Nonlinearity

PAPER • OPEN ACCESS

Two-layer sharply stratified Euler fluids in three dimensions: a Hamiltonian setting

To cite this article: R Camassa *et al* 2026 *Nonlinearity* **39** 085023

IOP Publishing Nonlinearity 39 (2026) 085023 https://doi.org/10.1088/1361-6544/ae9431

Nonlinearity

LONDON MATHEMATICAL SOCIETY EST. 1865 IOP Institute of Physics

Check for updates

OPEN ACCESS

RECEIVED
21 November 2025

REVISED
17 July 2026

ACCEPTED FOR PUBLICATION
3 August 2026

PUBLISHED
18 August 2026

PAPER

Two-layer sharply stratified Euler fluids in three dimensions: a Hamiltonian setting

R Camassa[1], G Falqui[2,5], G Ortenzi[3,6], M Pedroni[4,5] and E Sforza[2,5,*]

[1] University of North Carolina, 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 Matematica 'G.Peano', Università di Torino, Via Carlo Alberto 10, I-10123 Torino, Italy
[4] Dipartimento di Ingegneria Gestionale, dell'Informazione e della Produzione, Università di Bergamo, Viale Marconi 5, I-24044 Dalmine (BG), Italy
[5] INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[6] INFN, Sezione di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy
[*] Author to whom any correspondence should be addressed.

E-mail: e.sforza4@campus.unimib.it

Keywords: hamiltonian reductions, three-dimensional stratified fluids, asymptotic models, internal waves

Mathematics Subject Classification numbers: 37K58, 35Q31, 35Q35, 76D70, 76M45

Recommended by Dr Karima Khusnutdinova

Abstract

Three-dimensional two-layer incompressible Euler fluids are studied from a Hamiltonian perspective. A natural Hamiltonian structure for the effective 2D model described by the interface-value of the field variables is obtained by means of a Hamiltonian reduction process from the 3D Poisson structure. The problem of expressing the fluid's energy in terms of the reduced variables is considered, and it is shown that in the weakly nonlinear approximation our procedure gives rise to a so-called 2D Kaup–Broer–Kupershmidt Boussinesq (KBK-B) model with *'critical'* parameters. A model weakly dependent on one of the two horizontal directions is also discussed, whose unidirectionalization turns out to be the well-known Kadomtsev–Petviashvili (KP) equation. A Dirac-type reduction process of the Hamiltonian structure of the KBK-B model yields a natural Hamiltonian structure for KP *qua* 2+1-dimensional model.

1. Introduction

Density stratification is a crucial component of the dynamics of a myriad fluid phenomena over a wide range of scales, from those pertaining to geophysical applications in the ocean and atmosphere (see e.g. [6, 29]), to much smaller ones such as those that can occur in industrial processes [23, 26]. Under gravity, the displacement of fluid parcels from their neutral buoyancy position often results in internal wave motion, whose mathematical modelling is an active subject of investigation, as the fundamental governing equations, even under the simplifying assumptions of ideal, incompressible fluids, are not directly solvable, and are complex enough to 'hide' some of the basic mechanisms underlying the dynamics.

In this paper, we focus on three-dimensional settings, whereby constant density surfaces, or isopycnals, are viewed as graphs over two horizontal directions orthogonal to gravity, see figure 1. In the literature, these settings have received relatively less attention than their two-dimensional counterparts (see e.g. [3, 10, 13, 14, 17]), especially from an analytical and geometrical viewpoint. Our aim is to study the fundamental mathematical structures of the governing equations, and to use these as the starting point for the derivation of simplified mathematical models in a physically relevant setup. Specifically, the fluid is viewed as inviscid and incompressible, bound by two horizontal plates extending to infinity. Under these conditions, the equations of motion, sometimes referred to as the Euler–Boussinesq system, admit

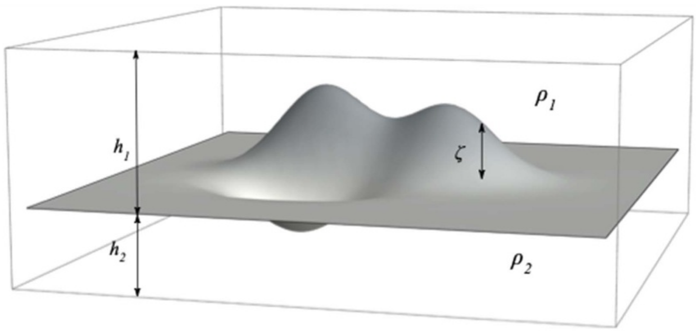

Figure 1. The geometry of a 3D two-layer configuration.

a natural Hamiltonian structure, as presented in [7]. This formulation, unlike other possibilities in the literature (e.g. [31, 35, 36]), avoids the use of potentials and utilizes dependent variables constructed directly from the measurable quantities of density and velocity. The Hamiltonian structure is arguably the preferred avenue for investigating conservation laws of a dynamical system, and is an excellent starting point for a systematic approach to its analysis from the perspective of Hamiltonian geometry that we shall henceforth adopt. Although other variational formulations of the problem are available in the literature (see, e.g. [2, 17, 30]) this framework is the preferred one to be exploited for systematic reductions to models that inherit the parent Hamiltonian structure and its conservation laws, and that can shed light on the dynamical evolution by providing closed-form computable solutions (e.g. travelling waves).

The following notation will be used throughout the paper:

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \boldsymbol{\nabla} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \equiv \begin{pmatrix} \nabla \\ \partial_z \end{pmatrix}$$

so that ∇ will refer to the gradient with respect to the two-dimensional horizontal components (x, y). The horizontal divergence will be denoted by $\nabla\cdot$, and the Laplacian will be understood to be the ordinary definition $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$, which, when applied to functions depending on the horizontal coordinates will collapse to $\partial_x^2 + \partial_y^2 = \nabla \cdot \nabla = \Delta$.

Three dimensional vector variables will be represented by boldface capital letters, and their components, in lower case, by three different letters, e.g.

$$\boldsymbol{U}(x,y,z,t) = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \qquad \boldsymbol{\Sigma}(x,y,z,t) = \begin{pmatrix} \sigma \\ \tau \\ \chi \end{pmatrix},$$

will be the velocity field and the weighted vorticity, respectively, while the lower index $i = 1,2$ will refer to the layer component, with $i = 1$ for the upper layer and $i = 2$ for the lower one. Values of the dependent variables restricted to the interface $z = \zeta(x,y,t)$ will be adorned by tildes, e.g.

$$\widetilde{\boldsymbol{U}}(x,y,t) = \boldsymbol{U}\big(x,y,\zeta(x,y,t),t\big) = \begin{pmatrix} \widetilde{u} \\ \widetilde{v} \\ \widetilde{w} \end{pmatrix}.$$

This paper is organized as follows: section 2 presents a quick overview of the Hamiltonian structure of the 3D Euler equations with density stratification, before moving on to consider the limiting configuration of two (homogeneous) density layers and discuss its Hamiltonian reduction in section 2.1. Next, section 3 focuses on the asymptotic expansion in the long wave limit of this Hamiltonian structure, and extends the Hamiltonian reduction, ultimately to derive the so-called two-dimensional Kaup–Broer–Kupershmidt Boussinesq (KBK-B) model governing the weakly nonlinear (WNL) regimes of dispersive waves propagation at the interface between the two fluids. Special solution classes of this model are discussed in section 4, as well as its well-posedness features determined by the dispersion relation around simple equilibria solutions. Section 5 continues the process of systematic Hamiltonian reduction in the presence of constraints to impose the nearly unidirectional asymptotic limit of propagation to the KBK-B model. This results in the appropriate version of the celebrated Kadomtsev–Petviashvili (KP) equation for our setup. Conclusions and perspectives on future work are presented in section 6.

The relation between the Hamiltonian structures of the parent equations already present in the literature, details on Hamiltonian reduction process, and the connection with the Dirichlet to Neumann approach are summarized in appendices.

2. The Benjamin-Bowman structure and its reduction

We consider a three dimensional stratified Euler fluid contained in an unbounded box

$$\mathcal{B} = \mathbb{R}^2 \times (-h_2, h_1) \ni (x, y, z)\,; \tag{2.1}$$

such a fluid is governed by the incompressible Euler equations for the velocity field $\boldsymbol{U} = (u, v, w)$ and non-constant density $\rho = \rho(x, y, z, t)$ in the presence of gravity $-g\boldsymbol{k}$:

$$\rho_t + \boldsymbol{U}\cdot\boldsymbol{\nabla}\rho = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{U} = 0, \qquad \boldsymbol{U}_t + (\boldsymbol{U}\cdot\boldsymbol{\nabla})\,\boldsymbol{U} + \frac{\boldsymbol{\nabla}p}{\rho} + g\boldsymbol{k} = \boldsymbol{0}, \tag{2.2}$$

with boundary conditions

$$\begin{aligned} &\boldsymbol{U}(r\to\infty, z, t) = \boldsymbol{0}, \quad \rho(r\to\infty, z, t) = \rho_0(z), \\ &w(x, y, -h_2, t) = w(x, y, h_1, t) = 0, \\ &r = \sqrt{x^2+y^2}, \quad z \in (-h_2, h_1), \quad t \in \mathbb{R}^+. \end{aligned} \tag{2.3}$$

We remark that $z = -h_2$ and $z = h_1$ are the locations of the bottom and top confining plates, respectively, and that $\rho_0(z)$ is a reference density, which can at most depend on the vertical coordinate z.

Generalizing to the physical three dimensions the results in two dimensions obtained by Benjamin in [3], Bowman [7] was able to endow equations (2.2) with a non-standard Hamiltonian structure, which can be described as follows. The Hamiltonian variables are to be taken as the pair $(\rho, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the *weighted* vorticity vector

$$\boldsymbol{\Sigma} = \boldsymbol{\nabla}\times(\rho\,\boldsymbol{U})\,. \tag{2.4}$$

The Hamiltonian functional is the total energy

$$\mathcal{H} = \int_{\mathcal{B}} \left(\frac{1}{2}\rho|\boldsymbol{U}|^2 + gz(\rho-\rho_0)\right) \mathrm{d}^3x, \tag{2.5}$$

and the Hamiltonian (Poisson) operator is the operator-valued 4×4 matrix in the evolution equations

$$\begin{pmatrix} \rho_t \\ \boldsymbol{\Sigma}_t \end{pmatrix} = -\begin{pmatrix} 0 & \boldsymbol{\nabla}\cdot(\rho\boldsymbol{\nabla}\times) \\ \boldsymbol{\nabla}\times(\rho\boldsymbol{\nabla}\cdot) & \boldsymbol{\nabla}\times(\boldsymbol{\Sigma}\times\boldsymbol{\nabla}\times) \end{pmatrix} \begin{pmatrix} \dfrac{\delta\mathcal{H}}{\delta\rho} \\ \dfrac{\delta\mathcal{H}}{\delta\boldsymbol{\Sigma}} \end{pmatrix}, \tag{2.6}$$

a system that can be shown [7] to be equivalent to the incompressible variable density Euler equations (2.2).

Remark 2.1. A Hamiltonian formulation of the full 3D Euler equations for vanishing helicity flows was found in [28, 35, 36] in terms of Clebsch potentials. The Benjamin-Bowman formulation we started from is tailored to the incompressible regime, but has the advantage of being formulated by means of physically measurable quantities, and not restricted to zero helicity. The two formulations are connected by a suitable ‘coordinate’ transformation. We further discuss such a relation in appendix A.

In the spirit of our previous papers [9, 10, 21] we now want to specialize the Hamiltonian structure (2.6) to the case of a sharply stratified two-layered configuration, in which a homogeneous fluid of constant density ρ_2 lies under another fluid of constant density ρ_1 (with $\rho_2 > \rho_1$ for stability). The two fluids are separated by an interface $z = \zeta(x, y, t)$, i.e. ζ is a function of the horizontal variables (x, y), and evolves with time t, a setup which we pictorially represent in figure 1.

The Euler variables are expressed via the Heaviside θ-function as

$$\begin{aligned} \rho(x, z, t) &= \rho_2 + (\rho_1-\rho_2)\theta(z-\zeta(x, y, t)) \\ \boldsymbol{U}(x, y, z, t) &= \boldsymbol{U}_2(x, y, z, t) + (\boldsymbol{U}_1(x, y, z, t) - \boldsymbol{U}_2(x, y, z, t))\,\theta(z-\zeta(x, y, t)), \end{aligned} \tag{2.7}$$

where $\boldsymbol{U}_1 = (u_1, v_1, w_1)$ and $\boldsymbol{U}_2 = (u_2, v_2, w_2)$ denote the velocity vector fields in the two domains, while the reference far-field density is

$$\rho_0(z) = \rho_2 + (\rho_1 - \rho_2)\theta(z), \quad z \in (-h_2, h_1). \tag{2.8}$$

Assuming the flow to be potential in the bulk of both domains, i.e.

$$\boldsymbol{U}_i = \boldsymbol{\nabla}\boldsymbol{\Phi}_i, \quad i = 1, 2, \tag{2.9}$$

the characteristic variable $\boldsymbol{\Sigma}$ is a 'vortex sheet' supported on the interface given by

$$\boldsymbol{\Sigma}(x,y,z;t) = \delta(z-\zeta)\begin{pmatrix} \rho_2 v_2 - \rho_1 v_1 + \zeta_y(\rho_2 w_2 - \rho_1 w_1) \\ -(\rho_2 u_2 - \rho_1 u_1 + \zeta_x(\rho_2 w_2 - \rho_1 w_1)) \\ \zeta_x(\rho_2 v_2 - \rho_1 v_1) - \zeta_y(\rho_2 u_2 - \rho_1 u_1) \end{pmatrix} = \delta(z-\zeta)\begin{pmatrix} \widetilde{\sigma}(x,y,t) \\ \widetilde{\tau}(x,y,t) \\ \widetilde{\chi}(x,y,t) \end{pmatrix}, \tag{2.10}$$

where we used the Dirac δ-function distribution, and we have defined the sheet variables according to the notation introduced in § 1. Equation (2.10) leads to the following

Proposition 2.2. *The interface vorticity $\widetilde{\boldsymbol{\Sigma}} = (\widetilde{\sigma}, \widetilde{\tau}, \widetilde{\chi})$ satisfies the following properties*

$$\widetilde{\chi} = \zeta_x \widetilde{\sigma} + \zeta_y \widetilde{\tau}, \qquad \nabla \cdot \begin{pmatrix} \widetilde{\sigma} \\ \widetilde{\tau} \end{pmatrix} = \widetilde{\sigma}_x + \widetilde{\tau}_y = 0. \tag{2.11}$$

Proof. The first property follows directly from (2.10).

For the second, we recall that by its very definition (2.4), $\boldsymbol{\nabla} \cdot \boldsymbol{\Sigma} = 0$. This yields

$$\begin{aligned} &\left[\partial_x(\rho_2 v_2 - \rho_1 v_1 + \zeta_y(\rho_2 w_2 - \rho_1 w_1)) - \partial_y(\rho_2 u_2 - \rho_1 u_1 + \zeta_x(\rho_2 w_2 - \rho_1 w_1))\right. \\ &\left. + \partial_z(\zeta_x(\rho_2 v_2 - \rho_1 v_1) - \zeta_y(\rho_2 u_2 - \rho_1 u_1))\right]\delta(z-\zeta) + \\ &\left[-\zeta_x(\rho_2 v_2 - \rho_1 v_1 + \zeta_y(\rho_2 w_2 - \rho_1 w_1)) + \zeta_y(\rho_2 u_2 - \rho_1 u_1 + \zeta_x(\rho_2 w_2 - \rho_1 w_1))\right. \\ &\left. + \zeta_x(\rho_2 v_2 - \rho_1 v_1) - \zeta_y(\rho_2 u_2 - \rho_1 u_1)\right]\delta'(z-\zeta) = 0. \end{aligned} \tag{2.12}$$

The vanishing of the $\delta'(z-\zeta)$ coefficient is just a restatement of the first property in (2.11). By straightforward algebraic manipulations, and taking into account that $\partial_z \zeta = 0$, the $\delta(z-\zeta)$ coefficient of (2.12) turns into

$$\begin{aligned} &\left[\partial_x(\rho_2 v_2 - \rho_1 v_1 + \zeta_y(\rho_2 w_2 - \rho_1 w_1)) + \zeta_x \partial_z(\rho_2 v_2 - \rho_1 v_1) + \right. \\ &\left. -\partial_y(\rho_2 u_2 - \rho_1 u_1 + \zeta_x(\rho_2 w_2 - \rho_1 w_1)) + \zeta_y \partial_z((\rho_2 u_2 - \rho_1 u_1))\right]. \end{aligned} \tag{2.13}$$

Adding and subtracting the term $\zeta_x \zeta_y(\rho_2 w_2 - \rho_1 w_1)$ and judiciously collecting terms we finally get

$$\widetilde{\sigma}_x + \widetilde{\tau}_y = 0, \tag{2.14}$$

that is, the second property in (2.11).

