, by (2.50), into the approximate mass conservation laws

$$\eta_{jt} + \partial_x(\eta_j \tilde{u}_j) + \frac{\epsilon^2}{3} \partial_x(\eta_j^3 \tilde{u}_{j,xx}) = O(\epsilon^4), \quad j = 1, 2. \tag{2.52}$$

Hence, the dynamic constraint (2.15) obtained by summing the two equation (2.51), taking into account the geometric constraint $\eta_1 + \eta_2 = h$ together with the far-field vanishing conditions, translates into the *approximate dynamical constraint*

$$\eta_1 \tilde{u}_1 + \eta_2 \tilde{u}_2 + \frac{\epsilon^2}{3} (\eta_1^3 \tilde{u}_{1,xx} + \eta_2^3 \tilde{u}_{2,xx}) = O(\epsilon^4). \tag{2.53}$$

2.6. The energy

Our next task is to write the explicit form (at order $O(\epsilon^2)$) of the energy. All the asymptotic manipulations needed are for the kinetic energy, the potential energy is straightforward and can be written out immediately at every order. The asymptotic analysis is somewhat equivalent to that used in other approaches, (see, e.g. [17]) starting from the different viewpoint of expanding, having assumed two-layer dynamics from the outset, the so-called Dirichlet-to-Neumann operator in each layer, and we can anticipate here that it will produce similar dispersive terms in the long-wave expansions below. However, besides the different starting point of geometric Hamiltonian reduction, our approach will also focus on the need to allow for different balances between nonlinearity and dispersion to capture, both qualitatively and quantitatively, fundamental features of the dynamics, while simultaneously striving for the simplest possible models.

Let us consider the lower fluid first. Its kinetic energy density reads

$$T_2 = \frac{\rho_2}{2} \int_{-h_2}^{\zeta} (u_2^2 + w_2^2) h dz, \tag{2.54}$$

(the dimensional factor h coming from the scaling of z). By Taylor-expanding, we have

$$u_2(x, z) = u_{20}(x) - \frac{\epsilon^2}{2} (z + h_2)^2 u_{20,xx}(x) + O(\epsilon^4). \tag{2.55}$$

By (2.47) we get

$$u_2(x, z) = \tilde{u}_2(x) + \frac{\epsilon^2}{2} (\eta_2^2(x) - (z + h_2)^2) \tilde{u}_{2,xx}(x) + O(\epsilon^4), \tag{2.56}$$

and by (2.42) and (2.47),

$$w_2(x, z) = -\epsilon(z + h_2) \tilde{u}_{2,x}(x) + O(\epsilon^2). \tag{2.57}$$

This leads to

$$\begin{aligned} T_2 &= \frac{h\rho_2}{2} \int_{-h_2}^{\zeta} \left[\tilde{u}_2^2 + \epsilon^2 (\tilde{u}_2 \tilde{u}_{2,xx} (\eta_2^2 - (z + h_2)^2) + \tilde{u}_{2,x}^2 (z + h_2)^2) + O(\epsilon^4) \right] dz \\ &= \frac{h\rho_2}{2} \left[\eta_2 \tilde{u}_2^2 + \frac{\epsilon^2}{3} \eta_2^3 (2\tilde{u}_2 \tilde{u}_{2,xx} + \tilde{u}_{2,x}^2) \right] + O(\epsilon^4). \end{aligned} \tag{2.58}$$

By the same arguments we obtain the contribution to the total kinetic energy density of the upper fluid as

$$T_1 = \frac{h\rho_2}{2} \int_{\zeta}^{h_1} (u_1^2 + w_1^2) dz = \frac{h\rho_1}{2} \left[\eta_1 \tilde{u}_1^2 + \frac{\epsilon^2}{3} \eta_2^3 (2\tilde{u}_1 \tilde{u}_{1,xx} + \tilde{u}_{1,x}^2) \right] + O(\epsilon^4). \tag{2.59}$$

In formulas (2.58) and (2.59) we used, respectively, $\eta_2 = \zeta + h_2$ and $\eta_1 = h_1 - \zeta$.

As mentioned above, the computation of the potential energy density is more direct: taking non-dimensionalization into account, we have

$$U = h^2 g \left(\int_{-h_2}^{\zeta} \rho_2 z dz + \int_{\zeta}^{h_1} \rho_1 z dz \right) = \frac{1}{2} h^2 \left(g(\rho_2 - \rho_1) \zeta^2 - \frac{1}{2} g(\rho_2 h_2^2 - \rho_1 h_1^2) \right), \tag{2.60}$$

where h is again the total distance between top and bottom plates.

3. Nonlinear asymptotics

In what follows, we shall deal with a simplified model, defined by the following requirements:

1. The interface displacement ζ will be understood to be scaled by its maximum value a , to yield the amplitude nondimensional small parameter $\alpha = \frac{a}{h} \ll 1$. Namely, the non-dimensional fluid thicknesses η_j will be written as

$$\eta_j = h_j + (-1)^j \alpha \zeta. \tag{3.1}$$

2. We shall make an asymptotic expansion in the small parameters α and ϵ and mainly consider the ‘mildly nonlinear’ (MNL) case, defined by the relative scaling $\epsilon^2 \ll \alpha \ll \epsilon$. We shall thus discard terms of order $\alpha\epsilon^2, \epsilon^3$ and higher, but retain terms of order α^2 . The usual WNL case (see, e.g. [37]), where $\alpha = O(\epsilon^2)$, can be seen as a special case of the MNL case (see section 4).

The first consequences of such scaling limits are the following:

- i) The slope ζ_x of the normalized interface is small and scales as $\frac{a}{hL} = O(\alpha\epsilon)$.
- ii) Since \tilde{w}_j scales as ϵ and, by the previous point, ζ_x scales as $\alpha\epsilon$, the Hamiltonian variable

$$\sigma = \rho_2 \tilde{u}_2 - \rho_1 \tilde{u}_1 + \zeta_x (\rho_2 \tilde{w}_2 - \rho_1 \tilde{w}_1) \tag{3.2}$$

within this asymptotics becomes

$$\sigma = \rho_2 \tilde{u}_2 - \rho_1 \tilde{u}_1, \tag{3.3}$$

which has the same form as that of the dispersionless approximation.

- iii) The approximate dynamical constraint (2.53) gets simplified as well, and reads

$$\eta_1 \tilde{u}_1 + \eta_2 \tilde{u}_2 + \frac{\epsilon^2}{3} (h_1^3 \tilde{u}_{1,xx} + h_2^3 \tilde{u}_{2,xx}) = O(\epsilon^4). \tag{3.4}$$

iv) There is a notable simplification in the kinetic energy densities (2.58) and (2.59). Indeed, for the lower fluid under approximations (2.58) one gets

$$T_2 = \frac{h\rho_2}{2} \left(\eta_2 \tilde{u}_2^2 + \frac{1}{3} \epsilon^2 h_2^3 \left(2\tilde{u}_2 \tilde{u}_{2,xx} + \tilde{u}_{2,x}^2 \right) \right). \tag{3.5}$$

Notice that the ϵ^2 term can be written as $\frac{h_2^3}{3} \tilde{u}_2 \tilde{u}_{2,xx} + \frac{h_2^3}{6} (\tilde{u}_2^2)_{,xx}$, and this second term, being a total derivative, does not contribute to the Hamiltonian formulation of the equations of motion.

Repeating the argument for the upper fluid, the total kinetic energy density, still within the same MLN asymptotics, can be written as

$$T = T_1 + T_2 = \frac{h}{2} \left(\rho_1 \left(\eta_1 \tilde{u}_1^2 + \frac{\epsilon^2}{3} h_1^3 \left(\tilde{u}_1 \tilde{u}_{1,xx} + \frac{1}{2} (\tilde{u}_1^2)_{,xx} \right) \right) + \rho_2 \left(\eta_2 \tilde{u}_2^2 + \frac{\epsilon^2}{3} h_2^3 \left(\tilde{u}_2 \tilde{u}_{2,xx} + \frac{1}{2} (\tilde{u}_2^2)_{,xx} \right) \right) \right). \tag{3.6}$$

3.1. The Hamiltonian in Darboux coordinates

Our next task is to express the Hamiltonian density $H = T + U$ of the asymptotic model in terms of the Darboux coordinates dictated by the Hamiltonian reduction process of § 2.3, that is, the pair ζ and σ given by (3.3). To this end, we make use of the geometrical constraint $\eta_1 + \eta_2 = h$ and the approximate dynamical constraint given by equation (2.53). Our strategy will be to use the weak nonlinearity assumption to simplify the dispersive terms first, and deal with small α expansion for the quasilinear terms afterwards, since the latter do not contain x -derivatives in the Hamiltonian, which can then be expanded in a standard Taylor series.

As remarked above, within the present asymptotic theory, the dynamical constraint reads

$$\eta_1 \tilde{u}_1 + \eta_2 \tilde{u}_2 + \frac{\epsilon^2}{3} (h_1^3 \tilde{u}_{1,xx} + h_2^3 \tilde{u}_{2,xx}) = 0. \tag{3.7}$$

Rewriting the latter in operator form as the equality

$$\eta_1 \left(\mathbf{1} + \frac{\epsilon^2}{3} h_1^2 \partial_x^2 \right) \tilde{u}_1 = -\eta_2 \left(\mathbf{1} + \frac{\epsilon^2}{3} h_2^2 \partial_x^2 \right) \tilde{u}_2, \tag{3.8}$$

which is correct up to terms of order $\alpha \epsilon^2$, and by using the approximate inversion formula for near-identity operators $(\mathbf{1} + \epsilon^2 \hat{A})^{-1} = \mathbf{1} - \epsilon^2 \hat{A} + O(\epsilon^4)$, we get

$$\tilde{u}_1 = - \left(\mathbf{1} - \frac{\epsilon^2}{3} h_1^2 \partial_x^2 \right) \left(\frac{\eta_2}{\eta_1} \left(\mathbf{1} + \frac{\epsilon^2}{3} h_2^2 \partial_x^2 \right) \right) \tilde{u}_2, \tag{3.9}$$

up to higher order terms in ϵ^2 . Since $\frac{\eta_2}{\eta_1} = \frac{h_2}{h_1} + O(\alpha)$ we arrive at the relation

$$\tilde{u}_1 = - \frac{\eta_2}{\eta_1} \tilde{u}_2 + \frac{\epsilon^2}{3} \frac{h_2}{h_1} (h_1^2 - h_2^2) \tilde{u}_{2,xx}. \tag{3.10}$$

Recall that the kinetic energy density is represented, at $O(\epsilon^2)$ and in this WNL asymptotics, by

$$T = \frac{h}{2} \left(\rho_1 \left(\eta_1 \tilde{u}_1^2 + \frac{\epsilon^2}{3} h_1^3 \tilde{u}_1 \tilde{u}_{1,xx} \right) + \rho_2 \left(\eta_2 \tilde{u}_2^2 + \frac{\epsilon^2}{3} h_2^3 \tilde{u}_2 \tilde{u}_{2,xx} \right) \right) \text{ plus total derivatives.} \tag{3.11}$$

At the order of approximation we are working with we can substitute

$$\tilde{u}_{1,xx} = -\frac{h_2}{h_1} \tilde{u}_{2,xx} \tag{3.12}$$

in the $O(\epsilon^2)$ terms of this expression as well as in the approximate dynamical constraint (3.7), which therefore turns into

$$\tilde{u}_1 \eta_1 + \tilde{u}_2 \eta_2 + \frac{1}{3} \epsilon^2 \tilde{u}_{2,xx} h_2 (h_2^2 - h_1^2) = 0. \tag{3.13}$$

By solving this algebraic constraint and the defining relation (3.3) with respect to the velocities, we get the implicit relations

$$\begin{aligned} \tilde{u}_1 &= -\frac{\sigma \eta_2}{\eta_1 \rho_2 + \eta_2 \rho_1} + \frac{\epsilon^2 \tilde{u}_{2,xx} h_2 \rho_2 (h_1^2 - h_2^2)}{3 (\eta_1 \rho_2 + \eta_2 \rho_1)} \\ \tilde{u}_2 &= \frac{\sigma \eta_1}{\eta_1 \rho_2 + \eta_2 \rho_1} + \frac{\epsilon^2 \tilde{u}_{2,xx} h_2 \rho_1 (h_1^2 - h_2^2)}{3 (\eta_1 \rho_2 + \eta_2 \rho_1)}. \end{aligned} \tag{3.14}$$

Now we can use the fact that the second derivative $\tilde{u}_{2,xx}$ appears only in terms $O(\epsilon^2)$, so that we can substitute (3.12) into

$$\sigma_{xx} = \rho_2 \tilde{u}_{2,xx} - \rho_1 \tilde{u}_{1,xx} \tag{3.15}$$

leading to

$$\tilde{u}_{2,xx} = \frac{\sigma_{xx} h_1}{h_1 \rho_2 + h_2 \rho_1}. \tag{3.16}$$

Hence, equation (3.14) become

$$\begin{aligned} \tilde{u}_1 &= -\frac{\eta_2 \sigma}{\eta_1 \rho_2 + \eta_2 \rho_1} + \rho_2 \frac{\epsilon^2}{3} \frac{h_1 h_2 (h_1^2 - h_2^2)}{(h_1 \rho_2 + h_2 \rho_1)(\eta_1 \rho_2 + \eta_2 \rho_1)} \sigma_{xx}, \\ \tilde{u}_2 &= \frac{\eta_1 \sigma}{\eta_1 \rho_2 + \eta_2 \rho_1} + \rho_1 \frac{\epsilon^2}{3} \frac{h_1 h_2 (h_1^2 - h_2^2)}{(h_1 \rho_2 + h_2 \rho_1)(\eta_1 \rho_2 + \eta_2 \rho_1)} \sigma_{xx}. \end{aligned} \tag{3.17}$$

Substituting these relations in the expression of the kinetic energy density (3.11) leads, dropping the total derivative terms, to the intermediate expression

$$T = \frac{h}{2} \left(\frac{\eta_1 \eta_2 \sigma^2}{\rho_2 \eta_1 + \rho_1 \eta_2} + \frac{\epsilon^2 h_1^2 h_2^2 (h_1 \rho_1 + h_2 \rho_2)}{3 (h_1 \rho_2 + h_2 \rho_1)^2} \sigma \sigma_{xx} \right). \tag{3.18}$$

Next, the first term in the kinetic energy must be expanded in powers of α to yield our final version of the kinetic energy density

$$\begin{aligned} T &= \frac{h}{2} \frac{h_1 h_2}{(h_1 \rho_2 + h_2 \rho_1)} \sigma^2 + \frac{\alpha h (h_1^2 \rho_2 - h_2^2 \rho_1)}{2 (h_1 \rho_2 + h_2 \rho_1)^2} \zeta \sigma^2 - \frac{\alpha^2}{2} \frac{h^3 \rho_1 \rho_2}{(h_1 \rho_2 + h_2 \rho_1)^3} \zeta^2 \sigma^2 \\ &+ \frac{\epsilon^2 h h_1^2 h_2^2 (h_1 \rho_1 + h_2 \rho_2)}{6 (h_1 \rho_2 + h_2 \rho_1)^2} \sigma \sigma_{xx} + O(\alpha^3, \alpha \epsilon^2, \epsilon^4). \end{aligned} \tag{3.19}$$

Therefore, with the potential energy expression (2.60), the total energy density at this order is

$$\begin{aligned} E &= h \left(\frac{1}{2} \frac{h_1 h_2}{(h_1 \rho_2 + h_2 \rho_1)} \sigma^2 + \frac{\alpha (h_1^2 \rho_2 - h_2^2 \rho_1)}{2 (h_1 \rho_2 + h_2 \rho_1)^2} \zeta \sigma^2 - \frac{\alpha^2}{2} \frac{h^2 \rho_1 \rho_2}{(h_1 \rho_2 + h_2 \rho_1)^3} \zeta^2 \sigma^2 \right) \\ &+ h \frac{\epsilon^2 h_1^2 h_2^2 (h_1 \rho_1 + h_2 \rho_2)}{6 (h_1 \rho_2 + h_2 \rho_1)^2} \sigma \sigma_{xx} + \frac{1}{2} h^2 g (\rho_2 - \rho_1) \zeta^2. \end{aligned} \tag{3.20}$$

It is convenient to introduce the non-dimensional momentum shear σ^* by

$$\sigma = \sqrt{hg(\rho_2 - \rho_1)(h_1\rho_2 + h_2\rho_1)} \sigma^* \tag{3.21}$$

so that the non-dimensional form E^* of the total energy is (immediately dropping asterisks for ease of notation)

$$\begin{aligned} E &= \frac{1}{2}h_1h_2\sigma^2 + \frac{\alpha}{2}\frac{h_1^2\rho_2 - h_2^2\rho_1}{(h_1\rho_2 + h_2\rho_1)}\zeta\sigma^2 - \frac{\alpha^2}{2}\frac{h^2\rho_1\rho_2}{(h_1\rho_2 + h_2\rho_1)^2}\zeta^2\sigma^2 \\ &+ \frac{\epsilon^2}{6}\frac{h_1^2h_2^2(h_1\rho_1 + h_2\rho_2)}{h_1\rho_2 + h_2\rho_1}\sigma\sigma_{xx} + \frac{1}{2}\zeta^2 \\ &= \frac{1}{2}(A\sigma^2 + \alpha B\zeta\sigma^2 - \alpha^2 C\zeta^2\sigma^2 + \zeta^2 + \epsilon^2\kappa\sigma\sigma_{xx}) \end{aligned} \tag{3.22}$$

where we denoted the constants by

$$A = h_1h_2, \quad B = \frac{h_1^2\rho_2 - h_2^2\rho_1}{(h_1\rho_2 + h_2\rho_1)}, \quad C = \frac{h^2\rho_1\rho_2}{(h_1\rho_2 + h_2\rho_1)^2}, \quad \kappa = \frac{1}{3}\frac{h_1^2h_2^2(h_1\rho_1 + h_2\rho_2)}{h_1\rho_2 + h_2\rho_1}. \tag{3.23}$$

Applying the Poisson tensor (2.32) to the variational differential of the energy (in fact, the Hamiltonian) $\mathcal{E} = \int E dx$ yields the equations of motion as

$$\begin{pmatrix} \zeta_t \\ \sigma_t \end{pmatrix} = - \begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix} \begin{pmatrix} \frac{\delta \mathcal{E}}{\delta \zeta} \\ \frac{\delta \mathcal{E}}{\delta \sigma} \end{pmatrix} \tag{3.24}$$

where t is the non-dimensional time, related with the physical time by

$$t \rightarrow \epsilon \sqrt{\frac{g(\rho_2 - \rho_1)}{h(h_1\rho_2 + h_2\rho_1)}} t. \tag{3.25}$$

This shows explicitly how the evolution proceeds in a slow time gauged by the dispersion parameter ϵ as required by the long-wave asymptotics.

The resulting system in conservation form is

$$\begin{cases} \zeta_t + (A\sigma + \alpha B\zeta\sigma - \alpha^2 C\zeta^2\sigma + \epsilon^2\kappa\sigma_{xx})_x = 0 \\ \sigma_t + \left(\zeta + \alpha\frac{B\sigma^2}{2} - \alpha^2 C\zeta\sigma^2\right)_x = 0 \end{cases}, \tag{3.26}$$

or, carrying out the relevant spatial differentiations explicitly,

$$\begin{cases} \zeta_t + A\sigma_x + \alpha B(\zeta\sigma)_x - \alpha^2 C(\zeta^2\sigma)_x + \epsilon^2\kappa\sigma_{xxx} = 0 \\ \sigma_t + \zeta_x + \alpha B\sigma\sigma_x - \alpha^2 C(\zeta\sigma^2)_x = 0 \end{cases}. \tag{3.27}$$

which from now on will be referred to as the *ABC-system*.