In summary, the assumption of a sharp 2-layer stratification and of the absence of bulk vorticity leaves us with a weighted vorticity sheet $\widetilde{\boldsymbol{\Sigma}}$ satisfying the 'constraints' of Proposition 2.1, i.e. with effective 2D variables $(\zeta, \widetilde{\sigma}, \widetilde{\tau})$. In the next section we shall reduce the Benjamin-Bowman Hamiltonian operator given by (2.6) to the 'manifold' parametrized by these effective variables.

Remark 2.3. It should be remembered that $(\widetilde{\sigma}, \widetilde{\tau})$ is evaluated at points on the interface. Hence the equation $\widetilde{\sigma}_x + \widetilde{\tau}_y = 0$ is to be carefully interpreted. As usual, this equation locally implies the existence of an interface-stream function $\widetilde{\psi}(x,y,t)$ such that

$$\widetilde{\sigma} = \widetilde{\psi}_y, \qquad \widetilde{\tau} = -\widetilde{\psi}_x. \tag{2.15}$$

In what follows, from section 4 onwards, we will make extensive use the fact that the orthogonal 2-vector $(-\widetilde{\tau}, \widetilde{\sigma})$ is a potential field.

2.1. The Hamiltonian reduction

As we already discussed, the incompressible non-homogeneous Euler equations for fluids in the box $\mathcal{B}$ are defined on the manifold (dropping t-dependence of the variables for ease of notation)

$$\mathcal{M} = \{(\rho(x,y,z), \mathbf{\Sigma}(x,y,z))\}, \quad (x,y) \in \mathbb{R}^2, z \in (-h_2, h_1). \tag{2.16}$$

This manifold is endowed with the Poisson tensor

$$\mathcal{P}_B = -\begin{pmatrix} 0 & \mathbf{\nabla}\cdot(\rho\,\mathbf{\nabla}\times\ ) \\ \mathbf{\nabla}\times(\rho\,\mathbf{\nabla}\cdot) & \mathbf{\nabla}\times(\mathbf{\Sigma}\times\mathbf{\nabla}\times) \end{pmatrix}. \tag{2.17}$$

We consider the submanifold of two-layer configurations

$$\begin{aligned}\mathcal{S} = \{\rho = \rho_2 + (\rho_1-\rho_2)\theta(z-\zeta(x,y)), \quad \sigma = \widetilde{\sigma}(x,y)\delta(z-\zeta(x,y)),\\ \tau = \widetilde{\tau}(x,y)\delta(z-\zeta(x,y)), \quad \chi = \widetilde{\chi}(x,y)\delta(z-\zeta(x,y)), \quad \widetilde{\chi} = \zeta_x\widetilde{\sigma}+\zeta_y\widetilde{\tau}\}.\end{aligned} \tag{2.18}$$

In order to equip $\mathcal{S}$ with a natural Hamiltonian structure, we use one of the specializations of the Marsden-Ratiu (MR) Hamiltonian reduction theorem [27], considering the reduction with respect to the distribution $\mathcal{D}$ defined as

$$\mathcal{D} = \mathcal{P}_B\left((T\mathcal{S})^0\right), \tag{2.19}$$

where $(T\mathcal{S})^0$ is the annihilator of the tangent space $T\mathcal{S} \subset T\mathcal{M}$ (this approach parallels what has successfully been applied in the 2D case in [10, 21]). Then we can obtain a Poisson tensor on the quotient $\mathcal{S}/(P_B(T\mathcal{S})^0)$ by means of the projection

$$\zeta = \frac{1}{\rho_2-\rho_1}\int_{-h_2}^{h_1}(\rho-\rho_1)\,\mathrm{d}z - h_2, \quad \widetilde{\sigma} = \int_{-h_2}^{h_1}\sigma\,\mathrm{d}z, \quad \widetilde{\tau} = \int_{-h_2}^{h_1}\tau\,\mathrm{d}z. \tag{2.20}$$

We refer to appendix B for details. Note the empty intersection

$$\mathcal{P}_B\left((T\mathcal{S})^0\right)\cap T\mathcal{S} = \{0\}, \tag{2.21}$$

so that the MR reduced Poisson manifold is indeed identical to $\mathcal{S}$, and thus parametrized by the triple $(\zeta,\widetilde{\sigma},\widetilde{\tau})$. In these coordinates the reduced Poisson tensor is given by the matrix of differential operators

$$\mathcal{P} = \begin{pmatrix} 0 & -\partial_y & \partial_x \\ -\partial_y & 0 & 0 \\ \partial_x & 0 & 0 \end{pmatrix}. \tag{2.22}$$

In summary, the geometric Poisson reduction procedure endows the phase space of 2-layered configurations with a simple Poisson tensor, independent of the density of the two layers as well as other hardware parameters of the model, leading to the Hamilton equations of motion

$$\begin{cases} \zeta_t = -\partial_y\dfrac{\delta\mathcal{H}}{\delta\widetilde{\sigma}} + \partial_x\dfrac{\delta\mathcal{H}}{\delta\widetilde{\tau}} \\ \widetilde{\sigma}_t = -\partial_y\dfrac{\delta\mathcal{H}}{\delta\zeta} \\ \widetilde{\tau}_t = \partial_x\dfrac{\delta\mathcal{H}}{\delta\zeta}\ , \end{cases} \tag{2.23}$$

supplemented by the condition $\widetilde{\sigma}_x + \widetilde{\tau}_y = 0$, where $\mathcal{H}$ is the energy (2.5).

The next step is to express the Hamiltonian of such configurations in terms of the Hamiltonian variables $(\zeta,\widetilde{\sigma},\widetilde{\tau})$, a task we face in section 3. We can anticipate that in the 3-dimensional setting, this process is quite more involved than in its 2-dimensional counterpart. However, restricting to the lowest non-trivial order in a natural asymptotic expansion yields a *local* Hamiltonian density, leading to the motion equations (3.53).

3. The WNL asymptotics and the KBK-Boussinesq 2D model

In this section we shall first tackle the problem of computing the Hamiltonian of the reduced 2-layer system whose (simple) Poisson structure was determined in section 2.1, in terms of the free variables $(\zeta,\widetilde{\sigma},\widetilde{\tau})$. Next, we shall restrict ourselves to the simplest dispersive asymptotic model, which we will denote as the WNL system.

3.1. Asymptotic expansion of the energy

The assumption of bulk irrotationality of the fluid flow allows us to introduce the bulk velocity potentials $\Phi_i(x,y,z)$. The Taylor series expansions with respect to the vertical variable z of these potentials at the plate locations $z_i \equiv (-1)^{i-1}h_i$ are

$$\Phi_i(x,y,z) = \sum_{n=0}^{\infty} \frac{(z+(-1)^i h_i)^n}{n!} \frac{\partial^n \Phi_i}{\partial z^n}(x,y,z_i)\,. \tag{3.1}$$

Every odd term of these series is zero, due to incompressibility, i.e. $\Delta\Phi_i = 0$, and to the kinematic boundary conditions at the bottom and at the top of the channel $\partial_z \Phi_i(x,y,z_i) = 0$ (see, e.g. [33]). Thus, with $n = 2m$

$$\partial_z^{2m}\Phi_i(x,y,z_i) = (-1)^m(\partial_x^2+\partial_y^2)^m \Phi_{i\,0}\,, \tag{3.2}$$

where we denoted by Φ_{i0} the bulk potentials at the boundary, i.e. $\Phi_{i0} = \Phi_i(x,y,z_i)$, and (3.1) reduces to

$$\Phi_i(x,y,z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(z+(-1)^i h_i)^{2n}(\partial_x^2+\partial_y^2)^n \Phi_{i\,0}\,. \tag{3.3}$$

The corresponding Taylor expansions of the velocity components are

$$\begin{aligned}
u_i(x,y,z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(z+(-1)^i h_i)^{2n}(\partial_x^2+\partial_y^2)^n u_{i\,0}\,,\\
v_i(x,y,z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(z+(-1)^i h_i)^{2n}(\partial_x^2+\partial_y^2)^n v_{i\,0}\,,\\
w_i(x,y,z) &= \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!}(z+(-1)^i h_i)^{2n+1}(\partial_x^2+\partial_y^2)^n (\partial_x u_{i\,0}+\partial_y v_{i\,0})\,,
\end{aligned} \tag{3.4}$$

and their counterpart for the interface velocities $\widetilde{\mathbf{U}}(x,y) = \mathbf{U}(x,y,\zeta(x,y))$ are

$$\begin{aligned}
\widetilde{u}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n u_{i\,0}\,,\\
\widetilde{v}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n v_{i\,0}\,,\\
\widetilde{w}_i(x,y) &= (-1)^i\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!}\eta_i^{2n+1}(\partial_x^2+\partial_y^2)^n (\partial_x u_{i\,0}+\partial_y v_{i\,0})\,,
\end{aligned} \tag{3.5}$$

since

$$\eta_1 = h_1 - \zeta, \quad \eta_2 = h_2 + \zeta\,. \tag{3.6}$$

The power series (3.4) and (3.5) can be used as the starting point of long-wave asymptotics, by introducing rescaled nondimensional independent variables,

$$x = Lx^*, \quad y = Ly^*, \quad z = hz^*\,. \tag{3.7}$$

Here L is assumed to be the typical horizontal wavelength of the dynamics, and $h = h_1 + h_2$ offers a natural vertical scale. Long wave asymptotics is defined assuming that the parameter $\epsilon = h/L$ is small. This can also be viewed as the weight of dispersion corrections to the well known shallow water wave theory. Moreover, the maximum displacement a from equilibrium can be used to nondimensionalize the smooth interface $\zeta(x,y)$,

$$\zeta = a\zeta^*\,. \tag{3.8}$$

This defines the parameter $\alpha = a/h$ whose smallness ultimately characterizes the WNL asymptotics to be discussed next.

Unless otherwise explicitly stated, from now on we will use non-dimensional variables and drop asterisks, but, when needed, with a slight abuse of notation the order symbol $O(\cdot)$ will also denote magnitude of bounded *dimensional* quantities whenever this can be done without generating confusion. With the above scalings, the formal power series (3.4) and (3.5) become

$$\begin{aligned}
u_i(x,y,z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\epsilon^{2n}\left(z+(-1)^i\frac{h_i}{h}\right)^{2n}(\partial_x^2+\partial_y^2)^n u_{i\,0}\,,\\
v_i(x,y,z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\epsilon^{2n}\left(z+(-1)^i\frac{h_i}{h}\right)^{2n}(\partial_x^2+\partial_y^2)^n v_{i\,0}\,,\\
w_i(x,y,z) &= \epsilon\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!}\epsilon^{2n}\left(z+(-1)^i\frac{h_i}{h}\right)^{2n+1}(\partial_x^2+\partial_y^2)^n(\partial_x u_{i\,0}+\partial_y v_{i\,0})\,,
\end{aligned} \tag{3.9}$$

and

$$\begin{aligned}
\widetilde{u}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\epsilon^{2n}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n u_{i\,0}\,,\\
\widetilde{v}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\epsilon^{2n}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n v_{i\,0}\,,\\
\widetilde{w}_i(x,y) &= (-1)^i\epsilon\sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!}\epsilon^{2n}\eta_i^{2n+1}(\partial_x^2+\partial_y^2)^n(\partial_x u_{i\,0}+\partial_y v_{i\,0})\,,
\end{aligned} \tag{3.10}$$

where the nondimensional thicknesses are now

$$\eta_1 = \frac{h_1}{h} - \alpha\zeta, \quad \eta_2 = \frac{h_2}{h} + \alpha\zeta\,. \tag{3.11}$$

For the next step we need the following relations

$$\begin{aligned}
u_{i0}(x,y) &= \widetilde{u}_i + \frac{\epsilon^2}{2}\eta_i^2\Delta\widetilde{u}_i + O(\epsilon^4)\,,\\
v_{i0}(x,y) &= \widetilde{v}_i + \frac{\epsilon^2}{2}\eta_i^2\Delta\widetilde{v}_i + O(\epsilon^4)\,.
\end{aligned} \tag{3.12}$$

In general, the inversion of such near identity operators leads to the asymptotic formula,

$$(\mathbf{1}+\epsilon^2\mathbf{D})^{-1} = \mathbf{1} - \epsilon^2\mathbf{D} + O(\epsilon^4). \tag{3.13}$$

In the following section we will also need the asymptotic expansions of the layer-averaged horizontal velocities, $\bar{u}_i(x,y)$, $\bar{v}_i(x,y)$ defined as

$$\bar{u}_i(x,y) := \frac{1}{\eta_i}\int_{[\eta_i]} u_i(x,y,z)\mathrm{d}z\,, \quad \bar{v}_i(x,y) := \frac{1}{\eta_i}\int_{[\eta_i]} v_i(x,y,z)\mathrm{d}z\,, \tag{3.14}$$

where definite integral notation stands for integration over z in each layer $i = 1,2$. By the Taylor expansions (3.9), these are

$$\begin{aligned}
\bar{u}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\epsilon^{2n}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n u_{i\,0}\,,\\
\bar{v}_i(x,y) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\epsilon^{2n}\eta_i^{2n}(\partial_x^2+\partial_y^2)^n v_{i\,0}\,,
\end{aligned} \tag{3.15}$$

which can be used to provide their expression in terms of the interfacial variables $\widetilde{u}_i, \widetilde{v}_i$.

Indeed, at the lowest order we have

$$\begin{aligned}
u_2(x,y,z) &= u_{20} - \frac{\epsilon^2}{2}\left(z+\frac{h_2}{h}\right)^2\Delta u_{20} + O(\epsilon^4)\,,\\
v_2(x,y,z) &= v_{20} - \frac{\epsilon^2}{2}\left(z+\frac{h_2}{h}\right)^2\Delta v_{20} + O(\epsilon^4)\,,
\end{aligned} \tag{3.16}$$

whence, by (3.10) we get

$$
\begin{aligned}
u_2(x,y,z) &= \widetilde{u}_2 + \frac{\epsilon^2}{2}\left(\eta_2^2 - (z+\frac{h_2}{h})^2\right)\Delta\widetilde{u}_2 + O(\epsilon^4)\,,\\
v_2(x,y,z) &= \widetilde{v}_2 + \frac{\epsilon^2}{2}\left(\eta_2^2 - (z+\frac{h_2}{h})^2\right)\Delta\widetilde{v}_2 + O(\epsilon^4)\,,
\end{aligned}
\tag{3.17}
$$

and, for the vertical velocity,

$$
w_2(x,y,z) = -\epsilon\left(z+\frac{h_2}{h}\right)\nabla\cdot\begin{pmatrix}\widetilde{u}_2\\ \widetilde{v}_2\end{pmatrix} + O(\epsilon^3)\,,
\tag{3.18}
$$

and similar scaling relations holding for the upper layer velocities. For the layer averaged velocities, the above relations imply

$$
\begin{aligned}
\bar{u}_i &= \widetilde{u}_i + \frac{\epsilon^2}{3}\frac{\eta_i^2}{h^2}\Delta\widetilde{u}_i + O(\epsilon^4) = \widetilde{u}_i + \frac{\epsilon^2}{3}\frac{h_i^2}{h^2}\Delta\widetilde{u}_i + O(\alpha\epsilon^2,\epsilon^4),\\
\bar{v}_i &= \widetilde{v}_i + \frac{\epsilon^2}{3}\frac{h_i^2}{h^2}\Delta\widetilde{v}_i + O(\epsilon^4) \ = \widetilde{v}_i + \frac{\epsilon^2}{3}\frac{h_i^2}{h^2}\Delta\widetilde{v}_i + O(\alpha\epsilon^2,\epsilon^4)\,.
\end{aligned}
\tag{3.19}
$$

Such asymptotic formulas will be widely used in what follows.