Remark 3.1. A few comments on the parameters A, B, C and their relations with the physical parameters ρ_1, ρ_1, h_1, h_2 are in order. First, the parameter A is just the square of the linear wave velocity; in nondimensional form it ranges from 0 to $\frac{1}{4}$, and could be set to unity by further rescaling σ . Next, note that the parameter κ is nonnegative, and vanishes only when $h_1 \rightarrow 0$ or $h_2 \rightarrow 0$. Similarly, the parameter C is nonnegative, and vanishes only in the air-water limit

$\rho_1 \rightarrow 0$. The most interesting parameter is B , which is non sign-definite and appears in front of the cubic term $\sigma^2 \zeta$ of the Hamiltonian. It vanishes at the critical ratio

$$\frac{\rho_1}{\rho_2} = \frac{h_1^2}{h_2^2}. \tag{3.28}$$

By denoting the density ratio parameter $r = \rho_1/\rho_2$, so that $0 < r < 1$, the definition of B shows that it is positive for $h_1 > \sqrt{r}h_2$, and negative for $h_1 < \sqrt{r}h_2$. One of the most relevant effects of this change in sign shows up in the existence and polarity of solitary travelling wave solutions of system (3.26), as we shall see below in section 5. In this regard, to account for the break down of the theory near the vanishing of the B coefficient of quadratic nonlinearity, we notice that under the WNL asymptotic scaling it would be necessary to compute a plethora of higher order terms for asymptotic consistency, as shown in [32]. Such terms involve higher order derivatives which make the search for travelling solutions a (mostly) numerical affair, whereas reasonable qualitative and somewhat quantitative agreement with Euler solutions can already be obtained under the present ABC model.

Remark 3.2. The Boussinesq approximation consists of retaining density differences in the potential (gravitational) energy density, while neglecting the associated inertial differences in the kinetic energy density, by setting in (3.6)

$$\rho_1 = \rho_2 = \bar{\rho}. \tag{3.29}$$

This Boussinesq approximation simplifies significantly the WNL or MNL asymptotics; indeed, the Hamiltonian variable for the weighted shear reduces to $\sigma = \bar{\rho}(\tilde{u}_2 - \tilde{u}_1)$, and the non-dimensional energy density of the system becomes

$$\begin{aligned} E_B &= \frac{1}{2}h_1 h_2 \sigma^2 + \frac{\alpha}{2}(h_1 - h_2)\zeta \sigma^2 - \frac{\alpha^2}{2}\zeta^2 \sigma^2 + \frac{\epsilon^2}{6}h_1^2 h_2^2 \sigma \sigma_{xx} + \frac{1}{2}\zeta^2 \\ &= \frac{1}{2}(A_B \sigma^2 + \alpha B_B \zeta \sigma^2 - \alpha^2 C_B \zeta^2 \sigma^2 + \zeta^2 + \epsilon^2 \kappa_B \sigma \sigma_{xx}), \end{aligned} \tag{3.30}$$

where

$$A_B = h_1 h_2, \quad B_B = h_1 - h_2, \quad C_B = 1, \quad \kappa_B = \frac{1}{3}h_1^2 h_2^2. \tag{3.31}$$

The Hamiltonian formulation of system (3.26) provides three additional constants of the motion besides the energy \mathcal{E} . They are the two Casimir functionals,

$$\mathcal{K}_1 = \int_{\mathbb{R}} \zeta \, dx, \quad \mathcal{K}_2 = \int_{\mathbb{R}} \sigma \, dx, \tag{3.32}$$

and the generator of the x -translation

$$\Pi = \int_{\mathbb{R}} \zeta \sigma \, dx. \tag{3.33}$$

They are conserved quantities for any choice of the parameters A, B and C . As we shall see in the following section, the WNL case is rather special, as it reduces to the so-called completely integrable Kaup–Boussinesq systems.

4. Two notable reductions and their complete integrability

It is worth considering further simplifications of the reduction (3.24) as they may be applicable to certain physical regimes and offer the unexpected bonus of being completely integrable. We first look at the WNL limit (WNL) in the context of the bidirectional system (3.26). We then

examine how the Hamiltonian reduction strategy can be used to derive unidirectional motion equations.

4.1. The WNL case and the Kaup–Boussinesq system

The WNL asymptotics mentioned above, where $\alpha = O(\epsilon^2)$, formally corresponds to dropping the C -term in the Hamiltonian, as the quartic terms of $O(\alpha^2)$ becomes subdominant with respect to other terms unless the ‘hardware’ parameters (depths and densities) are near the critical ratio (3.28), where the coefficient of the cubic term B vanishes. Away from the critical ratio, the WNL case leads to a ‘universal’ representative bidirectional system which can be viewed as standing at the same level as its unidirectional counterpart, the well known KdV equation.

Suppressing the order parameters α and ϵ for ease of notation, system (3.27) in the WNL limit becomes

$$\begin{cases} \zeta_t + A\sigma_x + B(\zeta\sigma)_x + \kappa\sigma_{xxx} = 0 \\ \sigma_t + \zeta_x + B\sigma\sigma_x = 0 \end{cases} \tag{4.1}$$

This system is a parametric version of the Boussinesq system for water waves equations, introduced by [5, 37]. It is asymptotically equivalent, up to terms of order $O(\alpha, \epsilon^2)$, to the nonlocal one reported in [12] (the form apparently favoured by Boussinesq [37], section 13.11), through the change of variables

$$\bar{\sigma} = \sigma + \kappa\sigma_{xx}, \quad \text{and} \quad \bar{\sigma} = \left(\frac{\rho_1 h_2 + \rho_2 h_1}{h_2} - \frac{\rho_2 \zeta}{h_2^2} \right) \bar{u}_1. \tag{4.2}$$

System (4.1) is completely integrable via the Inverse Scattering Method, as shown in [26], and further analysed in [27], where it was also shown how it can be derived in bi-Hamiltonian form. In our variables (which are related to those in [27] by a nontrivial Miura-like transformation) the corresponding Poisson pencil (see, e.g. [22, 24]) is

$$P(\lambda) = \lambda P_0 - P_1 = \begin{pmatrix} \frac{1}{2}B(\zeta\partial_x + \partial_x\zeta) + A\partial_x + \kappa\partial_x^3 & (\frac{1}{2}B\sigma - \lambda)\partial_x \\ \partial_x(\frac{1}{2}B\sigma - \lambda) & \partial_x \end{pmatrix}. \tag{4.3}$$

Indeed, the equations of motion (4.1) can be written as the Hamiltonian evolution

$$\begin{pmatrix} \zeta_t \\ \sigma_t \end{pmatrix} = P_1 \begin{pmatrix} \frac{\delta\Pi}{\delta\zeta} \\ \frac{\delta\Pi}{\delta\sigma} \end{pmatrix} = P_0 \begin{pmatrix} \frac{\delta\mathcal{E}}{\delta\zeta} \\ \frac{\delta\mathcal{E}}{\delta\sigma} \end{pmatrix}. \tag{4.4}$$

Throughout this section and the next one, the differential operator ∂_x is, as usual, meant to act on all quantities that stand to its right, e.g. $\partial_x \zeta\phi = (\zeta\phi)_x$. Also, in the above formula, Π is the generator of x -translations (3.33) and we renamed P_0 the tensor J_{red} of (2.32).

This bi-Hamiltonian formulation can be used to construct recursively an infinite family of constants of motion. We briefly review here the technique in [24], adapted to system (4.1). First, we seek the Casimir of the Poisson pencil (4.3), in the form of a series $\mathcal{H}(\lambda)$ in inverse powers of λ , $\mathcal{H}(\lambda) = \sum_{n=0}^{\infty} \mathcal{H}_n \lambda^{-n}$, whose variational gradient satisfies

$$P(\lambda) \cdot d\mathcal{H}(\lambda) = 0. \tag{4.5}$$

Denoting by (γ, β) the components of the gradient of $\mathcal{H}(\lambda)$ one gets the following system

$$\begin{cases} \frac{B}{2}(\zeta\gamma_x + (\zeta\gamma)_x) + A\gamma_x + \kappa\gamma_{xxx} + (\frac{B}{2}\sigma - \lambda)\beta_x = 0 \\ \frac{B}{2}(\sigma\gamma)_x - \lambda\gamma_x + \beta_x = 0 \end{cases} \tag{4.6}$$

Substituting $\beta_x = \lambda\gamma_x - B(\sigma\gamma)_x/2$ from the second equation into the first yields an expression for γ that can be manipulated, by multiplying it by γ , into the total x -derivative.

$$\begin{aligned} \frac{B}{2}[(\lambda\zeta\gamma_x + \gamma(\lambda\zeta)_x)] - \frac{B^2}{4}(\gamma\sigma)(\gamma\sigma)_x + \lambda\frac{B}{2}[(\gamma\sigma)\gamma_x + \gamma(\gamma\sigma)_x] \\ + (A - \lambda^2)\gamma\gamma_x + \kappa\gamma\gamma_{xxx} = 0. \end{aligned} \tag{4.7}$$

By integrating in x , system (4.6) can be replaced by

$$\begin{cases} \frac{1}{2}(-\lambda^2 + A + B(\lambda\sigma + \zeta - \frac{B}{4}\sigma^2))\gamma^2 + \kappa(\gamma\gamma_{xx} - \frac{1}{2}\gamma_x^2) = F(\lambda) \\ \frac{B}{2}\sigma\gamma - \lambda\gamma + \beta = G(\lambda) \end{cases}, \tag{4.8}$$

where $F(\lambda)$ and $G(\lambda)$ are the arbitrary constants of integration with respect to x . The corresponding inverse power series for $\gamma = 1 + O(\frac{1}{\lambda})$ and $\beta = O(\frac{1}{\lambda})$ can be obtained by setting

$$F(\lambda) = -\frac{\lambda^2}{2}, \quad G(\lambda) = -\lambda. \tag{4.9}$$

It is straightforward to check that with this choice system (4.8) can be solved iteratively. It remains to show that the one-form (γ, β) is exact. To this end we define

$$h(\lambda) := \frac{\lambda}{2\sqrt{\kappa}} \frac{1}{\gamma} + \frac{\gamma_x}{2\gamma} = \frac{\lambda}{2\sqrt{\kappa}} + h_0 + \frac{h_1}{\lambda} + \frac{h_2}{\lambda^2} + \frac{h_3}{\lambda^3} \dots \tag{4.10}$$

In terms of $h(\lambda)$ we can write (4.8), subject to the choice (4.9), as

$$\begin{cases} h(\lambda)_x + h(\lambda)^2 = \frac{1}{4\kappa} \left(\lambda^2 - A - B \left(\lambda\sigma + \zeta - \frac{B\sigma^2}{4} \right) \right) \\ \beta(\lambda) = \lambda(\gamma - 1) - \frac{B}{2}\sigma\gamma. \end{cases} \tag{4.11}$$

Let us consider the one-form (γ, β_γ) with β_γ given by the second equation of this system, and denote by $(\dot{\zeta}, \dot{\sigma})$ the tangent vector to a generic curve in the phase space (ζ, σ) . Then

$$\int_{\mathbb{R}} (\gamma\dot{\zeta} + \beta_\gamma\dot{\sigma}) dx = \int_{\mathbb{R}} \left(\gamma\dot{\zeta} + \left(\lambda(\gamma - 1) - \frac{B}{2}\gamma\sigma \right) \dot{\sigma} \right) dx. \tag{4.12}$$

From the first of (4.11) we get

$$\dot{h}_x + 2h\dot{h} = -\frac{B}{4\kappa} \left(\dot{\zeta} + \lambda\dot{\sigma} - \frac{B}{2}\sigma\dot{\sigma} \right). \tag{4.13}$$

Multiplying by γ , integrating by parts, and using the definition (4.10) finally yields

$$\int_{\mathbb{R}} (\gamma\dot{\zeta} + \beta_\gamma\dot{\sigma}) dx = -\lambda \frac{d}{dt} \int_{\mathbb{R}} \left(\sigma + \frac{4\sqrt{\kappa}}{B} h(\lambda) \right) dx. \tag{4.14}$$

We can conclude that

$$\mathcal{H}(\lambda) = - \int_{\mathbb{R}} \left(\sigma + \frac{4\sqrt{\kappa}}{B} h(\lambda) \right) dx \tag{4.15}$$

is a Casimir of the Poisson pencil (4.3); hence the coefficients of its expansion in inverse powers of λ are mutually commuting constants of the motion. The first conserved quantities are

$$\begin{aligned}
 \mathcal{H}_1 &= \int_{\mathbb{R}} \zeta \, dx, \\
 \mathcal{H}_2 &= \frac{B}{2} \int_{\mathbb{R}} \zeta \sigma \, dx, \\
 \mathcal{H}_3 &= \frac{B}{2} \int_{\mathbb{R}} \left(\frac{1}{2} A \sigma^2 + \frac{1}{2} B \zeta \sigma^2 + \frac{\zeta^2}{2} - \frac{1}{2} \kappa \sigma_x^2 \right) dx, \\
 \mathcal{H}_4 &= \frac{B^2}{8} \int_{\mathbb{R}} \left(A \sigma^3 + B \zeta \sigma^3 + 3 \zeta^2 \sigma - 3 \kappa \sigma \sigma_x^2 - \frac{4}{B} \kappa \zeta_x \sigma_x \right) dx.
 \end{aligned}
 \tag{4.16}$$

Note that \mathcal{H}_1 is a Casimir of P_0 , the quantity \mathcal{H}_2 was already identified with the generator of x -translations, while \mathcal{H}_3 is (up to a factor $1/2$) the energy (i.e. the Hamiltonian functional (3.22) for P_0). Together with $\int_{\mathbb{R}} \sigma \, dx$, the first three conserved quantities come from basic physical principles. The fourth, \mathcal{H}_4 , and all the higher order \mathcal{H} 's thus constructed are the conserved quantity more directly associated with the Liouville integrability of the mathematical problem and the bi-Hamiltonian formulation we have described. It is well known (see, e.g. [5]), that the energy \mathcal{H}_3 failing to be a positive-definite quantity implies that the corresponding equations of motion are not ‘well protected against short wave instability,’ the so-called *bad* Boussinesq equation being possibly the prototypical example of an integrable equation displaying such a drawback. Further comments on this phenomenon can be found in section 5.1.

We remark that the full ABC system (3.26), unlike its WNL reduced case, seems to fail the complete integrability property of a second local Hamiltonian structure. Following the WNL structure, one could make use of the conserved quantity (proportional to) \mathcal{H}_2 above to provide such a structure with the anti-symmetric operator

$$P_{ABC} = \begin{pmatrix} -\frac{1}{2} B (\zeta \partial_x + \partial_x \zeta) + A \partial_x + \kappa \partial_x^3 & -\frac{1}{2} B \sigma \partial_x + C \partial_x \sigma \zeta \\ -\frac{1}{2} \partial_x B \sigma + C \sigma \zeta \partial_x & \partial_x + C \sigma \partial_x \sigma \end{pmatrix}.
 \tag{4.17}$$

Used with the appropriate factor of \mathcal{H}_2 , this operator does yield the equations of motion (3.26); however, because its dispersionless limit is not associated with a flat metric, as detailed in [21], P_{ABC} fails to satisfy a necessary condition for fulfilling Jacobi identity, and hence cannot be used to generate a second Hamiltonian structure for system (3.26).

4.2. Unidirectional models

To obtain unidirectional nonlinear wave equations for our model, we at first observe that the rescaling $\sigma \rightarrow \sqrt{A} \sigma$ simplifies the Hamiltonian density (3.22) to

$$\tilde{\mathcal{H}} = \frac{1}{2} \left(\sigma^2 + \zeta^2 + \alpha \tilde{B} \zeta \sigma^2 - \alpha^2 \tilde{C} \zeta^2 \sigma^2 + \epsilon^2 \tilde{\kappa} \sigma \sigma_{xx} \right)
 \tag{4.18}$$

(with $\tilde{B} = \frac{B}{A}$ and so on and so forth), and the ensuing Hamiltonian equations of motion to

$$\begin{cases} \zeta_t = -(\sigma_x + \alpha \tilde{B} (\zeta \sigma)_x - \alpha^2 \tilde{C} (\zeta^2 \sigma)_x + \epsilon^2 \tilde{\kappa} \sigma_{xxx}) \\ \sigma_t = -(\zeta_x + \alpha \tilde{B} \sigma \sigma_x - \alpha^2 \tilde{C} (\zeta \sigma^2)_x) \end{cases}.
 \tag{4.19}$$

We seek (following, e.g. the classical steps of [37] section 13) for a relation $\sigma = \sigma(\zeta)$ of the form

$$\sigma = \zeta + \alpha F(\zeta) + \alpha^2 G(\zeta) + \epsilon^2 K(\zeta) \tag{4.20}$$

with F, G, K differential polynomials in ζ such that the resulting equations obtained substituting (4.20) in (4.19) coincide up to terms vanishing faster than α^2 and ϵ^2 in the limit $\alpha, \epsilon \rightarrow 0$, that is, at order $O(\alpha^2, \epsilon^2)$. This procedure can be carried on in a straightforward manner, the only difference with the derivation of the KdV equation of [37] being that at order $O(\alpha)$ one has to use the relation

$$\partial_t = -\partial_x - \frac{3}{2} \tilde{B} \alpha \zeta \partial_x. \tag{4.21}$$

The outcome is the following:

- i) The link between ζ and σ of equation (4.20) is explicitly given by

$$\sigma = \zeta - \frac{1}{4} \alpha \tilde{B} \zeta^2 + \frac{1}{8} \alpha^2 \tilde{B}^2 \zeta^3 - \frac{1}{2} \epsilon^2 \tilde{\kappa} \zeta_{xx}. \tag{4.22}$$

- ii) The resulting unidirectional equation of motion is a parametric form of the (defocusing) Gardner (or KdV-mKdV) equation

$$\zeta_t = -\zeta_x - \frac{3}{2} \alpha \tilde{B} \zeta \zeta_x + \left(3 \alpha^2 \tilde{C} + \frac{3}{8} \alpha^2 \tilde{B}^2 \right) \zeta^2 \zeta_x - \frac{1}{2} \epsilon^2 \tilde{\kappa} \zeta_{xxx}, \tag{4.23}$$

which was derived in the theory of stratified fluids, e.g. in [20].

We first notice that, in the WNL approximation, that is at $O(\alpha, \epsilon^2)$ with $\alpha = O(\epsilon^2)$, equation (4.23) becomes the Korteweg-deVries (KdV) equation, and relation (4.22) reduces to the one of [37] section 13. Moreover, although for $\tilde{C} = 0$ the Hamiltonian (4.18) becomes the Hamiltonian of the Kaup–Boussinesq system (4.1), the resulting unidirectional equation for ζ has a modified KdV (mKdV) term, given by $3 \alpha^2 \tilde{B}^2 \zeta^2 \zeta_x / 8$.