3.2. The energy

Our next task is to write the explicit form of the energy in terms of horizontal interfacial velocities $(\widetilde{u}_i,\widetilde{v}_i)$. While we can do this at various levels of approximations, for the time being we consider terms of order $O(\epsilon^2)$, discard terms of order $O(\epsilon^4)$ or higher, and retain all terms of every order in the parameter α. All the asymptotic manipulations are needed solely for the kinetic energy, the potential energy being straightforwardly computed.

Let us illustrate the above strategy with the lower fluid. Its kinetic energy density reads

$$
T_2 = \frac{\rho_2}{2}\int_{-h_2/h}^{\alpha\zeta}(u_2^2+v_2^2+w_2^2)\,h\,\mathrm{d}z\,,
\tag{3.20}
$$

the dimensional factor h coming from the scaling of the physical vertical coordinate z. This leads to

$$
\begin{aligned}
T_2 &= \frac{h\rho_2}{2}\int_{-\frac{h_2}{h}}^{\alpha\zeta}\left\{\widetilde{u}_2^2+\widetilde{v}_2^2+\epsilon^2\left[2\left(\widetilde{u}_2\Delta\widetilde{u}_2+\widetilde{v}_2\Delta\widetilde{v}_2\right)\left(\eta_2^2-\left(z+\frac{h_2}{h}\right)^2\right)+\right.\right.\\
&\quad\left.\left.+\left(\widetilde{u}_{2,x}+\widetilde{v}_{2,y}\right)^2\left(z+\frac{h_2}{h}\right)^2\right]+O(\epsilon^4)\right\}\mathrm{d}z\\
&= \frac{h\rho_2}{2}\left[\eta_2(\widetilde{u}_2^2+\widetilde{v}_2^2)+\frac{2\epsilon^2}{3}\eta_2^3\left(\widetilde{u}_2\Delta\widetilde{u}_2+\widetilde{v}_2\Delta\widetilde{v}_2+\frac{1}{2}\left(\widetilde{u}_{2x}+\widetilde{v}_{2y}\right)^2\right)\right]+O(\epsilon^4)\,.
\end{aligned}
\tag{3.21}
$$

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

$$
T_1 = \frac{h\rho_1}{2}\left[\eta_1(\widetilde{u}_1^2+\widetilde{v}_1^2)+\frac{2\epsilon^2}{3}\eta_1^3\left(\widetilde{u}_1\Delta\widetilde{u}_1+\widetilde{v}_1\Delta\widetilde{v}_1+\frac{1}{2}\left(\widetilde{u}_{1x}+\widetilde{v}_{1y}\right)^2\right)\right]+O(\epsilon^4)\,.
\tag{3.22}
$$

3.3. The constraint

As well known (see, e.g. [11]), in three dimensions the kinematic boundary conditions at the interface can be expressed as

$$
\nabla\cdot\left(\eta_1(x,y)\begin{pmatrix}\bar{u}_1(x,y)\\ \bar{v}_1(x,y)\end{pmatrix}+\eta_2(x,y)\begin{pmatrix}\bar{u}_2(x,y)\\ \bar{v}_2(x,y)\end{pmatrix}\right)=0\,,
\tag{3.23}
$$

where $(\bar{u}_i,\bar{v}_i)$ are the layer-averaged horizontal velocities. This implies the existence of a 'mass-stream'-function $\mu(x,y)$ such that

$$
\eta_1(x,y)\begin{pmatrix}\bar{u}_1(x,y)\\ \bar{v}_1(x,y)\end{pmatrix}+\eta_2(x,y)\begin{pmatrix}\bar{u}_2(x,y)\\ \bar{v}_2(x,y)\end{pmatrix}=\begin{pmatrix}\mu_y\\ -\mu_x\end{pmatrix}.
\tag{3.24}
$$

Since in general the mass-stream function cannot be assumed to vanish, a careful asymptotic analysis is needed to arrive at a workable form of the constraint (3.24), by fixing suitable scaling relations between the long-wave small parameter ϵ and the nonlinearity small parameter α.

Looking at the 3D velocity potentials $\Phi_i(x,y,z)$ and their interfacial values

$$\widetilde{\Phi}_i(x,y) = \Phi_i(x,y,z)\big|_{z=\zeta(x,y)}, \tag{3.25}$$

we have, e.g.

$$\widetilde{\Phi}_{1,x} = \partial_x(\Phi_1(x,y,z))\big|_{z=\zeta(x,y)} + \partial_z(\Phi_1(x,y,z))\big|_{z=\zeta(x,y)}\zeta_x = \widetilde{u}_1 + \widetilde{w}_1\zeta_x. \tag{3.26}$$

Since the slope of the normalized interface is small and scales as $O(\alpha\epsilon)$, and by the Taylor expansion (3.9) the vertical velocities are of order ϵ, in the WNL expansion we get $\widetilde{u}_1 = \widetilde{\Phi}_{1,x} + O(\alpha\epsilon^2)$. By the same argument we have

$$\begin{pmatrix}\widetilde{u}_i\\ \widetilde{v}_i\end{pmatrix} = \nabla\widetilde{\Phi}_i + O(\alpha\epsilon^2). \tag{3.27}$$

By using relation (3.19) between averaged and interface velocities together with (3.27) above, we can write the mass-balance equations (3.24) as the system

$$\begin{cases} \eta_1\widetilde{\Phi}_{1x} + \eta_2\widetilde{\Phi}_{2x} + \dfrac{\epsilon^2}{3h^3}\left(h_1^3(\widetilde{\Phi}_{1xxx} + \widetilde{\Phi}_{1yyx}) + h_2^3(\widetilde{\Phi}_{2xxx} + \widetilde{\Phi}_{2yyx})\right) = \mu_y + O(\epsilon^4,\alpha\epsilon^2)\\ \eta_1\widetilde{\Phi}_{1y} + \eta_2\widetilde{\Phi}_{2y} + \dfrac{\epsilon^2}{3h^3}\left(h_1^3(\widetilde{\Phi}_{1xxy} + \widetilde{\Phi}_{1yyy}) + h_2^3(\widetilde{\Phi}_{2xxy} + \widetilde{\Phi}_{2yyy})\right) = -\mu_x + O(\epsilon^4,\alpha\epsilon^2). \end{cases} \tag{3.28}$$

Cross differentiating and taking the difference we get that the mass-stream function satisfies the Poisson equation

$$\begin{aligned}\Delta\mu &= \alpha(-\zeta_y\widetilde{\Phi}_{1x} + \zeta_x\widetilde{\Phi}_{1y} + \zeta_y\widetilde{\Phi}_{2x} - \zeta_x\widetilde{\Phi}_{2,y}) + O(\epsilon^4,\alpha\epsilon^2)\\ &= -\alpha(\nabla\zeta\times\nabla(\widetilde{\Phi}_2 - \widetilde{\Phi}_1)) + O(\epsilon^4,\alpha\epsilon^2),\end{aligned} \tag{3.29}$$

where we considered the small amplitude parameter α entering the definition of the normalized interface ζ (see (3.11)). By restricting the model in the WNL asymptotics $\alpha \ll \epsilon$, i.e. considering terms of order α,ϵ^2 and ignoring terms of order $\alpha^2,\alpha\epsilon^2,\epsilon^4$, we can trade the general constraint equation (3.28) for the WNL constraint system:

$$\begin{cases} \eta_1\widetilde{u}_1 + \eta_2\widetilde{u}_2 + \dfrac{\epsilon^2}{3h^3}\left(h_1^3\Delta\widetilde{u}_1 + h_2^3\Delta\widetilde{u}_2\right) = -\alpha\,\partial_y\Delta^{-1}\left(\nabla\zeta\times\begin{pmatrix}\widetilde{u}_2-\widetilde{u}_1\\ \widetilde{v}_2-\widetilde{v}_1\end{pmatrix}\right) + O(\epsilon^4,\alpha\epsilon^2)\\ \eta_1\widetilde{v}_1 + \eta_2\widetilde{v}_2 + \dfrac{\epsilon^2}{3h^3}\left(h_1^3\Delta\widetilde{v}_1 + h_2^3\Delta\widetilde{v}_2\right) = \alpha\,\partial_x\Delta^{-1}\left(\nabla\zeta\times\begin{pmatrix}\widetilde{u}_2-\widetilde{u}_1\\ \widetilde{v}_2-\widetilde{v}_1\end{pmatrix}\right) + O(\epsilon^4,\alpha\epsilon^2). \end{cases} \tag{3.30}$$

This constraint can be manipulated to give explicit relations between the velocities of the two layers, better written in the vector form

$$\widetilde{\boldsymbol{u}}_2 = -\frac{h_1}{h_2}\widetilde{\boldsymbol{u}}_1 - \frac{\epsilon^2}{3}\frac{h_1}{h_2}\frac{h_1^2-h_2^2}{h^2}\Delta\widetilde{\boldsymbol{u}}_1 + +\alpha h\frac{h_1}{h_2}\left(\frac{1}{h_1}+\frac{1}{h_2}\right)\left(\zeta\widetilde{\boldsymbol{u}}_1 + \boldsymbol{\nabla}\times\Delta^{-1}(\boldsymbol{\nabla}\zeta\times\widetilde{\boldsymbol{u}}_1)\right) + O(\alpha^2,\epsilon^4,\alpha\epsilon^2), \tag{3.31}$$

where $\widetilde{\boldsymbol{u}}_i = \begin{pmatrix}\widetilde{u}_i\\ \widetilde{v}_i\\ 0\end{pmatrix}$.

Remark 3.1. Contrary to the 2-dimensional case, relation (3.31) introduces non-localities in the reduced theory, which parallel the ones obtained in the Dirichlet–Neumann (DN) operator approach of [5] for 3-dimensional 2-layer configurations with a rigid lid. As we will explicitly show in the next section, although the constraint (3.31) is non-local, the WNL approximation yields a local expression for the Hamiltonian in the canonical variables $(\widetilde{\sigma},\widetilde{\tau})$ prescribed by the geometric reduction approach of section 2.1. The comparison of our setting to that of [5] is presented in appendix C.

3.4. WNL case: the energy in Darboux coordinates

The final aim of this subsection is to express the energy of the system in terms of the free variables $(\zeta,\widetilde{\sigma},\widetilde{\tau})$ introduced in section 2.1, with the assumption that the two small parameters in the asymptotics are linked so that α scales as $O(\epsilon^2)$, and hence terms of order $O(\alpha^2,\alpha\epsilon^2,\epsilon^4)$ can be neglected. This will be done in a few steps.

The first one consists in noticing that in this asymptotics the kinetic energy densities (3.22), (3.21) preliminarily reduce to

$$T_{\mathrm{WNL}\,i}=\frac{h\rho_i}{2}\left[\eta_i(\widetilde{u}_i^2+\widetilde{v}_i^2)+\frac{\epsilon^2}{3}\frac{h_i^3}{h^3}\left(2(\widetilde{u}_i\Delta\widetilde{u}_i+\widetilde{v}_i\Delta\widetilde{v}_i)+\left(\widetilde{u}_{ix}+\widetilde{v}_{iy}\right)^2\right)\right]+O(\alpha\epsilon^2,\epsilon^4). \tag{3.32}$$

By using the (vector) Green identity

$$\boldsymbol{X}\cdot\Delta\boldsymbol{X}+(\boldsymbol{\nabla}\cdot\boldsymbol{X})^2=-(|\boldsymbol{\nabla}\times\boldsymbol{X}|)^2+\boldsymbol{\nabla}\cdot\left(\boldsymbol{X}\times(\boldsymbol{\nabla}\times\boldsymbol{X})+\boldsymbol{X}\,(\boldsymbol{\nabla}\cdot\boldsymbol{X})\right), \tag{3.33}$$

we rewrite (3.32) as

$$T_{\mathrm{WNL}\,i}=\frac{h\rho_i}{2}\left[\eta_i(\widetilde{u}_i^2+\widetilde{v}_i^2)+\frac{\epsilon^2}{3}\frac{h_i^3}{h^3}\left(\widetilde{u}_i\Delta\widetilde{u}_i+\widetilde{v}_i\Delta\widetilde{v}_i-\left(\widetilde{v}_{ix}-\widetilde{u}_{iy}\right)^2+T_{\mathrm{div}\,i}\right)\right]+O(\alpha\epsilon^2,\epsilon^4), \tag{3.34}$$

where in $T_{\mathrm{div}\,i}$ we collected the total divergence terms coming from the last summand in (3.33) with $\boldsymbol{X}=(\widetilde{u}_i,\widetilde{v}_i,0)$.

Next, we observe that the WNL approximation simplifies the effective link between $(\widetilde{\sigma},\widetilde{\tau})$ and the velocities. Indeed, (2.10) yields

$$\begin{pmatrix}\widetilde{\sigma}\\ \widetilde{\tau}\end{pmatrix}=\begin{pmatrix}\rho_2\widetilde{v}_2-\rho_1\widetilde{v}_1+\zeta_y(\rho_2\widetilde{w}_2-\rho_1\widetilde{w}_1)\\ -(\rho_2\widetilde{u}_2-\rho_1\widetilde{u}_1+\zeta_x(\rho_2\widetilde{w}_2-\rho_1\widetilde{w}_1))\end{pmatrix}=\begin{pmatrix}\rho_2\widetilde{v}_2-\rho_1\widetilde{v}_1\\ -\rho_2\widetilde{u}_2+\rho_1\widetilde{u}_1\end{pmatrix}+O(\alpha\epsilon^2), \tag{3.35}$$

since terms like $\zeta_y(\rho_2\widetilde{w}_2-\rho_1\widetilde{w}_1)$ scale as $O(\alpha\epsilon^2)$ thanks to (3.11) and (3.18). Moreover, combining (3.35) and the constraint (3.31) we arrive at the relations

$$\begin{cases}\widetilde{u}_1=\dfrac{h_2}{\rho_2h_1+\rho_1h_2}\,\widetilde{\tau}-\dfrac{\epsilon^2}{3}\dfrac{\rho_2h_1h_2(h_1-h_2)}{h(\rho_2h_1+\rho_1h_2)^2}\Delta\widetilde{\tau}+\alpha\dfrac{\rho_2h^2}{(\rho_2h_1+\rho_1h_2)^2}\,(\zeta\widetilde{\tau}+\mathcal{N}_1)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4)\\ \widetilde{u}_2=-\dfrac{h_1}{\rho_2h_1+\rho_1h_2}\,\widetilde{\tau}-\dfrac{\epsilon^2}{3}\dfrac{\rho_1h_1h_2(h_1-h_2)}{h(\rho_2h_1+\rho_1h_2)^2}\,\Delta\widetilde{\tau}+\alpha\dfrac{\rho_1h^2}{(\rho_2h_1+\rho_1h_2)^2}\,(\zeta\widetilde{\tau}+\mathcal{N}_1)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4),\end{cases} \tag{3.36}$$

and

$$\begin{cases}\widetilde{v}_1=-\dfrac{h_2}{\rho_2h_1+\rho_1h_2}\,\widetilde{\sigma}+\dfrac{\epsilon^2}{3}\dfrac{\rho_2h_1h_2(h_1-h_2)}{h(\rho_2h_1+\rho_1h_2)^2}\,\Delta\widetilde{\sigma}-\alpha\dfrac{\rho_2h^2}{(\rho_2h_1+\rho_1h_2)^2}\,(\zeta\widetilde{\sigma}+\mathcal{N}_2)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4)\\ \widetilde{v}_2=\dfrac{h_1}{\rho_2h_1+\rho_1h_2}\,\widetilde{\sigma}+\dfrac{\epsilon^2}{3}\dfrac{\rho_1h_1h_2(h_1-h_2)}{h(\rho_2h_1+\rho_1h_2)^2}\,\Delta\widetilde{\sigma}-\alpha\dfrac{\rho_1h^2}{(\rho_2h_1+\rho_1h_2)^2}\,(\zeta\widetilde{\sigma}+\mathcal{N}_2)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4),\end{cases} \tag{3.37}$$

where $\mathcal{N}_i$ is the ith component of the non-local vector

$$\boldsymbol{\mathcal{N}}=\boldsymbol{\nabla}\times\Delta^{-1}\left(\boldsymbol{\nabla}\zeta\times\begin{pmatrix}\widetilde{\tau}\\ \widetilde{\sigma}\\ 0\end{pmatrix}\right). \tag{3.38}$$

At the leading order, the previous relations yield

$$\begin{cases}\widetilde{u}_1=\dfrac{h_2\widetilde{\tau}}{\rho_2h_1+\rho_1h_2}+O(\alpha,\epsilon^2)\\ \widetilde{u}_2=-\dfrac{h_1\widetilde{\tau}}{\rho_2h_1+\rho_1h_2}+O(\alpha,\epsilon^2)\end{cases},\quad\begin{cases}\widetilde{v}_1=-\dfrac{h_2\widetilde{\sigma}}{\rho_2h_1+\rho_1h_2}+O(\alpha,\epsilon^2)\\ \widetilde{v}_2=\dfrac{h_1\widetilde{\sigma}}{\rho_2h_1+\rho_1h_2}+O(\alpha,\epsilon^2)\end{cases}. \tag{3.39}$$

Let us first consider the sum of the $O(\epsilon^2)$ terms $T^{(2)}_{\mathrm{WNL}\,i}$ in (3.34) (disregarding the total divergence terms $T_{\mathrm{div}\ i}$). Substituting the $O(1)$ relations (3.39) we obtain

$$\begin{aligned}\sum_{i=1}^{2}T^{(2)}_{\mathrm{WNL}\,i}&=\frac{1}{2}\frac{\epsilon^2}{3h^2}\frac{(\rho_1h_1+\rho_2h_2)h_1^2h_2^2}{(\rho_2h_1+\rho_1h_2)^2}\left(\widetilde{\sigma}\Delta\widetilde{\sigma}+\widetilde{\tau}\Delta\widetilde{\tau}-(\widetilde{\sigma}_x+\widetilde{\tau}_y)^2\right)+O(\epsilon^4,\alpha\epsilon^2)\\ &=\frac{1}{2}\frac{\epsilon^2}{3h^2}\frac{(\rho_1h_1+\rho_2h_2)h_1^2h_2^2}{(\rho_2h_1+\rho_1h_2)^2}\left(\widetilde{\sigma}\Delta\widetilde{\sigma}+\widetilde{\tau}\Delta\widetilde{\tau}\right)+O(\epsilon^4,\alpha\epsilon^2),\end{aligned} \tag{3.40}$$

owing to the solenoidal property (2.14).