These unidirectional equations can be given a Hamiltonian interpretation by providing an alternative strategy by a geometric reshaping of the argument in [33, 34] (which refers to a single layer Euler fluid, an inessential difference in this context). We regard equation (4.22) as an *asymptotic constraint* between the two dependent variables σ, ζ and we apply the Dirac theory of constraints [19], and its related Dirac Poisson brackets. First, a straightforward computation shows that, still in the $O(\epsilon^2, \alpha^2)$ asymptotics, no secondary constraint arise, that is, if we denote by

$$\Phi \equiv \sigma - \zeta + \frac{1}{4} \alpha \tilde{B} \zeta^2 - \frac{1}{8} \alpha^2 \tilde{B}^2 \zeta^3 + \frac{1}{2} \epsilon^2 \tilde{\kappa} \zeta_{xx} = 0 \tag{4.24}$$

the constraint, the equations of motion, the constraint equation (4.22) and relation (4.21) imply

$$\Phi_t \approx 0 \text{ at } O(\alpha^2, \epsilon^2), \tag{4.25}$$

where the ‘ \approx ’ symbol, as per the usual Dirac’s theory notation, stands for equality on the constrained manifold. Second, we notice that the pair $\zeta, \varphi = \sigma - g(\zeta)$, where $g(\zeta) = \zeta - \frac{1}{4} \alpha \tilde{B} \zeta^2 + \frac{1}{8} \alpha^2 \tilde{B}^2 \zeta^3 - \frac{1}{2} \epsilon^2 \tilde{\kappa} \zeta_{xx}$, is a set of coordinates equivalent to the pair ζ, σ , and we express the Poisson tensor (2.32) in these new coordinates. The result is the matrix of differential operators

$$\tilde{P} = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & \partial_x \cdot g'(\zeta) + g'(\zeta) \cdot \partial_x \end{pmatrix}, \tag{4.26}$$

where we denoted by $g'(\zeta)$ the Fréchet derivative of $g(\zeta)$, viz.

$$g' = 1 - \frac{1}{2}\tilde{B}\alpha\zeta + \frac{3}{8}\tilde{B}^2\alpha^2\zeta^2 - \frac{1}{2}\epsilon^2\tilde{\kappa}\partial_{xx}. \tag{4.27}$$

In analogy with the usual formula of the Dirac Poisson brackets for the finite N -dimensional case q_1, \dots, q_N , with a number $M < N$ of constraints Φ_1, \dots, Φ_M ,

$$\{q_i, q_j\}^D = \{q_i, q_j\} - \sum_{a,b=1}^M \{q_i, \Phi_a\}(\mathcal{C}^{-1})_{ab}\{\Phi_b, q_j\}, \tag{4.28}$$

where \mathcal{C} is the matrix with entries $\{\Phi_a, \Phi_b\}$, the Dirac tensor in the coordinates (ζ, φ) is given by

$$P^D \equiv \begin{pmatrix} \tilde{P}_{11} - \tilde{P}_{12}(\tilde{P}_{22})^{-1}\tilde{P}_{21} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -\partial_x(\tilde{P}_{22})^{-1}\partial_x & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.29}$$

with $\tilde{P}_{22} = \partial_x \cdot g'(\zeta) + g'(\zeta) \cdot \partial_x$. This yields the reduced Dirac Poisson tensor on the ‘constrained’ manifold of unidirectional right-moving waves,

$$P_R^D \equiv -\partial_x(\tilde{P}_{22})^{-1}\partial_x. \tag{4.30}$$

Our final task is to compute, still in the MNL asymptotics, the inverse of the operator \tilde{P}_{22} .

A direct computation in this asymptotics shows that such an inverse is given by the pseudo-differential operator

$$(\tilde{P}_{22})^{-1} = \frac{1}{4} \left(2\partial_x^{-1} + \frac{1}{2}\tilde{B}\alpha(\partial_x^{-1}\zeta + \zeta\partial_x^{-1}) - \frac{1}{8}\tilde{B}^2\alpha^2(2\partial_x^{-1}\zeta^2 + 2\zeta^2\partial_x^{-1} - \partial_x^{-1}\zeta\partial_x\zeta\partial_x^{-1} - \zeta\partial_x^{-1}\zeta) + \frac{1}{2}\epsilon^2\tilde{\kappa}\partial_x \right), \tag{4.31}$$

which yields, after some manipulation, the reduced Dirac tensor

$$P_R^D = -\frac{1}{2}\partial_x - \frac{1}{8}\alpha\tilde{B}(\zeta\partial_x + \partial_x\zeta) + \frac{1}{32}\alpha^2\tilde{B}^2(\zeta^2\partial_x + \partial_x\zeta^2 + \zeta_x\partial_x^{-1}\zeta_x) - \frac{1}{4}\epsilon^2\tilde{\kappa}\partial_{xxx}. \tag{4.32}$$

The Hamiltonian density reduces, on the constrained manifold $\sigma = \zeta - \frac{1}{4}\alpha\tilde{B}\zeta^2 + \frac{1}{8}\alpha^2\tilde{B}^2\zeta^3 - \frac{1}{2}\epsilon^2\tilde{\kappa}\zeta_{xx}$ and in the MNL asymptotics, to

$$\tilde{H}^D = \zeta^2 + \frac{1}{4}\alpha\tilde{B}\zeta^3 - \alpha^2 \left(\frac{3}{32}\tilde{B}^2 + \frac{1}{2}\tilde{C} \right) \zeta^4. \tag{4.33}$$

Finally, it is easy to verify that the combination

$$P_R^D \frac{\delta}{\delta \zeta} \int_{\mathbb{R}} \tilde{H}^D dx \tag{4.34}$$

yields the unidirectional equation of motion (4.23).

It is remarkable that, in the WNL asymptotics $\alpha = O(\epsilon^2)$, the operator P_R^D yields the bi-Hamiltonian structure for the KdV equation. Indeed,

$$P_{R,KdV}^D = \frac{1}{2}\partial_x - \frac{1}{8}\alpha\tilde{B}(\zeta\partial_x + \partial_x\zeta) - \frac{1}{2}\epsilon^2\tilde{\kappa}\partial_{xxx} \tag{4.35}$$

can be written, after suitable rescaling of the variables, as

$$P_{R,KdV}^D = \partial_x - \epsilon^2(\partial_{xxx} + \zeta\partial_x + \partial_x\zeta), \tag{4.36}$$

which is the Magri Poisson pencil for KdV, where the usual role of the pencil parameter is here played by the square of the inverse dispersion parameter ϵ^{-2} .

5. Special solutions

In this Section, we investigate some properties of the motion equation (3.26) which are relevant to their actual applicability as models of wave propagation, viz. their dispersive behaviour and their traveling wave solutions. This is a first step, which can be carried out without resorting to numerical methods, necessary to assess the performance of the models we have derived with respect to established results for the parent Euler equations.

5.1. Linearization and the dispersion relation

Linearizing system (3.26) around the constant solution (Z, S) , $\zeta = Z + z(x, t)$ and $\sigma = S + s(x, t)$ say, with the functions z, s treated as infinitesimal, yields

$$\begin{cases} z_t + As_x + \alpha B(Zs_x + Sz_x) - \alpha^2 C(Z^2 s_x + 2SZs_x) + \epsilon^2 \kappa s_{xxx} = 0 \\ s_t + z_x + \alpha BSs_x - \alpha^2 C(S^2 z_x + 2SZs_x) = 0 \end{cases} \quad (5.1)$$

Looking for sinusoidal wave solutions of the form $(z, s) = (a_z, a_s)e^{i(kx - \omega t)}$ leads to the following algebraic eigenvalue problem for the phase speed $c_p \equiv \omega/k$ as eigenvalue,

$$\begin{bmatrix} \alpha BS - 2\alpha^2 CSZ - c_p & A + \alpha BZ - \alpha^2 CZ^2 - \epsilon^2 \kappa k^2 \\ 1 - \alpha^2 CS^2 & \alpha BS - 2\alpha^2 CZS - c_p \end{bmatrix} \begin{bmatrix} a_z \\ a_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (5.2)$$

giving the following dispersion relation

$$c_p = \alpha(BS - 2\alpha CSZ) \pm \sqrt{(1 - \alpha^2 CS^2)(A + \alpha BZ - \alpha^2 CZ^2 - \epsilon^2 \kappa k^2)}, \quad (5.3)$$

whereby the critical threshold wavenumber

$$k_c^2 = (A + \alpha BZ - \alpha^2 CZ^2)/(\epsilon^2 \kappa) \quad (5.4)$$

is identified, past which the system becomes Hadamard ill-posed with $k > k_c$. Note that the factor $1 - \alpha^2 CS^2$ needs to be positive, for stability of long waves, i.e. $k \ll 1$, where the asymptotic model applies. This puts a bound on the admissible values of the equilibrium momentum shear S at the critical values $S = \pm 1/(\alpha\sqrt{C})$, with $-1/(\alpha\sqrt{C}) < S < 1/(\alpha\sqrt{C})$. Such a bound, and in particular the fact that the threshold wavenumber for (3.27) is independent of the magnitude of the shear σ , limits the applicability of this system in possible numerical applications. However, it is inherent to our local Hamiltonian setting and will consistently recur in what follows, e.g. in the definitions of the domain of hyperbolicity of the dispersionless limit, as well as in the analysis of the travelling wave solutions of the dispersive case.