Let us now consider the remaining terms in (3.34) that add up to

$$\sum_{i=1}^{2} T^{(0)}_{\mathrm{WNL}\,i} = \frac{h}{2}\left(\rho_1\eta_1(\widetilde{u}_1^2+\widetilde{v}_1^2)+\rho_2\eta_2(\widetilde{u}_2^2+\widetilde{v}_2^2)\right). \tag{3.41}$$

Recalling that $\eta_i = \dfrac{h_i}{h} + (-1)^i\alpha\zeta$ and expanding in α we get

$$\sum_{i=1}^{2} T^{(0)}_{\mathrm{WNL}\,i} = \frac{1}{2}\left(\rho_1 h_1(\widetilde{u}_1^2+\widetilde{v}_1^2)+\rho_2 h_2(\widetilde{u}_2^2+\widetilde{v}_2^2)\right) + \alpha\frac{h\zeta}{2}\left(\rho_2(\widetilde{u}_2^2+\widetilde{v}_2^2)-\rho_1(\widetilde{u}_1^2+\widetilde{v}_1^2)\right). \tag{3.42}$$

In the second summand we can substitute (3.39) to obtain

$$\alpha\frac{h\zeta}{2}\left(\rho_2(\widetilde{u}_2^2+\widetilde{v}_2^2)-\rho_1(\widetilde{u}_1^2+\widetilde{v}_1^2)\right) = \alpha\frac{h}{2}\,\frac{h_1^2\rho_2-h_2^2\rho_1}{(\rho_2h_1+\rho_1h_2)^2}\,\zeta\,(\widetilde{\sigma}^2+\widetilde{\tau}^2)+O(\alpha^2,\alpha\epsilon^2). \tag{3.43}$$

To express the first term in (3.42) in terms of $(\zeta,\widetilde{\sigma},\widetilde{\tau})$ we need to use the full relations (3.36) and (3.37). The contributions of the $O(\alpha,\epsilon^2)$ terms in $\frac{1}{2}\left(\rho_1h_1(\widetilde{u}_1^2+\widetilde{v}_1^2)+\rho_2h_2(\widetilde{u}_2^2+\widetilde{v}_2^2)\right)$ are readily seen to cancel out thanks to the interplay of signs and h_i,ρ_i factors, and so the end result is simply

$$\frac{1}{2}\left(\rho_1h_1(\widetilde{u}_1^2+\widetilde{v}_1^2)+\rho_2h_2(\widetilde{u}_2^2+\widetilde{v}_2^2)\right) = \frac{1}{2}\,\frac{h_1h_2}{\rho_2h_1+\rho_1h_2}\,(\widetilde{\sigma}^2+\widetilde{\tau}^2)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4). \tag{3.44}$$

Notice that the effective kinetic energy is a local expression in the field variables. Finally, adding the terms (3.40). (3.43) and (3.44), and discarding total divergences we arrive at the final expression of the total effective kinetic energy density as

$$\begin{aligned} T^{\mathrm{eff}}_{\mathrm{WNL}} = \frac{h}{2}\Bigg(&\frac{h_1h_2}{h(\rho_2h_1+\rho_1h_2)}\,(\widetilde{\sigma}^2+\widetilde{\tau}^2)+\alpha\frac{\rho_2h_1^2-\rho_1h_2^2}{(\rho_2h_1+\rho_1h_2)^2}\zeta\,(\widetilde{\sigma}^2+\widetilde{\tau}^2) \\ &+\frac{\epsilon^2}{3}\,\frac{h_1^2h_2^2(\rho_2h_2+\rho_1h_1)}{h^3(\rho_2h_1+\rho_1h_2)^2}\,(\widetilde{\sigma}\Delta\widetilde{\sigma}+\widetilde{\tau}\Delta\widetilde{\tau})\Bigg)+O(\alpha^2,\alpha\epsilon^2,\epsilon^4). \end{aligned} \tag{3.45}$$

The Hamiltonian energy of the system is given by

$$\mathcal{H}_{\mathrm{WNL}} = L^2\int (T^{\mathrm{eff}}_{\mathrm{WNL}}+U)\,\mathrm{d}x\,\mathrm{d}y \tag{3.46}$$

where U is the potential energy density,

$$U = \frac{1}{2}gh^2(\rho_2-\rho_1)\zeta^2. \tag{3.47}$$

By non-dimensionalizing $(\widetilde{\sigma},\widetilde{\tau})$ according to

$$\widetilde{\sigma} = \widetilde{\sigma}^*\sqrt{g(\rho_2-\rho_1)(\rho_2h_1+\rho_1h_2)}, \qquad \widetilde{\tau} = \widetilde{\tau}^*\sqrt{g(\rho_2-\rho_1)(\rho_2h_1+\rho_1h_2)}, \tag{3.48}$$

discarding $O(\alpha^2,\alpha\epsilon^2,\epsilon^4)$ terms and dropping *'s gives the Hamiltonian density as

$$\begin{aligned} H_{\mathrm{WNL}} = \frac{gh^2(\rho_2-\rho_1)}{2}\Bigg(&\frac{h_1h_2}{h^2}\,(\widetilde{\sigma}^2+\widetilde{\tau}^2)+\zeta^2+\alpha\frac{\rho_2h_1^2-\rho_1h_2^2}{h(\rho_2h_1+\rho_1h_2)}\zeta\,(\widetilde{\sigma}^2+\widetilde{\tau}^2) \\ &+\frac{\epsilon^2}{3}\,\frac{h_1^2h_2^2(\rho_2h_2+\rho_1h_1)}{h^4(\rho_2h_1+\rho_1h_2)}\,(\widetilde{\sigma}\Delta\widetilde{\sigma}+\widetilde{\tau}\Delta\widetilde{\tau})\Bigg). \end{aligned} \tag{3.49}$$

Setting hereafter

$$A = \frac{h_1h_2}{h^2}, \quad B = \frac{\rho_2h_1^2-\rho_1h_2^2}{h(\rho_2h_1+\rho_1h_2)}, \quad \kappa = \frac{h_1^2h_2^2(\rho_1h_1+\rho_2h_2)}{h^4(\rho_2h_1+\rho_1h_2)}, \tag{3.50}$$

and dividing the Hamiltonian density by the factor $D = gh^2(\rho_2 - \rho_1)$ (which is tantamount to a time-rescaling), the variational derivatives of (3.49) can now be computed as

$$\begin{aligned}\frac{\delta\mathcal{H}}{\delta\zeta} &= \zeta + \frac{\alpha}{2}B(\widetilde{\sigma}^2 + \widetilde{\tau}^2),\\ \frac{\delta\mathcal{H}}{\delta\widetilde{\sigma}} &= A\widetilde{\sigma} + \alpha B\zeta\widetilde{\sigma} + \frac{\epsilon^2}{3}\kappa\Delta\widetilde{\sigma},\\ \frac{\delta\mathcal{H}}{\delta\widetilde{\tau}} &= A\widetilde{\tau} + \alpha B\zeta\widetilde{\tau} + \frac{\epsilon^2}{3}\kappa\Delta\widetilde{\tau},\end{aligned} \tag{3.51}$$

where $\mathcal{H} = \frac{1}{D}\int H_{\mathrm{WNL}}\,\mathrm{d}x\,\mathrm{d}y$. Equations of motion can be found using the Hamiltonian structure (2.22) as

$$\begin{pmatrix}\zeta_t\\ \widetilde{\sigma}_t\\ \widetilde{\tau}_t\end{pmatrix} = \begin{pmatrix}0 & -\partial_y & \partial_x\\ -\partial_y & 0 & 0\\ \partial_x & 0 & 0\end{pmatrix}\begin{pmatrix}\frac{\delta\mathcal{H}}{\delta\zeta}\\ \frac{\delta\mathcal{H}}{\delta\widetilde{\sigma}}\\ \frac{\delta\mathcal{H}}{\delta\widetilde{\tau}}\end{pmatrix}, \tag{3.52}$$

to obtain, explicitly,

$$\begin{cases}\zeta_t = A(\widetilde{\tau}_x - \widetilde{\sigma}_y) + \alpha B((\zeta\widetilde{\tau})_x - (\zeta\widetilde{\sigma})_y) + \frac{\epsilon^2}{3}\kappa((\Delta\widetilde{\tau})_x - (\Delta\widetilde{\sigma})_y)\\ \widetilde{\sigma}_t = -\zeta_y - \alpha B(\widetilde{\sigma}\widetilde{\sigma}_y + \widetilde{\tau}\widetilde{\tau}_y)\\ \widetilde{\tau}_t = \zeta_x + \alpha B(\widetilde{\sigma}\widetilde{\sigma}_x + \widetilde{\tau}\widetilde{\tau}_x)\,.\end{cases} \tag{3.53}$$

This system has to be augmented with the constraint of zero-divergence condition

$$\widetilde{\sigma}_x + \widetilde{\tau}_y = 0\,. \tag{3.54}$$

In terms of the 2D vector

$$\boldsymbol{\gamma} = \begin{pmatrix}-\widetilde{\tau}\\ \widetilde{\sigma}\end{pmatrix}, \tag{3.55}$$

system (3.53) can be written in the more compact form

$$\begin{cases}\zeta_t = -A\nabla\cdot\boldsymbol{\gamma} - \alpha B\nabla\cdot(\zeta\boldsymbol{\gamma}) - \frac{\epsilon^2}{3}\kappa\nabla\cdot(\Delta\boldsymbol{\gamma})\\ \boldsymbol{\gamma}_t = -g'\nabla\zeta - \frac{\alpha}{2}B\nabla|\boldsymbol{\gamma}|^2\end{cases} \tag{3.56}$$

with the additional constraint (3.54) becoming the zero-curl condition

$$\nabla\times\boldsymbol{\gamma} = 0\,. \tag{3.57}$$

With the variable $\boldsymbol{\gamma}$ thus defined, the Hamiltonian structure (2.22) assumes the more elegant and compact form

$$\mathcal{P}^{\mathrm{red}} = \begin{pmatrix}0 & -\nabla\cdot\\ -\nabla & 0\end{pmatrix}, \tag{3.58}$$

with Hamiltonian

$$\mathcal{H} = \frac{1}{2}\int\left((A + \alpha B\zeta)|\boldsymbol{\gamma}|^2 + \frac{\epsilon^2}{3}\kappa\Big(\boldsymbol{\gamma}\cdot\Delta\boldsymbol{\gamma}\Big) + g'\zeta^2\right)\mathrm{d}x\,\mathrm{d}y\,. \tag{3.59}$$

Here and in what follows it is useful to insert the non-dimensional parameter g' as a place-holder to highlight the role of the reduced gravity in setting the dynamical evolution. With these variables, this form of the WNL system is sometimes referred to in the literature as the Kaup–Broer–Kupershmidt-Boussinesq (KBK-Boussinesq) equations [24] (see also [5, 17]) and is the two-dimensional generalization of the WNL system in [10], where the Hamiltonian reduction approach was used.

The Hamiltonian approach helps display a few conservation laws associated with (3.56). From the structure of the Poisson tensor, it is clear that the quantities $\int\zeta\,\mathrm{d}x\,\mathrm{d}y$, $\int\boldsymbol{\gamma}\,\mathrm{d}x\,\mathrm{d}y$ are (trivial) constants of the motion, as obviously is the Hamiltonian (3.59) itself. Translational invariance of our problem guarantees the conservation of the generators of x- and y-translations. It is immediate to show (using the irrotationality condition $\gamma_{1y} = \gamma_{2x}$) that they are given by the expressions

$$\mathcal{M}_1 = \int \zeta\gamma_1 \, dx\, dy, \quad \mathcal{M}_2 = \int \zeta\gamma_2 \, dx\, dy. \tag{3.60}$$

The generator of the rotational invariance in the horizontal plane can be found as well, and it is given by

$$\mathcal{L} \equiv \int \zeta\,(x\gamma_2 - y\gamma_1)\, dx\, dy, \tag{3.61}$$

where we assume that all fields vanish sufficiently fast at spatial infinity.

4. Special solutions

We first look at the linearization of system (3.56) around constant states which gives dispersion relations for these equations. Next, we will obtain time-independent solutions of (3.56), and, last, examine travelling wave solutions.