With regard to the last point, it is worth noting that an asymptotic step, akin to the well know KdV to BBM near identity relation, can here be used advantageously to circumvent the hindrance to numerical applications due the lack of well posedness of (3.27). Indeed, a shift of the dependent variable σ , similar to (4.2),

$$\bar{\sigma} \equiv \sigma + \epsilon^2 \bar{\kappa} \sigma_{xx} \implies \sigma = \bar{\sigma} - \epsilon^2 \bar{\kappa} \bar{\sigma}_{xx} + O(\epsilon^4), \quad (5.5)$$

where $\bar{\kappa} = \kappa/A$, takes system (3.27) into the asymptotically equivalent form

$$\begin{cases} \zeta_t + A\bar{\sigma}_x + \alpha B(\zeta\bar{\sigma})_x - \alpha^2 C(\zeta^2\bar{\sigma})_x = 0 \\ \bar{\sigma}_t + \zeta_x + \alpha B\bar{\sigma}\bar{\sigma}_x - \alpha^2 C(\zeta\bar{\sigma}^2)_x = \epsilon^2 \bar{\kappa} \bar{\sigma}_{xxx} \end{cases} \quad (5.6)$$

As hinted by the notation used, this step is equivalent to that of using layer averaged velocities, in defining the density weighted vorticity, instead of the velocities at the interface between layers used in the definition (3.15) of σ . The dispersion relation for system (5.6) linearized around constant states $\bar{\sigma} = S$ and $\zeta = Z$ is readily obtained from modifying (5.2) above:

$$\begin{bmatrix} \alpha BS - 2\alpha^2 CSZ - c_p & A + \alpha BZ - \alpha^2 CZ^2 \\ 1 - \alpha^2 CS^2 & \alpha BS - 2\alpha^2 CZS - c_p(1 + \epsilon^2 \bar{\kappa} k^2) \end{bmatrix} \begin{bmatrix} a_z \\ a_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{5.7}$$

This is more resilient for well-posedness than (5.3). In fact, the eigenvalue c_p is given by the solution of the quadratic equation

$$(1 + \epsilon^2 \bar{\kappa} k^2) c_p^2 - q_1(2 + \epsilon^2 \bar{\kappa} k^2) c_p - q_2 + q_1^2 = 0, \tag{5.8}$$

where we have introduced the shorthand notation

$$q_1 = \alpha BS - 2\alpha^2 CSZ, \quad q_2 = (A + \alpha BZ - \alpha^2 CZ^2)(1 - \alpha^2 CS^2).$$

The asymptotic expansion of the discriminant of (5.8) is

$$\Delta = 4(A + \alpha BZ - \alpha^2 C(AS^2 + Z^2) + A\epsilon^2 \bar{\kappa} k^2) + O(\alpha^3, \alpha \epsilon^2),$$

which is certainly asymptotically positive for values of S and Z of order $O(1)$ with respect to the small α parameter.

The standard dispersion relations for infinitesimal disturbances around the quiescent state $\zeta = \sigma = 0$ are obtained from the above relations setting $Z = S = 0$,

$$c_p^2 = A - \epsilon^2 \kappa k^2, \tag{5.9}$$

and

$$c_p^2 = \frac{A}{1 + \epsilon^2 \bar{\kappa} k^2}, \tag{5.10}$$

for systems (3.27) and (5.6) respectively. The role of the coefficient A as the limiting long wave phase speed, and the different behaviours of these speeds in the large wavenumber k limit, are especially transparent in this case.

Remark 5.1. The analysis of the dispersionless counterpart of the system (3.26) goes as follows (see [9] for the fully nonlinear dispersionless case). The dispersionless Hamiltonian density can be read off equation (3.22),

$$H_d = \frac{1}{2} (A\sigma^2 + \alpha B\zeta\sigma^2 - \alpha^2 C\zeta^2\sigma^2 + \zeta^2). \tag{5.11}$$

Hence, the dispersionless equations can be written as

$$\begin{pmatrix} \zeta_t \\ \sigma_t \end{pmatrix} + \begin{pmatrix} H_{d,\zeta\sigma} & H_{d,\sigma\sigma} \\ H_{d,\zeta\zeta} & H_{d,\zeta\sigma} \end{pmatrix} \begin{pmatrix} \zeta_x \\ \sigma_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.12}$$

The characteristic matrix of the system is, explicitly,

$$V = \begin{pmatrix} \alpha B\sigma - 2\alpha^2 C\sigma\zeta & -\alpha^2 C\zeta^2 + \alpha B\zeta + A \\ 1 - \alpha^2 C\sigma^2 & \alpha B\sigma - 2\alpha^2 C\sigma\zeta \end{pmatrix}, \tag{5.13}$$

and so the characteristic velocities are given by

$$v_{\pm} = \alpha B\sigma - 2\alpha^2 C\sigma\zeta \pm \sqrt{(-\alpha^2 C\zeta^2 + \alpha B\zeta + A)(1 - \alpha^2 C\sigma^2)}. \tag{5.14}$$

The hyperbolicity domain is thus the rectangular region in the hodograph space

$$(\zeta, \sigma) \in \left(\frac{B - \sqrt{4CA + B^2}}{2\alpha C}, \frac{B + \sqrt{4CA + B^2}}{2\alpha C} \right) \times \left(-\frac{1}{\alpha\sqrt{C}}, \frac{1}{\alpha\sqrt{C}} \right). \quad (5.15)$$

Remark 5.2. The regularization (5.5) is different from those used in [4, 13, 28], where the change of variables leading to the shear is done through the choice of a reference height of the horizontal velocities in the layers. The resulting models can be still ill-posed with respect to a dispersion critical wavenumber, however with an optimal choice of the reference height (typically, that of the bottom and top layer) the critical wavenumber can be made to maximize the well posed interval of the dispersion relation.

5.2. Travelling wave solutions and their properties

Travelling waves for the ABC-system,

$$\begin{cases} \zeta_t + (A\sigma + B\zeta\sigma - C\zeta^2\sigma + \kappa\sigma_{xx})_x = 0 \\ \sigma_t + \left(\zeta + \frac{B}{2}\sigma^2 - C\zeta\sigma^2 \right)_x = 0 \end{cases}, \quad (5.16)$$

rewritten here dropping stars and setting $\alpha = \epsilon = 1$ in (3.26), are obtained via the ansatz $\zeta(t, x) = \zeta(x - ct)$, $\sigma(t, x) = \sigma(x - ct)$ as the solution of the system

$$\begin{cases} -c\zeta + A\sigma + B\zeta\sigma - C\zeta^2\sigma + \kappa\sigma_{xx} = K_1 \\ -c\sigma + \zeta + \frac{B}{2}\sigma^2 - C\zeta\sigma^2 = K_2 \end{cases}, \quad (5.17)$$

K_1 and K_2 being integration constants. We limit ourselves to seek solitary wave solutions propagating into a quiescent state, i.e. $\zeta \rightarrow 0$ and $\sigma \rightarrow 0$ as $x \rightarrow \infty$, which sets $K_1 = K_2 = 0$. The second equation in (5.17) yields the relation between ζ and σ ,

$$\zeta = \sigma \frac{c - B\sigma/2}{1 - C\sigma^2}, \quad (5.18)$$

and substituting this into the first of (5.17) provides the quadrature formula

$$\kappa\sigma_x^2 = -\sigma^2 \left(A - \frac{1}{4} \frac{(B\sigma - 2c)^2}{1 - C\sigma^2} \right), \quad (5.19)$$

which can be interpreted as the mechanical analog of particle of mass 2κ in a potential well $U(\sigma)$,

$$U(\sigma) \equiv \sigma^2 \left(A - \frac{1}{4} \frac{(B\sigma - 2c)^2}{1 - C\sigma^2} \right). \quad (5.20)$$

An exact expression for x as a function of σ for the solution of (5.19) can be found in terms of elliptic functions, but it is not particularly illuminating and will not be reported here. Once such an expression is obtained, its counterpart for the displacement ζ follows immediately from (5.18). Typical wave solution profiles are shown in figure 2, where they are compared with those of the same amplitudes for the two-layer model of [12]; this in turn is known to provide good approximations to full Euler solutions in a broad range of physical parameters, including large nonlinearity. As this figure suggests, the differences between the two models become smaller for decreasing amplitudes, in agreement with the mild nonlinearity assumption

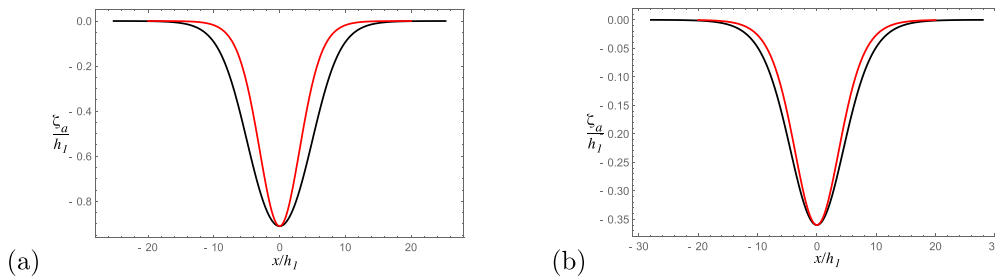


Figure 2. Comparison of internal solitary wave profiles $\zeta(x)$ of different amplitudes: red—solution of (5.16), black—solution of strongly nonlinear model of [12]. Here the physical parameters are $g = 981 \text{ cm s}^{-2}$, $h_1 = 15 \text{ cm}$, $h_2 = 62 \text{ cm}$, $\rho_1 = 0.999 \text{ g cm}^{-3}$, $\rho_2 = 1.022 \text{ g cm}^{-3}$ (used in the experiment of [25]), and the maximum displacement is measured in units of the upper layer thickness h_1 : (a) $\zeta_a = 0.91h_1$; (b) $\zeta_a = 0.36h_1$.

underlying system (5.16). The potential (5.20) has been normalized to have a double zero for $\sigma = 0$, and limits to $-\infty$ when σ tends to $1/\sqrt{C}$ from the right and to $-1/\sqrt{C}$ from the left. Solitary waves—always associated with the null value of the energy of the corresponding mechanical system—exist when $\sigma = 0$ is a local maximum for $U(\sigma)$, and $U(\sigma)$ has two more distinct non-zero roots $\sigma_{1,2}^*$ in the interval $(-1/\sqrt{C}, 1/\sqrt{C})$, so that $-U(\sigma)$ is non-negative for σ between 0 and the smallest (in absolute value) of these additional roots (the limiting case of $\sigma_1^* \rightarrow \sigma_2^*$ corresponds to soliton solutions degenerating to front-like solutions, as the orbit becomes heteroclinic). This implies that (smooth) solitary waves $(\zeta(x - ct), \sigma(x - ct))$ form a one-parameter family with respect to the speed c in the interval

$$A < c^2 < A + \frac{B^2}{4C}. \tag{5.21}$$

Since $c_0 \equiv \sqrt{A}$ can be interpreted as the linearized speed of interfacial long-waves in a two-fluid system, this shows that nonlinear, solitary waves move faster than c_0 up to a limiting maximum speed defined by

$$c_m^2 \equiv c_0^2 + \frac{B^2}{4C}.$$

The dependence on the speed parameter c of the maximum displacements of solitary waves from equilibrium is (taking right-moving waves and $B > 0$ to fix ideas)

$$\sigma_a = \frac{2Bc - 2\sqrt{A(B^2 - 4C(c^2 - A))}}{B^2 + 4AC}, \quad \zeta_a = \frac{-2AB + 2c\sqrt{A(B^2 - 4C(c^2 - A))}}{B^2 - 4c^2C}, \tag{5.22}$$

respectively for σ and ζ . (If $B < 0$, and with right moving waves, the opposite sign of the square roots in the above formulae needs to be taken.) Regardless of the sign of B , at $c = c_m$ these waves degenerate into fronts, with amplitude of displacement given by

$$\sigma_m = \frac{B}{\sqrt{C(B^2 + 4AC)}}, \quad \zeta_m = \frac{B}{2C}. \tag{5.23}$$