4.1. The dispersion relation

Linearizing system (3.56) around the state $(0, \boldsymbol{\gamma_0}) \equiv (0, \gamma_{1,0}, \gamma_{2,0})$,

$$\begin{cases} \zeta_t + A\nabla\cdot\boldsymbol{\gamma} + \alpha B\boldsymbol{\gamma_0}\cdot\nabla(\zeta) + \frac{\epsilon^2}{3}\kappa\Delta(\nabla\cdot\boldsymbol{\gamma}) = 0 \\ \gamma_{1,t} + g'\zeta_x + \alpha B(\boldsymbol{\gamma_0}\cdot\boldsymbol{\gamma})_x = 0 \\ \gamma_{2,t} + g'\zeta_y + \alpha B(\boldsymbol{\gamma_0}\cdot\boldsymbol{\gamma})_y = 0 \end{cases}, \tag{4.1}$$

and seeking solutions by Fourier transform leads to the dispersion relation. Thus, by setting

$$\zeta(x,y,t) = \hat{\zeta}e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \quad \gamma_1(x,y,t) = \hat{\gamma}_1 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \quad \gamma_2(x,y,t) = \hat{\gamma}_2 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{4.2}$$

where $\boldsymbol{k} = (k, \ell)$, $\boldsymbol{x} = (x, y)$, $|\boldsymbol{k}|^2 = k^2 + \ell^2$ and the hatted quantities are constants, the left hand side of system (4.1) becomes, after simple manipulations, the matrix multiplication

$$\begin{pmatrix} \omega - \alpha B\boldsymbol{k}\cdot\boldsymbol{\gamma_0} & k\left(\frac{\epsilon^2}{3}\kappa|\boldsymbol{k}|^2 - A\right) & \ell\left(\frac{\epsilon^2}{3}\kappa|\boldsymbol{k}|^2 - A\right) \\ -g'k & \omega - \alpha Bk\gamma_{1,0} & -\alpha Bk\gamma_{2,0} \\ -g'\ell & -\alpha B\ell\gamma_{1,0} & \omega - \alpha B\ell\gamma_{2,0} \end{pmatrix} \begin{pmatrix} \hat{\zeta} \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}. \tag{4.3}$$

Nontrivial solutions of the system require the determinant of this matrix to be zero, which leads to

$$\omega^3 - 2\alpha B(\boldsymbol{k}\cdot\gamma_0)\omega^2 + \alpha^2 B^2(\boldsymbol{k}\cdot\gamma_0)^2\omega + (g'\kappa\frac{\epsilon^2}{3}|\boldsymbol{k}|^4 - g'A|\boldsymbol{k}|^2)\omega = 0. \tag{4.4}$$

The dispersion relation is therefore

$$(\omega - \alpha B\boldsymbol{k}\cdot\boldsymbol{\gamma_0})^2 = g'|\boldsymbol{k}|^2\left(A - \frac{\epsilon^2}{3}\kappa|\boldsymbol{k}|^2\right), \tag{4.5}$$

which shows that the linearized system (4.1) is ill-posed, as the RHS of this equation becomes negative and is of order $O(|\boldsymbol{k}|^4)$ as $|\boldsymbol{k}| \to \infty$. As in the two dimensional case [10], a regularization of system (4.1) can be achieved by the change of dependent variables

$$\bar{\gamma} = \gamma + \frac{\epsilon^2}{3}\frac{\kappa}{A}\Delta\gamma, \quad \text{i.e., at } O(\epsilon^2) \quad \gamma = \bar{\gamma} - \frac{\epsilon^2}{3}\frac{\kappa}{A}\Delta\bar{\gamma}, \tag{4.6}$$

where inverse operator are computed thanks to their near identity asymptotic weights.

After linearizing the evolution equations around the constant state $\boldsymbol{\gamma_0}$ with these new variables, the dispersion relation of the corresponding non purely evolutionary system becomes

$$(\omega - \alpha B\boldsymbol{k}\cdot\boldsymbol{\gamma_0})^2 = \frac{g'(|\boldsymbol{k}|^2)A}{1 + \frac{\epsilon^2}{3}\frac{\kappa}{A}(|\boldsymbol{k}|^2)}, \tag{4.7}$$

which is well-posed for all $\boldsymbol{k}$'s.

Note that the change of coordinates (4.6) between $\boldsymbol{\gamma}$ and $\bar{\boldsymbol{\gamma}}$ corresponds to the change of coordinates (3.19) between interfacial and layer averaged variables. In fact, coming back to dimensional variables, with the definition

$$\begin{pmatrix}\bar{\gamma}_1\\ \bar{\gamma}_2\end{pmatrix} \equiv \begin{pmatrix}\rho_2\bar{u}_2-\rho_1\bar{u}_1\\ \rho_2\bar{v}_2-\rho_1\bar{v}_1\end{pmatrix} \tag{4.8}$$

and (3.19) we have, e.g. for the second component,

$$\begin{aligned}\gamma_2 &= \rho_2\widetilde{v}_2-\rho_1\widetilde{v}_1\\ &= \rho_2\bar{v}_2-\rho_1\bar{v}_1-\rho_2\frac{\epsilon^2}{3}\frac{h_2^2}{h^2}\Delta\bar{v}_2+\rho_1\frac{\epsilon^2}{3}\frac{h_1^2}{h^2}\Delta\bar{v}_1\,.\end{aligned} \tag{4.9}$$

The definition of $\bar{\gamma}_2$, and use of the WNL constraint at leading order,

$$h_1\widetilde{v}_1 = -h_2\widetilde{v}_2+O(\alpha)\,, \tag{4.10}$$

yields

$$\begin{aligned}\bar{\gamma}_2 &= \frac{\rho_2h_1+\rho_1h_2}{h_1}\bar{v}_2+O(\alpha)\\ \bar{\gamma}_2 &= -\frac{\rho_2h_1+\rho_1h_2}{h_2}\bar{v}_1+O(\alpha)\,.\end{aligned} \tag{4.11}$$

This can be inverted and plugged into equation (4.9) (dropping terms of order $\alpha\epsilon^2$) to find

$$\gamma_2 = \bar{\gamma}_2-\frac{\epsilon^2}{3}\frac{h_2h_1(\rho_2h_2+\rho_1h_1)}{h^2(\rho_2h_1+\rho_1h_2)}\Delta\bar{\gamma}_2\,, \tag{4.12}$$

which plainly holds also for non-dimensional variables. The coefficient of the Laplacian in this relation is

$$\frac{h_2h_1(\rho_2h_2+\rho_1h_1)}{h^2(\rho_2h_1+\rho_1h_2)} = \frac{\kappa}{A}\,, \tag{4.13}$$

which yields (4.6). Similar manipulations hold for the first component. This shows that the regularizing variable $\bar{\boldsymbol{\gamma}}$ represents the jump between the averaged momentum density vectors of each fluid.

Remark 4.1. The regularization (4.6) is different from the one used in [1], where the change of variables leading to the shear is done through the choice of the bottom and top layer as reference height of the horizontal velocities in the layers. The resulting model is however still ill-posed with respect to a dispersion critical wavenumber depending on the velocity shear. It has to be remarked that a velocity shear is intrinsic in our choice of the $\boldsymbol{\gamma}$-coordinates, which basically represent a momentum shear, and with the regularizing choice (4.6) all Fourier modes of the linearized WNL equation are stable, avoiding instabilities that would hamper possible numerical approaches to the study of the system. In the $O(\alpha,\epsilon^2)$ asymptotics, all these models are (asymptotically) equivalent, albeit with different dispersion relations that become relevant for large wave numbers $|\boldsymbol{k}|$, that is, beyond the expected applicability ranges of the WNL approximation.

The variable change (4.6) is not necessarily confined to the analysis of the linearized model. Indeed, substituting (4.6) in the full WNL equations (4.1) gives rise, still at $O(\alpha,\epsilon^2)$, to the following non local system

$$\begin{cases}\zeta_t = -A\nabla\cdot\bar{\boldsymbol{\gamma}}-\alpha B\nabla\cdot(\zeta\bar{\boldsymbol{\gamma}})\\ \bar{\boldsymbol{\gamma}}_t-\frac{\epsilon^2\kappa}{3A}\Delta\bar{\boldsymbol{\gamma}}_t = -g'\nabla\zeta-\frac{\alpha}{2}B\nabla|\bar{\boldsymbol{\gamma}}|^2\end{cases} \tag{4.14}$$

supplemented with the additional constraint $\nabla\times\bar{\boldsymbol{\gamma}}=0$.

From the Hamiltonian point of view, the effects of the variable change (4.6) are quite subtle and can be described as follows. Since the (Fréchet) Jacobian of (4.6) is

$$\begin{pmatrix}\mathbf{1} & 0 & 0\\ 0 & \mathbf{1}+\frac{\epsilon^2\kappa}{3A}\Delta & 0\\ 0 & 0 & \mathbf{1}+\frac{\epsilon^2\kappa}{3A}\Delta\end{pmatrix}, \tag{4.15}$$

the Darboux Poisson tensor (3.58) is transformed into the non-canonical one

$$\bar{\mathcal{P}} = -\begin{pmatrix} 0 & \partial_x + \frac{\epsilon^2\kappa}{3A}\Delta\partial_x & \partial_y + \frac{\epsilon^2\kappa}{3A}\Delta\partial_y \\ \partial_x + \frac{\epsilon^2\kappa}{3A}\Delta\partial_x & 0 & 0 \\ \partial_y + \frac{\epsilon^2\kappa}{3A}\Delta\partial_y & 0 & 0 \end{pmatrix}, \tag{4.16}$$

where Δ is the horizontal Laplacian, we used the commutativity properties of the constant-coefficients differential operators, and we fully displayed the tensor to avoid cumbersome notations. To obtain system (4.14) within this Hamiltonian formalism, one has to use the full inversion formula to the first of (4.6), namely

$$\boldsymbol{\gamma} = (\boldsymbol{\mathcal{T}})^{-1}\bar{\boldsymbol{\gamma}} \quad \text{with } \boldsymbol{\mathcal{T}} = \mathbf{1} + \frac{\epsilon^2\kappa}{3A}\Delta\,. \tag{4.17}$$

Under such a transformation, the Hamiltonian (3.59) transforms into the non-local expression

$$\bar{\mathcal{H}} = \frac{1}{2}\int\left((A+\alpha B\zeta)|\boldsymbol{\mathcal{T}}^{-1}(\bar{\boldsymbol{\gamma}})|^2 - \frac{\epsilon^2}{3}\kappa\left(\boldsymbol{\mathcal{T}}^{-1}(\bar{\boldsymbol{\gamma}})\cdot\Delta\boldsymbol{\mathcal{T}}^{-1}(\bar{\boldsymbol{\gamma}})\right) + g'\zeta^2\right)\mathrm{d}x\,\mathrm{d}y\,. \tag{4.18}$$

Further study of this system falls outside the size of this paper, and we leave this to future work.

4.2. Time-independent solutions

We now consider time-independent solutions of the WNL motion equations (3.56). Dropping the time derivatives, and setting the parameters ϵ and α to 1, we get the system

$$\begin{cases} A\nabla\cdot\boldsymbol{\gamma} + B\nabla\cdot(\zeta\boldsymbol{\gamma}) + \dfrac{\kappa}{3}\Delta(\nabla\cdot\boldsymbol{\gamma}) = 0\,, \\ g'\nabla\zeta + \dfrac{B}{2}\nabla|\boldsymbol{\gamma}|^2 = 0\,. \end{cases} \tag{4.19}$$

The second equation for $\nabla\zeta$ gives the explicit dependence of ζ on $|\boldsymbol{\gamma}|$ and the parameter B,

$$\zeta = -\frac{B}{2g'}|\boldsymbol{\gamma}|^2\,, \tag{4.20}$$

under the assumption of zero velocities and interface displacement at infinity.

The coefficient B is not sign definite, and vanishes at the critical ratio $h_1/h_2 = \sqrt{\rho_1/\rho_2}$, so that elevation or depression of the interface depends on these 'hardware' parameters. Indeed, from (4.20), the elevation ζ of the interface has the opposite sign of that of the parameter B in (3.52). Hence, the role of B makes the interface move away from the deeper fluid layer, at this order of approximation. Note that this attractive effect of the 'nearest' horizontal boundary dictated by B is curiously opposite to its counterpart in the travelling wave solutions, well known for the 2D case (see also the next section). Whether this phenomenon of 'nearest-boundary attraction' for this class of nontrivial stationary solutions in the two-layer case is a mere artefact of our WNL asymptotics, or it has a deeper physical meaning, is a problem that deserves a more detailed study, but one that falls beyond the aims of the present paper.

The stationary solutions can also be associated with an explicit expression for the interfacial pressure P. The layer averaged momentum WNL equation (4.4) of [12], specialized to the case of flat bottom (see, also, equation (2.20) of [1]), yields for the lower fluid interface velocity $\widetilde{\boldsymbol{u}} = (\widetilde{u}_2, \widetilde{v}_2)$ the evolution

$$\widetilde{\boldsymbol{u}}_t + (\widetilde{\boldsymbol{u}}\cdot\nabla)\widetilde{\boldsymbol{u}} + g(1-\frac{\rho_1}{\rho_2})\nabla\zeta = -\frac{\nabla P}{\rho_2} + \frac{\rho_1}{\rho_2}(\frac{h_1^2}{2}\nabla(\nabla\cdot\widetilde{\boldsymbol{u}})_t + h_1\zeta_{tt}) + o(\epsilon^2,\alpha)\,. \tag{4.21}$$

Restoring the asymptotic small parameter $\alpha \simeq \epsilon^2$ in the quadratic inertia term, and considering stationary flows, yields the Bernoulli relation

$$P = -(\rho_2-\rho_1)g\zeta - \frac{\alpha}{2}\rho_2(\widetilde{u}_2^2 + \widetilde{v}_2^2)\,, \tag{4.22}$$

since the interface is a streamline. From relations (3.39) we get

$$P = -(\rho_2-\rho_1)g\zeta - \frac{\alpha}{2}\frac{h_1^2\rho_2}{(\rho_2 h_1+\rho_1 h_2)^2}|\boldsymbol{\gamma}|^2\,. \tag{4.23}$$

This relation shows that to leading order the pressure is hydrostatic, i.e. follows the interfacial displacement, and the weighted vorticity contributes a weak, $O(\alpha)$, depressive term.

Remark 4.2. Substituting (4.20) in the first of (4.19) we arrive at the following equation for $\boldsymbol{\gamma}$

$$\nabla\cdot\left(A\boldsymbol{\gamma}-\frac{B^2}{2g'}(|\boldsymbol{\gamma}|^2\boldsymbol{\gamma})+\frac{\kappa}{3}\Delta\boldsymbol{\gamma}\right)=0\,. \tag{4.24}$$

Standing, irrotational wave solutions with frequency A of the defocusing vector Nonlinear Schrödinger equation

$$i\partial_t\boldsymbol{\gamma}-\frac{B^2}{2g'}(|\boldsymbol{\gamma}|^2\boldsymbol{\gamma})+\frac{\kappa}{3}\Delta\boldsymbol{\gamma}=0 \tag{4.25}$$

satisfying suitable boundary conditions provide non-trivial solutions to (4.24), as noticed in the recent paper [24]. Further study on this connection is ongoing.

4.3. Travelling solitary plane wave solutions

By using rotational invariance, we can reduce the plane-travelling wave ansatz for system (3.56) to

$$\zeta(x,y,t)=\zeta(x-ct)\,,\quad \gamma_1(x,y,t)=\gamma_1(x-ct)\,,\quad \gamma_2(x,y,t)=\gamma_2(x-ct)\,. \tag{4.26}$$

Integrating once with respect to x and taking into account the far-field vanishing boundary condition turns the system into

$$\begin{cases}-c\zeta+A\gamma+B\zeta\gamma+\dfrac{\kappa}{3}\gamma_{xx}=0\\ -c\gamma+g'\zeta+\dfrac{B}{2}\gamma^2=0\end{cases} \tag{4.27}$$

for the pair $(\zeta,\gamma\equiv\gamma_1)$, since under the plane-travelling wave ansatz (4.26) the y component γ_2 of the vector $\boldsymbol{\gamma}$ vanishes. This is the travelling wave system associated with the $1+1$ dimensional system of PDEs

$$\begin{cases}\zeta_t+A\gamma_x+B(\zeta\gamma)_x+\frac{\kappa}{3}\gamma_{xxx}=0\\ \gamma_t+g'\zeta_x+B\gamma\gamma_x=0\end{cases}, \tag{4.28}$$

which is the parametric form of the classical $1+1$ dimensional Kaup–Broer–Kupershmidt integrable shallow water equation considered in [10]. Its one-soliton solution can be found in a few simple steps, which we report here for the sake of completeness. Solving the second algebraic equation of (4.27) we get

$$\zeta=\frac{1}{g'}\left(c\gamma-\frac{B}{2}\gamma^2\right), \tag{4.29}$$

so that we obtain for γ the ODE

$$\frac{2g'\kappa}{3}\,\gamma_{xx}=B^2\gamma^3-3c\gamma^2B+2\left(c^2-Ag'\right)\gamma \tag{4.30}$$

equivalent to a Newton Second Law of motion for a point-particle of mass $\mu=\dfrac{2g'\kappa}{3}$ in the potential energy

$$U=\gamma^2\left((Ag'-c^2)+Bc\gamma-\frac{1}{4}B^2\gamma^2\right). \tag{4.31}$$

The first condition to have a travelling wave is the usual one, stating that solitons travel faster than the linear waves $c_0=\sqrt{g'A}$, i.e.

$$c>c_0\,. \tag{4.32}$$

Then we can *a priori* notice that the amplitude of the γ-peak is given by

$$\gamma_m=\frac{2}{B}(c-c_0)\,, \tag{4.33}$$

while the explicit solution expression is

$$\gamma(x,t) = \frac{4\left(c^2 - c_0^2\right)}{B} \frac{\mathrm{e}^{-(x-ct)c_\Delta}}{\left(2c\,\mathrm{e}^{-(x-ct)c_\Delta} + c_0\mathrm{e}^{-2(x-ct)c_\Delta} + c_0\right)}, \tag{4.34}$$

and the corresponding interface elevation solution is given by

$$\zeta(x,t) = \frac{4A\left(c^2 - c_0^2\right)}{c_0 B} \frac{\mathrm{e}^{-(x-ct)c_\Delta}\left(2c_0\,\mathrm{e}^{-(x-ct)c_\Delta} + c\,\mathrm{e}^{-2(x-ct)c_\Delta} + c\right)}{\left(2c\,\mathrm{e}^{-(x-ct)c_\Delta} + c_0\mathrm{e}^{-2(x-ct)c_\Delta} + c_0\right)^2}, \tag{4.35}$$

where $c_\Delta = \sqrt{(c^2 - c_0^2)/\mu}$. This solution is well known from the theory of the KBK-B system. The remarkable fact in the 2-layer setting is that waves are of elevation for $B > 0$ and of depression for $B < 0$, that is, we have solitary waves of elevation ('bright') when $h_1 > \sqrt{\frac{\rho_1}{\rho_2}}\,h_2$, and of depression ('dark') in the opposite case.