Note that the sign of the B coefficient in these relations determines whether the displacement ζ of internal solitary waves is positive (waves of elevation) or negative (waves of depression), for $B > 0$ and $B < 0$, respectively. This sign is in turn determined by the value of the ratio h_1/h_2 with respect to the critical value $\sqrt{\rho_1/\rho_2}$ (see remark 3.1). Notice that for the WNL limit, which, as mentioned above, yields the Kaup–Boussinesq system (4.1) by formally taking $C = 0$, the relation between wave amplitude and speed does not have a limiting extremum (at which the solitary waves degenerate into fronts), and this relation is linear,

$$\zeta_a = 2\sqrt{A} \frac{c - \sqrt{A}}{B}.$$

In terms of the ‘hardware parameters’ h ’s and ρ ’s the maximum speed c_m and displacement amplitude ζ_m read, in the original dimensional variables,

$$c_m^2 = c_0^2 \left(1 + \frac{(\rho_1 h_2^2 - \rho_2 h_1^2)^2}{4(h_1 + h_2)^2 h_1 h_2 \rho_1 \rho_2} \right), \quad \zeta_m = \frac{(\rho_1 h_2 + \rho_2 h_1)(\rho_2 h_1^2 - \rho_1 h_2^2)}{2\rho_1 \rho_2 (h_1 + h_2)^2}, \tag{5.24}$$

where, in dimensional form, the internal long-wave linear speed is given by

$$c_0^2 = g(\rho_2 - \rho_1) \frac{h_1 h_2}{\rho_1 h_2 + \rho_2 h_1}.$$

As expected for the present asymptotic theory, carried out under the assumption of weak non-linearity, these limiting values will in general be different from their exact counterparts of the two-layer Euler system which coincide with the ones obtained with the fully nonlinear model [12]. The latter, c_m^E and ζ_m^E say, are (see, e.g. [12])

$$(c_m^E)^2 = c_0^2 \frac{(h_1 + h_2)(\rho_1 h_2 + \rho_2 h_1)}{h_1 h_2 (\rho_1 + \rho_2 + 2\sqrt{\rho_1 \rho_2})}, \quad \zeta_m^E = \frac{\rho_2 h_1^2 - \rho_1 h_2^2}{\rho_2 h_1 + \rho_1 h_2 + (h_1 + h_2)\sqrt{\rho_1 \rho_2}},$$

and they can be expected to be asymptotically close to c_m and ζ_m as the critical ratio of depths and densities $\frac{h_1}{h_2} = \sqrt{\frac{\rho_1}{\rho_2}}$ is approached. This is due to the fact that the MNL model includes the α^2 -term which is dominant in this regime, since the coefficient of the term of order α vanishes in this limit. Thus, the travelling solitary wave solutions of the present asymptotic model can be expected to provide a good approximation for their exact counterparts in the whole amplitude range, even approaching their front limit, when the depths and densities are such that the critical aspect ratio is approached to within an error of order $O(\alpha)$; more precisely, it can be shown that

$$\frac{h_1}{h_2} - \sqrt{\frac{\rho_1}{\rho_2}} = \alpha \implies \frac{c_m^E - c_m}{c_m^E - c_0} = O(\alpha), \quad \text{and} \quad \frac{\zeta_m - \zeta_m^E}{\zeta_m^E} = O(\alpha), \tag{5.25}$$

in the limit $\alpha \rightarrow 0$, for which $c_m^E, c_m \rightarrow c_0$ and $\zeta_m^E, \zeta_m \rightarrow 0$. These observations are exemplified by figure 3 where the so-called effective wavelength

$$\lambda_l \equiv \frac{1}{\zeta_a} \int_0^{+\infty} \zeta(x) dx$$

is plotted vs. ζ_a for two cases, one corresponding to the hardware parameters used in the experiment in [25] and the other where the depth of the upper layer h_1 is adjusted to be close to the critical ratio as in (5.25) with $\alpha = 0.1$. Figure 3, and its companion figure 4 where the so called nonlinear dispersion relation curve is also depicted, show a comparison with the analogous curves from the strongly nonlinear model [12]. It is remarkable that the limiting

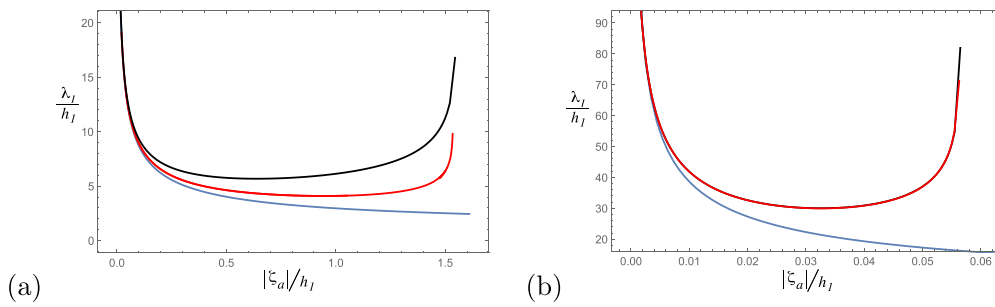


Figure 3. Effective wavelength vs. maximum displacement amplitude for internal solitary waves modelled by (5.16) (red), by (4.1) (blue), and by the strongly nonlinear model of [12] (black). (a) $h_1 = 15$ cm, $h_2 = 62$ cm, $g = 981$ cm s⁻², $\rho_1 = 0.999$ g cm⁻³, $\rho_2 = 1.022$ g cm⁻³. (b) $h_1 = 55.1$ cm, all other parameters as in (a). The axes are scaled with respect to the top layer thickness h_1 . Note the different axis range in the two cases: for (b) the amplitude range is much smaller and the waves are generally longer. Also note how in both cases the location of the vertical asymptote at the front value ζ_m is captured reasonably well by model (5.16), while the Kaup–Boussinesq system (4.1) lacks the front limiting case. As expected, for case (a) which is far from the critical ratio, the intermediate amplitudes are off, while for the near-critical ratio case (b) the black curve is graphically indistinguishable from the red one.

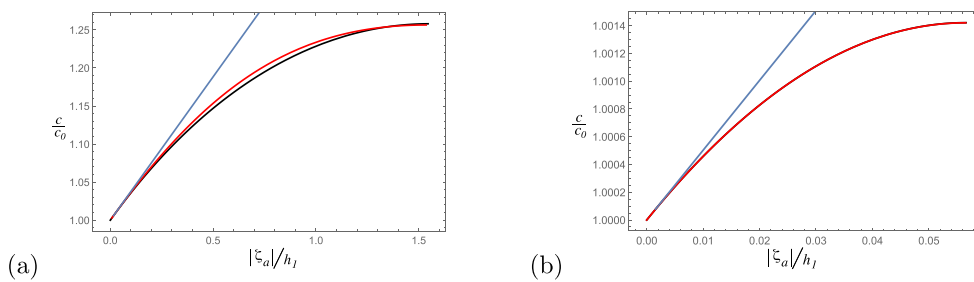


Figure 4. Same as figure 3, but for speed of internal solitary waves vs. maximum displacement amplitude modelled by (5.16) (red), by (4.1) (blue), and by the strongly nonlinear model of [12] (black). Note that the discrepancy between (5.16) and the strongly nonlinear model [12] is far less pronounced for speed than its effective wavelength counterpart, while the Kaup–Boussinesq (4.1) model is only asymptotically valid for small amplitudes in all cases. As for figure 3, notice the different range of case (a) with respect to (b), and the undistinguishable overlapping between the red and the black curves.

values where solitary waves degenerate into fronts are somewhat accurately captured by the model even when these values fall well beyond the model’s asymptotic validity. It is remarkable that this agreement occurs without recourse to an ad-hoc adjustment of the coefficients of the various term in the model, as done in [14] in the context of a unidirectional reduction. For the example provided in figure 4(a), for instance, the limiting values are, respectively for the model and Euler systems, $c_m/c_0 = 1.2565$ and $c_m^E/c_0 = 1.2580$, and $\zeta_m/h_1 = -1.5361$ and $\zeta_m^E/h_1 = -1.5521$. Also, the nonlinear dispersion relation curve representing the wave velocity dependence on amplitude, $c = f(\zeta_a)$ with the function f determined by the second equation in (5.22), remains close to that of the strongly nonlinear system throughout the range of admissible displacement amplitudes, as seen in figure 4(a). Once again, notice how all differences

between models become graphically undetectable as the critical ratio is approached, as demonstrated by the (b) panels of these figures.

6. Conclusions and perspectives

We have applied the technique of Hamiltonian reduction in its entirety, including the handling of constraints when present, to derive model equations that inherit their structure from the parent Benjamin Hamiltonian formulation of a density stratified ideal fluid, under the asymptotic scalings of small amplitudes of fluid parcel displacements from their equilibrium positions and under slow variations in their horizontal positions, i.e. long wave approximation and small dispersion.