The no-contact condition with the plates of the channel translates into

$$\begin{cases} \zeta_m < \frac{h_1}{h} & \text{for } B > 0\,, \\ \zeta_m > -\frac{h_2}{h} & \text{for } B < 0\,, \end{cases} \tag{4.36}$$

where the maximum ζ amplitude is computed from (4.35) as

$$\zeta_m = \frac{2A}{B}\frac{c - c_0}{c_0}\,.$$

Condition (4.36) together with the characteristic condition (4.32) for the existence of solitary wave solutions is satisfied for

$$\begin{cases} c_0 < c < c_0\left(1 + \dfrac{B h_1}{2A\,h}\right) & \text{for } B > 0\,, \\ c_0 < c < c_0\left(1 - \dfrac{B h_2}{2A\,h}\right) & \text{for } B < 0\,, \end{cases} \tag{4.37}$$

whence there exists a range of speeds (and corresponding amplitudes) supporting travelling 'line' solitary wave solutions in the 2-layer case. We remark that, notwithstanding the intrinsic limitations associated with our WNL approximation, the expected appearance of bright ($B > 0$) and dark ($B < 0$) solitary wave is recovered. It should also be noticed that B is small whenever $h_1/h \sim h_2/h = O(1)$, so that the range of admissible wave speeds is quite limited. Moreover the model does not support heteroclinic orbits and hence limiting conjugate-state-like solutions [13].

5. The KP regime

In this section we show how the so called Kadomtsev–Petviashvili equation arises in the process of unidirectionalization of the KBK-Boussinesq system (3.56), by assuming a slow variation in the y-direction and a weak nonlinearity. We will first derive the KP equation directly from the KBK-Boussinesq system, and then equip it with the reduced Hamiltonian structure associated with the asymptotic process with the systematic Dirac reduction technique.

The KP asymptotic regime is obtained by breaking the rotational symmetry of the KBK-B equations, and assuming that the horizontal independent original variables have different scaling laws along the two horizontal directions, namely $x = Lx^*$ and $y = L'y^*$, with $L/L' = O(\epsilon)$. In terms of the rescaled variables we are using we have

$$y' = \beta y \Rightarrow \partial_y = \beta\partial_{y'} \qquad \text{with} \quad \beta = O(\epsilon)\,. \tag{5.1}$$

As a notable consequence of this modified y-scaling the second component γ_2 is to be replaced by

$$\gamma_2' = \frac{1}{\beta}\gamma_2\,, \tag{5.2}$$

owing to the fact that $\boldsymbol{\gamma}$ is irrotational, due to its definition (3.55) and the divergence free condition $\widetilde{\sigma}_x + \widetilde{\tau}_y = 0$.

Under the WNL assumptions the KBK-Boussinesq model (3.56) becomes

$$\begin{cases} \zeta_t + A\gamma_{1\,x} + \beta^2 A\gamma'_{2\,y'} + \alpha B(\zeta\gamma_1)_x + \dfrac{\epsilon^2}{3}\kappa\gamma_{1\,xxx} = 0 \\ \gamma_{1\,t} + g'\zeta_x + \alpha\frac{B}{2}(\gamma_1^2)_x = 0 \\ \beta\gamma'_{2\,t} + \beta g'\zeta_{y'} + \alpha\beta\frac{B}{2}(\gamma_1^2)_{y'} = 0\,. \end{cases} \tag{5.3}$$

This system has to be augmented by the irrotationality condition that now reads $\beta\gamma'_{2\,x} = \beta\gamma_{1\,y'}$. Dropping primes from y-derivatives and γ_2 we finally arrive at

$$\begin{cases} \zeta_t + A\gamma_{1\,x} + \beta^2 A\gamma_{2\,y} + \alpha B(\zeta\gamma_1)_x + \frac{\epsilon^2}{3}\kappa\gamma_{1\,xxx} = 0 \\ \gamma_{1\,t} + g'\zeta_x + \alpha\frac{B}{2}(\gamma_1^2)_x = 0 \\ \gamma_{2\,t} + g'\zeta_y + \alpha\frac{B}{2}(\gamma_1^2)_y = 0 \\ \gamma_{2\,x} = \gamma_{1\,y}\,. \end{cases} \tag{5.4}$$

We now proceed with the derivation of the unidirectional model, similarly to the case of the KdV equation in [32], by seeking a relation between the dependent variables that makes the first two equations of system (5.4) identical. The rationale for this is that the third equation of system (5.4) is simply a compatibility condition on the components of $\boldsymbol{\gamma}$, satisfied at the asymptotic order we are considering.

We thus seek a relation (akin to a Riemann invariant) that keeps track of the contribution of weak nonlinearity and dispersion to the right going linearized restricted system,

$$\begin{cases} \zeta_t + A\gamma_{1\,x} = 0 \\ \gamma_{1\,t} + g'\zeta_x = 0\,, \end{cases} \tag{5.5}$$

as

$$\gamma_1 = \sqrt{\frac{g'}{A}}\zeta + \alpha F(\zeta) + \epsilon^2 G(\zeta) + O(\alpha\epsilon^2, \alpha^2, \epsilon^4, \beta^2)\,, \tag{5.6}$$

where $F(\zeta)$ and $G(\zeta)$ are differential polynomials in the variable ζ. Making the first two equations in (5.4) equivalent up to order $O(\beta^2)$ yields

$$\gamma_1 = \sqrt{\frac{g'}{A}}\zeta - \frac{\alpha}{4}\frac{B}{A}\sqrt{\frac{g'}{A}}\zeta^2 - \frac{\epsilon^2}{6}\frac{\kappa}{A}\sqrt{\frac{g'}{A}}\zeta_{xx} + O(\beta^2)\,. \tag{5.7}$$

Next, we account for the $O(\beta^2)$ terms in the system

$$\begin{cases} \zeta_t + A\gamma_{1\,x} + \beta^2 A\partial_x^{-1}\gamma_{1\,yy} + \alpha B(\zeta\gamma_1)_x + \frac{\epsilon^2}{3}\kappa\gamma_{1\,xxx} = 0 \\ \gamma_{1\,t} + g'\zeta_x + \alpha\frac{B}{2}(\gamma_1^2)_x = 0\,. \end{cases} \tag{5.8}$$

The relation between γ_1 and ζ to make the two equations equivalent at order $O(\beta^2)$ can now be sought as

$$\gamma_1 = \sqrt{\frac{g'}{A}}\zeta - \frac{\alpha}{4}\frac{B}{A}\sqrt{\frac{g'}{A}}\zeta^2 - \frac{\epsilon^2}{6}\frac{\kappa}{A}\sqrt{\frac{g'}{A}}\zeta_{xx} + \beta^2 K(\zeta)\,, \tag{5.9}$$

which, once inserted in (5.8), results in

$$\gamma_1 = \sqrt{\frac{g'}{A}}\left(\zeta - \frac{\alpha}{4}\frac{B}{A}\zeta^2 - \frac{\epsilon^2}{6}\frac{\kappa}{A}\zeta_{xx} - \frac{\beta^2}{2}\partial_x^{-2}\zeta_{yy}\right)\,. \tag{5.10}$$

The final form of the sought-after unidirectional equation, after taking an x-derivative, is thus

$$\left(\zeta_t + \sqrt{Ag'}\zeta_x + \frac{3}{4}\alpha\sqrt{\frac{g'}{A}}B(\zeta^2)_x + \frac{\epsilon^2}{6}\kappa\sqrt{\frac{g'}{A}}\zeta_{xxx}\right)_x + \frac{\beta^2}{2}\sqrt{Ag'}\,\zeta_{yy} = 0\,, \tag{5.11}$$

which is a Kadomtsev–Petviashvili (KP-II) equation with the coefficients determined by the hardware parameters ρ_i, h_i. Note that coefficient B of the nonlinear term can change sign across the critical height relation

$$\frac{h_1}{h_2} = \sqrt{\frac{\rho_1}{\rho_2}}\,. \tag{5.12}$$

5.1. The KP Hamiltonian structure

The KP bi-Hamiltonian structure can be obtained within our scheme by means of a suitable Dirac Hamiltonian reduction [20]. By applying the scaling (5.1), with $\beta \sim \epsilon \sim \sqrt{\alpha}$ to the Hamiltonian functional, taking into account (5.2), and discarding terms of order higher than $\alpha, \epsilon^2, \beta^2$, we obtain the asymptotic Hamiltonian in the KP regime

$$\mathcal{H}_{\mathrm{KP}} = \frac{1}{2}\int \left(A\gamma_1^2 + \beta^2 A\gamma_2^2 + \alpha B\zeta\gamma_1^2 + \frac{\epsilon^2}{3}\kappa\gamma_1\gamma_{1\,xx} + g'\zeta^2 \right) \mathrm{d}x\,\mathrm{d}y, \tag{5.13}$$

whose Hamiltonian equations of motion given by the Poisson operator (3.58) and supplemented by the irrotationality condition $\gamma_{2\,x} = \gamma_{1\,y}$ are

$$\begin{cases} \zeta_t + A\gamma_{1\,x} + \beta^2 A\gamma_{2\,y} + \alpha B(\zeta\gamma_1)_x + \frac{\epsilon^2}{3}\kappa\gamma_{1\,xxx} = 0 \\ \gamma_{1\,t} + g'\zeta_x + \alpha\frac{B}{2}(\gamma_1^2)_x = 0 \\ \gamma_{2\,t} + g'\zeta_y + \alpha\frac{B}{2}(\gamma_1^2)_y = 0 \\ \gamma_{2\,x} = \gamma_{1\,y}\,. \end{cases} \tag{5.14}$$

The system thus obtained coincides with the direct asymptotic expansion (5.4) of the KBK-B model we presented above, a consequence that comes all the way back from the constant nature of the Hamiltonian operator (2.22) from the first two-layer reduction. The idea is to consider the irrotationality condition and the unidirectional constraint in the form given by (5.10) as constraints tying ζ and γ_1, and find the Hamiltonian structure(s) for the KP equation (5.11) by using the extension to the infinite dimensional case of the Dirac reduction procedure of the standard Poisson tensor (3.58). To this end, we recall that the celebrated Dirac reduced bracket [20] for a finite-dimensional Hamiltonian system with coordinates $(y_1, \ldots, y_N)$ subject to M constraints $\Phi_a = 0, a = 1, \ldots, M, M < N$, such that the determinant of the matrix of the Poisson brackets of the constraints $\mathcal{C}_{a,b} = \{\Phi_a, \Phi_b\}$ be invertible, reads

$$\{y_i, y_j\}^D = \{y_i, y_j\} - \sum_{a,b=1}^{M} \{y_i, \Phi_a\}(\mathcal{C}^{-1})_{ab}\{\Phi_b, y_j\}\,. \tag{5.15}$$

It is quickly realized that the constraints Φ_a, $a = 1, \ldots, M$, are Casimir functions for the Dirac bracket, which implies that it restricts directly to the constrained manifold defined by the equations $\Phi_a = 0$, $a = 1, \ldots, M$. This observation suggests that a suitable procedure to obtain the Poisson tensor corresponding to the Dirac bracket (5.15) can be the following:

1. Pick a set of coordinates adapted to the constraints, that is, consider the coordinate set $(y_1, \ldots, y_{N-M}, \Phi_1, \ldots, \Phi_M)$.
2. Write the Poisson tensor P corresponding to the original bracket $\{\cdot,\cdot\}$ in these new coordinates to obtain the new matrix representation

$$\widetilde{P} = \begin{pmatrix} \{y_k, y_l\} & \{y_i, \Phi_a\} \\ \{\Phi_a, y_\ell\} & \{\Phi_a, \Phi_b\} \end{pmatrix} \equiv \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ -\mathcal{B}^{\mathrm{T}} & \mathcal{C} \end{pmatrix}. \tag{5.16}$$

3. Apply formula (5.15) to read the Dirac reduced tensor as

$$\widetilde{P}^D = \begin{pmatrix} \mathcal{A} - \mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B}^{\mathrm{T}} & 0 \\ 0 & 0 \end{pmatrix}, \tag{5.17}$$

whose restriction to the constrained manifold, parametrized by the coordinates $(y_1, \ldots, y_{N-M})$, is given by the $(N-M)\times(N-M)$ upper left block $\mathcal{A} - \mathcal{B}\cdot\mathcal{C}^{-1}\cdot\mathcal{B}^{\mathrm{T}}$.

Such a procedure can be easily implemented (*mutatis mutandis* and with a grain of salt) to the field theoretical model we are herewith studying as follows.

We consider the expression of the constraints (5.10) as defining two new coordinates as

$$\begin{aligned} \Phi_1 &\equiv \gamma_1 - \sqrt{\frac{g'}{A}}\zeta + \frac{\alpha}{4}\frac{B}{A}\sqrt{\frac{g'}{A}}\zeta^2 + \frac{\epsilon^2}{6}\frac{\kappa}{A}\sqrt{\frac{g'}{A}}\zeta_{xx} + \frac{\beta^2}{2}\sqrt{\frac{g'}{A}}\partial_x^{-2}\zeta_{yy} \\ \Phi_2 &\equiv \gamma_{2\,x} - \gamma_{1\,y}\,. \end{aligned} \tag{5.18}$$

As a first step in the Dirac reduction procedure we express the Poisson structure (3.58) in the coordinates (ζ, Φ_1, Φ_2). The Jacobian operator-valued matrix of the transformation $(\zeta, \gamma_1, \gamma_2) \mapsto (\zeta, \Phi_1, \Phi_2)$ is

$$J = \begin{pmatrix} 1 & 0 & 0 \\ \delta_\zeta \Phi_1 & 1 & 0 \\ 0 & -\partial_y & \partial_x \end{pmatrix}, \tag{5.19}$$

where $\delta_\zeta \Phi_1$ is the Fréchet derivative of Φ_1 with respect to ζ, given by

$$\delta_\zeta \Phi_1 = -\frac{g'}{c_0} + \frac{\alpha}{2} \frac{B}{A} \frac{g'}{c_0} \zeta + \frac{\epsilon^2}{6} \frac{\kappa}{A} \frac{g'}{c_0} \partial_{xx} + \frac{\beta^2}{2} \frac{g'}{c_0} \partial_x^{-2} \partial_y^2 . \tag{5.20}$$

Here to improve readability we used the relation $c_0 = \sqrt{A g'}$ tying the speed c_0 of the linear waves to the 'mass' and gravity parameters of the problem.