The resulting main model generalizes to a bidirectional system the properties of the so-called Gardner unidirectional wave propagation equation, by allowing a quartic nonlinear term to enter the equations and provide the necessary nonlinear contribution to set a critical maximum displacement at which travelling solitary wave solutions degenerate into fronts, as well handle wave dynamics in a neighbourhood of critical depth ratio. In contrast to its unidirectional counterpart, the additional nonlinearity within the bidirectional system makes wave properties such as the maximum amplitude ζ_m and the nonlinear dispersion relation $c(\zeta_a)$ close to their exact Euler counterparts, at least for the parameter range we have explored, even though the quartic term is formally subdominant to the other terms in the Hamiltonian with respect to the small asymptotic parameters carried by the coefficients. This is an unexpected feature of the MNL model that would have been difficult to anticipate based on the derivation alone. Of course, based on this metrics of travelling wave solutions the fidelity of the strongly nonlinear model [12] with respect to the parent Euler system remains unmatched. However, we should stress that the reasonable agreement is obtained here with a substantially simpler, local structure of the model. Further, the Hamiltonian reduction techniques also allow for a systematic derivation of completely integrable models and in particular of Magri's bi-Hamiltonian structure [30] of the KdV equation from the parent Euler system, a program that fulfils the goal posed in [33, 34] (for single layer fluids).

Future work will address reductions that are closer to the physical system by retaining higher order nonlinearity and dispersion, as well as remedy the drawbacks of ill-posedness injected by the asymptotic truncations. The subtleties related to the double scaling limits with the two small parameters α and ϵ with respect to the physical hardware parameters densities ρ 's and depths h 's (see, e.g. [3]), deserve further investigation which will be reported elsewhere.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

We thank R Barros, P Lorenzoni, and R Vitolo for useful discussions. Thanks are also due to the anonymous referees, for providing remarks and suggestions which helped improve the presentation, and for suggesting additional references. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant No. 778010 *IPaDEGAN*. We also gratefully acknowledge the auspices of the GNFM section of INdAM, under which part of this work was carried out, and the financial support of the project MMNLP (Mathematical Methods in Nonlinear Physics) of the

INFN. RC thanks the support by the National Science Foundation under Grants RTG DMS-0943851, CMG ARC-1025523, DMS-1009750, DMS-1517879, DMS-1910824, and by the Office of Naval Research under Grants N00014-18-1-2490, N00014-23-1-2478, and DURIP N00014-12-1-0749. R C and M P thank the Department of Mathematics and its Applications of the University of Milano-Bicocca for its hospitality. R C would also like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme HYD2 where work on this paper was completed, with support by EPSRC Grant EP/R014604/1.

ORCID iDs

R Camassa  <https://orcid.org/0000-0002-1607-5950>

G Falqui  <https://orcid.org/0000-0002-4893-9186>

G Ortenzi  <https://orcid.org/0000-0003-2192-6737>

M Pedroni  <https://orcid.org/0000-0002-7358-0945>

T T Vu Ho  <https://orcid.org/0000-0003-4695-1103>

References

- [1] Benjamin T B 1986 On the Boussinesq model for two-dimensional wave motions in heterogeneous fluids *J. Fluid Mech.* **165** 445–74
- [2] Benjamin T B and Bridges T B 1997 Reappraisal of the Kelvin–Helmholtz problem. Part 1. Hamiltonian structure *J. Fluid Mech.* **333** 301–25
- [3] Bona J L, Lannes D and Saut J-C 2008 Asymptotic models for internal waves *J. Math. Pures Appl.* **89** 538–66
- [4] Boonkasame A and Milewski P A 2014 A model for strongly nonlinear long interfacial waves with background shear *Stud. Appl. Math.* **133** 182–213
- [5] Broer L J F 1975 Approximate equations for long water waves *Appl. Sci. Res.* **31** 377–95
- [6] Camassa R, Choi W, Michallet H, Rusan P-O and Sveen J K 2006 On the realm of validity of strongly nonlinear asymptotic approximations for internal waves *J. Fluid Mech.* **549** 1–23
- [7] Camassa R, Chen S, Falqui G, Ortenzi G and Pedroni M 2012 An inertia ‘paradox’ for incompressible stratified Euler fluids *J. Fluid Mech.* **695** 330–40
- [8] Camassa R, Chen S, Falqui G, Ortenzi G and Pedroni M 2013 Effects of inertia and stratification in incompressible ideal fluids: pressure imbalances by rigid confinement *J. Fluid Mech.* **726** 404–38
- [9] Camassa R, Falqui G and Ortenzi G 2017 Two-layer interfacial flows beyond the Boussinesq approximation: a Hamiltonian approach *Nonlinearity* **30** 466–91
- [10] Chesnokov A A, El G A, Gavriluk S L and Pavlov M V 2017 Stability of shear shallow water flows with free surface *SIAM J. Appl. Math.* **77** 1068–87
- [11] Choi W and Camassa R 1996 Weakly nonlinear internal waves in a two-fluid system *J. Fluid Mech.* **313** 83–103
- [12] Choi W and Camassa R 1999 Fully nonlinear internal waves in a two-fluid system *J. Fluid Mech.* **396** 1–36
- [13] Choi W, Barros R and Jo T-C 2009 A regularized model for strongly nonlinear internal solitary waves *J. Fluid Mech.* **629** 73–85
- [14] Choi W, Zhi C and Barros R 2020 High-order unidirectional model with adjusted coefficients for large-amplitude long internal waves *Ocean Model.* **151** 101643
- [15] Chumakova L, Menzaque F E, Milewski P A, Rosales R R, Tabak E G and Turner C V 2009 Shear instability for stratified hydrostatic flows *Commun. Pure Appl. Math.* **62** 183–97
- [16] Chumakova L, Menzaque F E, Milewski P A, Rosales R R, Tabak E G and Turner C V 2009 Stability properties and nonlinear mappings of two and three-layer stratified flows *Stud. Appl. Math.* **122** 123–37
- [17] Craig W, Guyenne P and Kalisch H 2005 Hamiltonian long-wave expansions for free surfaces and interfaces *Commun. Pure Appl. Math.* **58** 1587–641
- [18] Craig W and Sulem C 1993 Numerical simulation of gravity waves *J. Comput. Phys.* **108** 73–83

- [19] Dirac P A M 1950 Generalized Hamiltonian dynamics *Can. J. Math.* **2** 129–48
- [20] Djordjevic V D and Redekopp L G 1978 The fission and disintegration of internal solitary waves moving over two-dimensional topography *J. Phys. Oceanogr.* **8** 1016–24
- [21] Dubrovin B and Novikov S 1989 Hydrodynamics of weakly deformed soliton lattices. Differential geometry and Hamiltonian theory *Russ. Math. Surv.* **44** 35–124
- [22] Dubrovin B and Zhang Y 2010 Normal forms of hierarchies of integrable PDEs, Frobenius manifolds and Gromov–Witten invariants pp 1–189 (arXiv:0108160)
- [23] Duchêne V, Israwi S and Talhouk R 2016 A new class of two-layer Green–Naghdi systems with improved frequency dispersion *Stud. Appl. Math.* **137** 356–415
- [24] Falqui G, Magri F and Pedroni M 1998 BiHamiltonian geometry, Darboux coverings and linearization of the KP hierarchy *Commun. Math. Phys.* **197** 303–24
- [25] Grue J, Jensen A, Rupas P E and Sveen J K 1999 Properties of large-amplitude internal waves *J. Fluid Mech.* **380** 257–78
- [26] Kaup D J 1975 A higher-order water-wave equation and the method for solving it *Prog. Theor. Phys.* **54** 396–408
- [27] Kupershmidt B A 1985 Mathematics of dispersive water-waves *Commun. Math. Phys.* **99** 51–73
- [28] Lannes D and Ming M 2015 The Kelvin–Helmholtz instabilities in two-fluids shallow water models *Hamiltonian Partial Differential Equations and Applications (Fields Institute Communications vol 75)* ed P Guyenne, D Nicholls and C Sulem (Springer)
- [29] Lvov Y V and Tabak E 2001 Hamiltonian formalism and the Garrett–Munk spectrum of internal waves in the ocean *Phys. Rev. Lett.* **87** 168501
- [30] Magri F 1978 A simple model of the integrable hamiltonian equation *J. Math. Phys.* **19** 1156–62
- [31] Marsden J E and Ratiu T 1986 Reduction of poisson manifolds *Lett. Math. Phys.* **11** 161–9
- [32] Nguyen H Y and Dias F 2008 A Boussinesq system for two-way propagation of interfacial waves *Physica D* **237** 2365–89
- [33] Olver P J 1984 Hamiltonian and non-Hamiltonian models in water waves *Trends and Applications of Pure Mathematics to Mechanics (Lecture Notes in Physics vol 195)* ed P G Ciarlet and M Roseau (Springer)
- [34] Olver P J 1984 Hamiltonian perturbation theory and water waves *Fluids and Plasmas: Geometry and Dynamics (Contemporary Mathematics vol 28)* ed J E Marsden (AMS)
- [35] Ovsyannikov L V 1979 Two-layer ‘shallow water’ model *J. Appl. Mech. Tech. Phys.* **20** 127–35
- [36] Percival J R, Cotter C J and Holm D D 2008 A Euler–Poincaré framework for the multilayer Green–Naghdi equations *J. Phys. A* **41** 344018
- [37] Whitham G B 2000 *Linear and Nonlinear Waves* (Wiley)
- [38] Wu T Y 1981 Long waves in ocean and coastal waters *J. Eng. Mech. Div.* **107** 501–22
- [39] Wu T Y 1998 Nonlinear waves and solitons in water *Physica D* **123** 48–63
- [40] Wu T Y 2000 On modeling unsteady fully nonlinear dispersive interfacial waves *Fluid Dynamics at Interfaces* ed W Shyy and R Narayanan (Cambridge University Press)
- [41] Zakharov V E 1968 Stability of periodic waves of finite amplitude on the surface of a deep fluid *J. Appl. Mech. Tech. Phys.* **9** 190–4
- [42] Zakharov V E, Musher S L and Rubenchik A M 1985 Hamiltonian approach to the description of non-linear plasma phenomena *Phys. Rep.* **129** 285–366