The Poisson tensor (3.58) is given in terms of the new coordinates (ζ, Φ_1, Φ_2) by

$$\mathcal{P}_{\mathrm{KP}} = J \mathcal{P} J^{\mathrm{T}} = \begin{pmatrix} 0 & -\partial_x & 0 \\ -\partial_x & \mathcal{P}_{22}(\zeta, \partial_x^{-1}, \partial_y) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{5.21}$$

where, explicitly,

$$\begin{aligned} \mathcal{P}_{22}(\zeta, \partial_x^{-1}, \partial_y) &= -(\partial_x (\delta_\zeta \Phi_1)^{\mathrm{T}} + (\delta_\zeta \Phi_1) \partial_x) \\ &= 2 \frac{g'}{c_0} \partial_x - \frac{\alpha}{2} \frac{B}{A} \frac{g'}{c_0} (\partial_x \zeta + \zeta \partial_x) - \frac{\epsilon^2}{3} \frac{\kappa}{A} \frac{g'}{c_0} \partial_{xxx} - \beta^2 \frac{g'}{c_0} \partial_y \partial_x^{-1} \partial_y . \end{aligned} \tag{5.22}$$

From this expression we have that Φ_2 is already a Casimir of $\mathcal{P}_{\mathrm{KP}}$, so that we need to use a Dirac reduction only on the constraint Φ_1, and obtain the reduced Dirac tensor on the constrained manifold defined by $\Phi_1 = 0, \Phi_2 = 0$, parametrized by the free coordinate ζ, by the effective formula

$$\mathcal{P}^D_{\mathrm{KP}} = \partial_x \left(\partial_x \delta_\zeta \Phi_1 + \delta_\zeta \Phi_1 \partial_x \right)^{-1} \partial_x . \tag{5.23}$$

The operator $\partial_x (\delta_\zeta \Phi_1)^{\mathrm{T}} + (\delta_\zeta \Phi_1) \partial_x \equiv \partial_x (\delta_\zeta \Phi_1) + (\delta_\zeta \Phi_1) \partial_x$ turns out to be invertible in the ring of microdifferential operators, with an inverse that at this asymptotic order is given by

$$(\partial_x (\delta_\zeta \Phi_1) + (\delta_\zeta \Phi_1) \partial_x)^{-1} = -\frac{1}{4} \frac{c_0}{g'} \left(2 \partial_x^{-1} + \frac{\alpha}{2} \frac{B}{A} (\partial_x^{-1} \zeta + \zeta \partial_x^{-1}) + \frac{\epsilon^2}{3} \frac{\kappa}{A} \partial_x + \beta^2 \partial_y \partial_x^{-3} \partial_y \right), \tag{5.24}$$

as can be shown by direct computation.

By inserting the inverse operator (5.24) in the expression (5.23) we finally find

$$\mathcal{P}^D_{\mathrm{KP}} = \frac{1}{4} \frac{c_0}{g'} \left(-2 \partial_x - \frac{\alpha}{2} \frac{B}{A} (\zeta \partial_x + \partial_x \zeta) - \frac{\epsilon^2}{3} \frac{\kappa}{A} \partial_{xxx} - \beta^2 \partial_y \partial_x^{-1} \partial_y \right). \tag{5.25}$$

We remark that, consistently collecting the three small parameters $\alpha \sim \epsilon^2 \sim \beta^2$ and suitably rescaling variables to get rid of the hardware parameters A and κ, this expression reduces to

$$\mathcal{P}^D_{\mathrm{KP}} \simeq -2 \partial_x - \alpha \left(\partial_{xxx} + 2 \zeta \partial_x + \zeta_x + \partial_y \partial_x^{-1} \partial_y \right), \tag{5.26}$$

which brings it closer to the notation in the literature for the Poisson pair of the KP equation, viewed as evolutionary non-local partial differential equation in 2-space dimensions (see also [22]), where α plays the role of (inverse) parameter of the Poisson pencil.

As a final step, we have to determine the asymptotic expansion of the constrained Hamiltonian. Enforcing the constraints $\Phi_1 = 0, \Phi_2 = 0$ in the expression of the functional (5.13) we obtain the preliminary form of the constrained Hamiltonian as

$$\mathcal{H}^{(c)}_{\mathrm{KP}} = \int \left(g' \zeta^2 + \frac{\alpha}{4} B \frac{g'^2}{c_0^2} \zeta^3 + \frac{\beta^2}{2} A \frac{g'^2}{c_0^2} \left((\partial_x^{-1} \partial_y \zeta)^2 - \zeta \partial_x^{-1} \partial_y \partial_x^{-1} \partial_y \zeta \right) \right) \mathrm{d}x \, \mathrm{d}y, \tag{5.27}$$

where we used (5.10), and $\gamma_2 = \partial_x^{-1} \partial_y \gamma_1$ so that $\gamma_2 = \dfrac{g'}{c_0} \partial_x^{-1} \partial_y \zeta + O(\alpha, \epsilon^2, \beta^2)$, and the fact that $A g'^2 / c_0^2 \equiv g'$.

Thanks to the symmetry of the operator $\partial_x^{-1}\partial_y$, the $O(\beta^2)$-term in (5.27) can be discarded, to conclude that the restricted Hamiltonian is

$$\mathcal{H}_{\mathrm{KP}}^{(c)} = \int \left(g'\zeta^2 + \frac{\alpha}{4} B \frac{g'^2}{c_0^2} \zeta^3 \right) \mathrm{d}x\,\mathrm{d}y\,. \tag{5.28}$$

As a final check, we note that the variational derivative of the constrained Hamiltonian (5.28) with respect to ζ is

$$\frac{\delta \mathcal{H}_{\mathrm{KP}}^{(c)}}{\delta \zeta} = 2g'\zeta + \alpha \frac{3}{4} \frac{g'^2}{c_0^2} B\zeta^2\,. \tag{5.29}$$

By applying the Dirac-reduced Poisson tensor (5.25) to this differential we obtain

$$\zeta_t + c_0\zeta_x + \alpha \frac{3}{4} B \frac{g'}{c_0} (\zeta^2)_x + \frac{\epsilon^2}{6} \kappa \frac{g'}{c_0} \zeta_{xxx} + \frac{\beta^2}{2} c_0 \partial_x^{-1} \zeta_{yy} = 0\,, \tag{5.30}$$

which, after taking an x-derivative, is exactly the KP-II equation (5.11).

5.2. KP travelling waves in the channel

For the sake of completeness, we shall now derive the solitary wave solution to the KP equation, by means of the travelling-wave ansatz

$$\zeta(x,y,t) = \zeta(x - ct + qy)\,, \tag{5.31}$$

where $x + qy = 0$ is the equation defining the 'line' travelling wave, where q, according to the weak transversal dependence assumption, is small. The KP equation (5.11) transforms into

$$(c_0 - c)\zeta_{xx} + \frac{3}{4}\alpha \frac{g'}{c_0} B(\zeta^2)_{xx} + \frac{\epsilon^2}{6} \frac{g'}{c_0} \kappa \zeta_{xxxx} + \frac{\beta^2}{2} c_0 q^2 \zeta_{xx} = 0\,, \tag{5.32}$$

where we kept the definition of $c_0 = \sqrt{Ag'}$. Setting for simplicity $\alpha = \epsilon^2 = \beta^2 = 1$ and integrating we obtain the usual Newton zero-energy relation for a point mass nonlinear oscillator with a cubic potential of the form

$$\frac{\kappa g'}{12c_0} \zeta_x^2 = -\left(\frac{c_0(1 + \frac{q^2}{2}) - c}{2} \zeta^2 + \frac{1}{4} \frac{g'}{c_0} B\zeta^3 \right)\,. \tag{5.33}$$

Defining by

$$\zeta_m = \frac{2c_0(c - c_0(1 + \frac{q^2}{2}))}{g'B} \tag{5.34}$$

the non-zero root of the cubic potential, we can rewrite the previous equation (5.33) compactly as

$$\zeta_x^2 = \frac{3B}{\kappa} \zeta^2 (\zeta_m - \zeta) \tag{5.35}$$

as in the KdV case (see [32]). Similarly to the full system's solitary wave discussion, for solutions to exist the potential's concavity at the origin has to be negative which forces the speed c to be higher than that of the linearized equation,

$$c > c_0 \left(1 + \frac{q^2}{2} \right)\,. \tag{5.36}$$

As in the plane wave solitons of the full system discussed in section 4.3, the sign of ζ_m depends on the sign of B: when B is positive, we have a soliton of elevation, while when B is negative the soliton is of depression.

The explicit line soliton formula (see, e.g. [32]) is, as well known,

$$\zeta(x,y,t) = \zeta_m \,\mathrm{sech}^2 \left(\sqrt{\frac{3B\zeta_m}{4\kappa}} (x - ct + qy) \right)\,. \tag{5.37}$$

We notice that, thanks to (5.34) and the condition (5.36), the argument of the hyperbolic function is always positive.

A non-trivial restriction on the velocity is however given by the parameters of the problem. Indeed we have to impose *a posteriori* that the solitary wave does not contact the physical boundaries, where, incidentally, the whole formalism breaks down from the very beginning at the level of the parent equations [8]. That is, we have to require

$$\begin{cases} \zeta_m > -\dfrac{h_2}{h} & \text{for } B < 0 \\ \zeta_m < \dfrac{h_1}{h} & \text{for } B > 0\,. \end{cases} \tag{5.38}$$

These conditions translate into the following constraints on the speed c of the soliton:

$$\begin{aligned} &c_0\left(1+\frac{q^2}{2}\right) < c < c_0\left(1+\frac{q^2}{2}\right)\left(1-\frac{h_2 B}{2Ah(1+\frac{q^2}{2})}\right) \quad \text{if} \quad B<0, \\ &c_0\left(1+\frac{q^2}{2}\right) < c < c_0\left(1+\frac{q^2}{2}\right)\left(1+\frac{h_1 B}{2Ah(1+\frac{q^2}{2})}\right) \quad \text{if} \quad B>0\,. \end{aligned} \tag{5.39}$$

Therefore, enforcing these 'engineering' constraints set an upper limit to the soliton speed, and ensures the existence of a range of speeds (and amplitudes) with which KP-II solitons can describe travelling waves in our two-fluid system. The model inherits the same limitations of WNL assumptions briefly addressed at the end of section 4 for the parent model.

6. Conclusions and perspectives

In this paper we studied the problem of sharply stratified 2-layer Euler fluids in 3 dimensions from a Hamiltonian perspective. Our starting point was the Hamiltonian structure for incompressible fluids with general variable density introduced in [7] (a generalization to 3D of that of Benjamin [3] for 2D settings) which we termed Benjamin-Bowman Hamiltonian structure. We considered the sharply stratified two-layer case as a submanifold of the general field configuration and followed MR [27] to a Hamiltonian reduction that yielded: (*i*) the description of the effective free parameters of the theory, i.e. the interface displacement ζ and the components $(\widetilde{\sigma},\widetilde{\tau})$ of the tangential momentum jump, or shear, across the interface (which can be viewed as a density-weighted vorticity supported on the interface), and (*ii*) a simple reduced structure for these Hamiltonian variables.

We then turned to the problem of expressing the energy of the system via these free variables. We worked in the long wave asymptotic limit, in which the ratio ϵ between the typical vertical scales characterized by h and the horizontal ones by L is small. Contrary to the 2D case (see e.g. [25]), it is known that the long wave asymptotics is not enough to obtain a local Hamiltonian density. A local theory, which is the one we focussed here, can be obtained in the WNL approximation, defined by the balance of small parameters $\alpha = a/h = O(\epsilon^2)$, a being a typical wave amplitude. In this case the resulting system coincides with the one known in the literature as the Kaup-Broer-Kupershimdt-Boussinesq (KBK-Boussinesq) model for 3D water waves. In the above approximations, the relevant Hamiltonian variables are the interfacial displacement and the *horizontal* momentum shear vector. An important feature of the two-layer model is that the coefficient

$$B = \frac{\rho_2 h_1^2 - \rho_1 h_2^2}{h\,(\rho_2 h_1 + \rho_1 h_2)}$$

of the nonlinear term in the effective Hamiltonian may change sign according to the 'hardware parameters' of the physical system, viz. the layer densities ρ_j and the asymptotic heights h_j, $j = 1, 2$. While this sign-change of nonlinearity already occurs in 2D, its implications in 3D settings are perhaps more remarkable. This was brought forth by considering the simplest solutions to this model. We derived the dispersion relations and then discussed travelling waves, showing how the sign of B affects solitary wave solutions which, just as in the 2D case, turned out to be wave of depression when $B < 0$ and of elevation for $B > 0$. However, we also briefly touched upon the problem of stationary solutions, which naturally arise for nonlinear equations in more that one space dimensions. Again, the parameter B played a crucial role, but curiously its effects are somehow reversed with respect to the travelling wave solutions. Indeed, since for a stationary solution $\zeta = -B|\boldsymbol{\gamma}|^2/2$, we see that when $B < 0$ the interface is attracted by the upper plate, while for $B > 0$ by the lower plate. Further study of this mathematical phenomenon and its possible physical implications will be addressed in a future work.

We finally considered the so-called KP-regime, obtaining by breaking the $SO(2)$ symmetry of the problem and assuming that the horizontal independent variables (x, y) have different scaling laws along the two horizontal directions, namely $x = Lx^*$ and $y = L'y^*$, with $L/L' = O(\epsilon)$. Unidirectionalization of the resulting model leads to a Kadomtsev–Petviashvili-II equation for the interface elevation ζ, with B dictating the sign of the nonlinear term. Finally, interpreting the unidirectionalization process as a set of constraints on the dependent variables, we show how the Hamiltonian structures of the KP equation can be obtained from that of the KBK-Boussinesq parent system by means of a Dirac reduction process.

In summary, our work has framed the problem of evolution of internal waves in 3D settings from a Hamiltonian perspective. Not unexpectedly, the effective equations in the weakly nonlinear asymptotics reproduce well known models in the theory of water waves, i.e. the KBK-Boussinesq and the KP-II systems. These models share the notable feature, typical of internal waves, of the zero crossing of the coefficient of the nonlinear terms with respect to variations of the densities and asymptotic layer depths, which signals that higher order nonlinearity would come to play a relevant role. From a theoretical point of view, the challenge is to try to go beyond the WNL approximation. In this respect, the mastering of the non localities inherent to the theory, and their framing within the general Hamiltonian reduction scheme herewith presented, is crucial. From a more applicative standpoint, future work will have to assess how the predictions afforded by the models will be reflected, at least qualitatively, by actual experiments and those *in silico*, with attention paid to (quasi) stationary situations.[7]

Acknowledgments

We thank D. Noja for useful discussions. This project was carried out with support by the National Science Foundation under Grants RTG DMS-0943851, CMG ARC-1025523, DMS-1009750, DMS-1517879, DMS-1910824, DMS-2308063, by the Office of Naval Research under Grants N00014-18-1-2490, N00014-23-1-22478 and DURIP N00014-12-1-0749. This project has received fundings by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant no 778010 *IPaDEGAN* and by the PRIN 2022TEB52W-PE1 Project 'The charm of integrability: from nonlinear waves to random matrices.' We also gratefully acknowledge the financial support for RC's visit to Milano-Bicocca by 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 & MP thank the Department of Mathematics and Applications of the University of Milano-Bicocca, ES thanks the Mathematics Department of UNC, and GF thanks SISSA-Trieste for their hospitality where part of this study was carried out. We finally thank the anonymous referees for their comments which helped improve our presentation.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Canonical versus Benjamin-Bowman's Hamiltonian structure

The Benjamin-Bowman Hamiltonian structure we used in this paper can be seen as the canonical Zakharov–Kuznetsov (ZK) Hamiltonian structure [31, 35, 36] for the Euler equations. Indeed, the ZK approach makes use, along with the density variable ρ, of the so-called Clebsch potentials(Φ, λ, μ), that define the Eulerian velocity field by

$$\boldsymbol{v} = \frac{\lambda}{\rho}\,\boldsymbol{\nabla}\mu + \boldsymbol{\nabla}\Phi\,. \tag{A.1}$$

[7] Note added in proof: After submission of this paper, we became aware of the relevant reference by Holm D D, Hu R and Wang H, 2026 Surface Wave Solutions in 1D and 2D for the Broer–Kaup–Boussinesq–Kupershmidt System *Stud. Appl. Math.* https://doi.org/10.1111/sapm.70165

In the ZK representation the pairs (ρ,Φ) and (λ,μ) are canonically conjugated, so that in the 4-tuple (ρ,Φ,λ,μ) (which we shall later denote as $(\rho,\boldsymbol{C})$, i.e.$(C^1,C^2,C^3)=(\Phi,\lambda,\mu)$) the Poisson tensor is the canonical one,

$$J_{\mathrm{can}}=\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}. \tag{A.2}$$

The Benjamin-Bowman variables $(\rho,\boldsymbol{\Sigma})$ are related with the Clebsch potentials by

$$\boldsymbol{\Sigma}=\boldsymbol{\nabla}\times(\rho\boldsymbol{v})=\boldsymbol{\nabla}\rho\times\boldsymbol{\nabla}\Phi+\boldsymbol{\nabla}\lambda\times\boldsymbol{\nabla}\mu. \tag{A.3}$$

Proposition A.1. *Let $\boldsymbol{F}$ denote the change of coordinates*

$$(\rho,\Phi,\lambda,\mu)\xrightarrow{\boldsymbol{F}}(\rho,\boldsymbol{\Sigma}), \tag{A.4}$$

with $\boldsymbol{\Sigma}$ given by (A.3). Then the canonical Poisson tensor (A.2) and the Benjamin-Bowman Poisson tensor (2.6), i.e.

$$P_B=-\begin{pmatrix}0&\boldsymbol{\nabla}\cdot(\rho\boldsymbol{\nabla}\times)\\\boldsymbol{\nabla}\times(\rho\boldsymbol{\nabla}\cdot)&\boldsymbol{\nabla}\times(\boldsymbol{\Sigma}\times\boldsymbol{\nabla}\times)\end{pmatrix}\begin{pmatrix}\dfrac{\delta\mathcal{H}}{\delta\rho}\\\dfrac{\delta\mathcal{H}}{\delta\boldsymbol{\Sigma}}\end{pmatrix} \tag{A.5}$$

are $\boldsymbol{F}$-related.

Proof. The proposition can be proven by tedious but straightforward computations, whose key points are collected below.

Denoting by $\boldsymbol{F}_*$ the (Fréchet) Jacobian of the map (A.4), and with $\boldsymbol{F}^*$ its transpose, we have to prove the equality

$$P_B=\boldsymbol{F}_*\cdot J_{\mathrm{can}}\cdot\boldsymbol{F}^*, \tag{A.6}$$

where $\cdot$ stands for both matrix multiplication and composition of differential operators.

We have (using Einstein's convention of summation over repeated indices that run from 1 to 3) that the Jacobian matrix is

$$\boldsymbol{F}_*=\begin{pmatrix}1&0&0&0\\\epsilon^{1lm}\partial_m\Phi\partial_l&\epsilon^{1lm}\partial_l\rho\partial_m&\epsilon^{1lm}\partial_m\mu\partial_l&\epsilon^{1lm}\partial_l\lambda\partial_m\\\epsilon^{2lm}\partial_m\Phi\partial_l&\epsilon^{2lm}\partial_l\rho\partial_m&\epsilon^{2lm}\partial_m\mu\partial_l&\epsilon^{2lm}\partial_l\lambda\partial_m\\\epsilon^{3lm}\partial_m\Phi\partial_l&\epsilon^{3lm}\partial_l\rho\partial_m&\epsilon^{3lm}\partial_m\mu\partial_l&\epsilon^{3lm}\partial_l\lambda\partial_m\end{pmatrix} \tag{A.7}$$

and, thanks to the antisymmetry of the Levi-Civita ϵ symbol, its transpose is

$$\boldsymbol{F}^*=\begin{pmatrix}1&-\epsilon^{1lm}\partial_m\Phi\partial_l&-\epsilon^{2lm}\partial_m\Phi\partial_l&-\epsilon^{3lm}\partial_m\Phi\partial_l\\0&-\epsilon^{1lm}\partial_l\rho\partial_m&-\epsilon^{2lm}\partial_l\rho\partial_m&-\epsilon^{3lm}\partial_l\rho\partial_m\\0&-\epsilon^{1lm}\partial_m\mu\partial_l&-\epsilon^{2lm}\partial_m\mu\partial_l&-\epsilon^{3lm}\partial_m\mu\partial_l\\0&-\epsilon^{1lm}\partial_l\lambda\partial_m&-\epsilon^{2lm}\partial_l\lambda\partial_m&-\epsilon^{3lm}\partial_l\lambda\partial_m\end{pmatrix}. \tag{A.8}$$

Carefully considering the operator multiplications and using the well known identity

$$\epsilon^{lmk}\epsilon^{kij}=\delta^{li}\delta^{mj}-\delta^{lj}\delta^{mi}, \tag{A.9}$$

the assertion follows.

Appendix B. The Hamiltonian reduction procedure

In this section we give some details on the reduction procedure presented in section 2.1. We consider the submanifold $\mathcal{S}\subset\mathcal{M}$, see (2.18) and (2.16), and we restrict the Poisson tensor P_B, see (2.17), on $\mathcal{S}$.

Our first task is to compute the annihilator $(T\mathcal{S})^0$. It is easily seen that it coincides with the kernel of $i^*: T^*\mathcal{M}\to T^*\mathcal{S}$ at the points $(\zeta,\widetilde{\sigma},\widetilde{\tau})\in\mathcal{S}$, where $i:\mathcal{S}\to\mathcal{M}$ is the canonical injection. Since the differential $i_*: T\mathcal{S}\to T\mathcal{M}$ is given by

$$
\begin{aligned}
(\dot\zeta,\dot{\widetilde{\sigma}},\dot{\widetilde{\tau}})\mapsto\Big((\rho_2-\rho_1)\dot\zeta\delta(z-\zeta),\quad \dot{\widetilde{\sigma}}\delta(z-\zeta)-\widetilde{\sigma}\dot\zeta\delta'(z-\zeta),\quad \dot{\widetilde{\tau}}\delta(z-\zeta)-\widetilde{\tau}\dot\zeta\delta'(z-\zeta),\\
(\dot{\widetilde{\sigma}}\zeta_x+\dot{\widetilde{\tau}}\zeta_y+\widetilde{\sigma}\dot\zeta_x+\widetilde{\tau}\dot\zeta_y)\delta(z-\zeta)-\dot\zeta(\widetilde{\sigma}\zeta_x+\widetilde{\tau}\zeta_y)\delta'(z-\zeta)\Big),
\end{aligned}
\tag{B.1}
$$

one can check that the image of $(\alpha_\rho,\alpha_\sigma,\alpha_\tau,\alpha_\chi)\in T^*\mathcal{M}$ under i^* is the covector $(\mu_\zeta,\mu_{\widetilde{\sigma}},\mu_{\widetilde{\tau}})$, where

$$\mu_\zeta=(\rho_2-\rho_1)\widetilde{\alpha_\rho}+\widetilde{\sigma}\left(\widetilde{(\alpha_\sigma)_z}-\widetilde{(\alpha_\chi)_x}\right)+\widetilde{\tau}\left(\widetilde{(\alpha_\tau)_z}-\widetilde{(\alpha_\chi)_y}\right)\tag{B.2}$$

$$\mu_{\widetilde{\sigma}}=\widetilde{\alpha_\sigma}+\zeta_x\widetilde{\alpha_\chi}\tag{B.3}$$

$$\mu_{\widetilde{\tau}}=\widetilde{\alpha_\tau}+\zeta_y\widetilde{\alpha_\chi}\tag{B.4}$$

and, as usual, we denoted with a tilde the evaluation $z=\zeta(x,y)$, i.e. the restriction to the interface between the two fluids. Hence we can conclude that $(\alpha_\rho,\alpha_\sigma,\alpha_\tau,\alpha_\chi)$ belongs to $(T\mathcal{S})^0=\operatorname{Ker}i^*$ if and only if the conditions

$$(\rho_2-\rho_1)\widetilde{\alpha_\rho}+\widetilde{\sigma}\left(\widetilde{(\alpha_\sigma)_z}-\widetilde{(\alpha_\chi)_x}\right)+\widetilde{\tau}\left(\widetilde{(\alpha_\tau)_z}-\widetilde{(\alpha_\chi)_y}\right)=0\tag{B.5}$$

$$\widetilde{\alpha_\sigma}+\zeta_x\widetilde{\alpha_\chi}=0\tag{B.6}$$

$$\widetilde{\alpha_\tau}+\zeta_y\widetilde{\alpha_\chi}=0\tag{B.7}$$

are fulfilled.

Now we show that $\mathcal{P}_B(T\mathcal{S})^0=\{0\}$, i.e. the kernel of $\mathcal{P}_B$ contains $(T\mathcal{S})^0$ at every point of $\mathcal{S}$. Let $(\dot\rho,\dot\sigma,\dot\tau,\dot\chi)$ be the image of $(\alpha_\rho,\alpha_\sigma,\alpha_\tau,\alpha_\chi)\in(T\mathcal{S})^0$ under $\mathcal{P}_B$. Then, using the expression (2.17) of $\mathcal{P}_B$ at the points of $\mathcal{S}$ yields

$$
\begin{aligned}
\dot\rho&=(\rho_1-\rho_2)\left[\zeta_x\left((\alpha_\chi)_y-(\alpha_\tau)_z\right)+\zeta_y\left((\alpha_\sigma)_z-(\alpha_\chi)_x\right)-(\alpha_\tau)_x+(\alpha_\sigma)_y\right]\delta(z-\zeta)\\
&=(\rho_1-\rho_2)\left[(\widetilde{\alpha_\sigma}+\zeta_x\widetilde{\alpha_\chi})_y-(\widetilde{\alpha_\tau}+\zeta_y\widetilde{\alpha_\chi})_x\right]\delta(z-\zeta)\\
&=(\rho_1-\rho_2)\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta(z-\zeta),
\end{aligned}
\tag{B.8}
$$

where we used (B.3) and (B.4) in the last equality. Then (B.6) and (B.7) imply that $\dot\rho=0$. With similar computations, one can show that

$$
\begin{aligned}
\dot\sigma&=-(\mu_\zeta)_y\delta(z-\zeta)+\sigma\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta'(z-\zeta)\\
\dot\tau&=(\mu_\zeta)_x\delta(z-\zeta)+\tau\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta'(z-\zeta)\\
\dot\chi&=\left[\zeta_y(\mu_\zeta)_x-\zeta_x(\mu_\zeta)_y\right]\delta(z-\zeta)+\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right](\widetilde{\sigma}\zeta_x+\widetilde{\tau}\zeta_y)\delta'(z-\zeta),
\end{aligned}
\tag{B.9}
$$

entailing that $\dot\sigma=\dot\tau=\dot\chi=0$ as a consequence of (B.5), (B.6), and (B.7). Hence we have shown that (2.21) holds, so that the reduced Poisson manifold coincides with $\mathcal{S}$, which is parametrized by the triple $(\zeta,\widetilde{\sigma},\widetilde{\tau})$.

The final step is to find the reduced Poisson tensor on $\mathcal{S}$. To this aim, we will use the projection $\pi:\mathcal{M}\to\mathcal{S}$ given by (2.20), seeing $\mathcal{S}$ as a quotient manifold rather than a submanifold of $\mathcal{M}$. We know that $(\alpha_\rho,\alpha_\sigma,\alpha_\tau,\alpha_\chi)\in T^*\mathcal{M}$ is an extension of $(\mu_\zeta,\mu_{\widetilde{\sigma}},\mu_{\widetilde{\tau}})$, i.e. one of its preimages of under i^*, if and only if (B.2), (B.3), and (B.4) hold. We also know — see (B.8) and the first two equations in (B.9) — that the evaluation of P_B on such a covector in $T^*\mathcal{M}$ gives

$$\begin{pmatrix}\dot\rho\\ \dot\sigma\\ \dot\tau\end{pmatrix}=\begin{pmatrix}(\rho_1-\rho_2)\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta(z-\zeta)\\ -(\mu_\zeta)_y\delta(z-\zeta)+\sigma\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta'(z-\zeta)\\ (\mu_\zeta)_x\delta(z-\zeta)+\tau\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta'(z-\zeta)\end{pmatrix}.$$

Applying the differential of π, we obtain

$$\begin{aligned}\dot{\zeta} &= \frac{1}{\rho_2-\rho_1}\int_{-h_2}^{h_1}\dot{\rho}\,\mathrm{d}z = -(\mu_{\widetilde{\sigma}})_y + (\mu_{\widetilde{\tau}})_x \\ \dot{\widetilde{\sigma}} &= \int_{-h_2}^{h_1}\dot{\sigma}\,\mathrm{d}z = -\int_{-h_2}^{h_1}(\mu_\zeta)_y\delta(z-\zeta)\,\mathrm{d}z + \int_{-h_2}^{h_1}\sigma\left[(\mu_{\widetilde{\sigma}})_y-(\mu_{\widetilde{\tau}})_x\right]\delta'(z-\zeta)\,\mathrm{d}z = -(\mu_\zeta)_y,\end{aligned} \tag{B.10}$$

where the last integral in the second line vanishes since the interface does not touch the plates, i.e. $-h_2 < \zeta < h_1$. In the same way, one finds that

$$\dot{\widetilde{\tau}} = (\mu_\zeta)_x, \tag{B.11}$$

so that the reduced Poisson tensor on $\mathcal{S}$ is given by (2.22).

Appendix C. On the Dirichlet–Neumann (DN) operator approach

In this appendix we compare the approach through the DN operator of [17] and [5] (see, also, e.g. [15, 18, 19]) to the computations of section 3.3.

Following the seminal approach of [34] one can write the Euler equations for the two-layer case through the potentials $\Phi_i(x,y,z)$,

$$\boldsymbol{\Delta}\Phi_i = 0 \quad \text{in} \quad \Omega_i\,, \tag{C.1}$$

$$\Phi_{i\,t} + \frac{1}{2}|\boldsymbol{\nabla}\Phi_i|^2 = -\frac{p}{\rho_i} - gz \quad \text{in} \quad \Omega_i\,, \tag{C.2}$$

where Ω_i is the i-th fluid domain, $\boldsymbol{\Delta} = \partial_x^2+\partial_y^2+\partial_z^2$, $\boldsymbol{\nabla} = (\partial_x,\partial_y,\partial_z)$, equipped with the following set of boundary conditions

$$\Phi_{i\,z} = 0 \quad \text{at} \quad z = (-1)^{i-1}h_i\,, \tag{C.3}$$

$$\zeta_t = \sqrt{1+\zeta_x^2+\zeta_y^2}\;\partial_n\Phi_2 \quad \text{at} \quad z=\zeta\,, \tag{C.4}$$

$$\partial_n\Phi_1 = \partial_n\Phi_2 \quad \text{at} \quad z=\zeta\,, \tag{C.5}$$

where ∂_n is the normal derivative at the interface associated with the lower fluid,

$$\mathbf{n} = \frac{(-\zeta_x,-\zeta_y,1)}{\sqrt{1+|(\zeta_x,\zeta_y)|^2}}\,, \qquad \partial_n = \mathbf{n}\cdot\boldsymbol{\nabla}\,. \tag{C.6}$$

'Trace' variables are defined as $\widetilde{\Phi}_i = \Phi_i|_{z=\zeta}$. The starting point of the DN approach relies on the observation (see, e.g. [5]) that, once given $\widetilde{\Phi}_1(x,y)$, we can formally find both the bulk potentials Φ_i's. Indeed, knowing $\widetilde{\Phi}_1$, we can solve the following Laplace problem with mixed boundary conditions for Φ_1,

$$\begin{cases}\boldsymbol{\Delta}\Phi_1 = 0 & \text{in} \quad \Omega_1\\ \Phi_{1\,z} = 0 & \text{at} \quad z = h_1\\ \Phi_1 = \widetilde{\Phi}_1 & \text{at} \quad z=\zeta\,.\end{cases} \tag{C.7}$$

Then, using the boundary condition (C.5), we can consider the Neumann boundary value problem

$$\begin{cases}\boldsymbol{\Delta}\Phi_2 = 0 & \text{in} \quad \Omega_2\\ \Phi_{2\,z} = 0 & \text{at} \quad z = -h_2\\ \partial_n\Phi_2 = \partial_n\Phi_1 & \text{at} \quad z=\zeta\,.\end{cases} \tag{C.8}$$

The compatibility condition of this Neumann problem is satisfied thanks to (C.7), so that it has a unique solution up to constants.

It is known that the problem can be formulated within a canonical formalism, the Hamiltonian of the problem formally being given by the analogue of the 2-dimensional counterpart found in [17], that is,

$$\mathcal{H}[\zeta,\widetilde{\Phi}_1,\widetilde{\Phi}_2] = \frac{1}{2}\int_{\mathbb{R}^2}\left(\rho_1\widetilde{\Phi}_1 G_1[\zeta]\widetilde{\Phi}_1 + \rho_2\widetilde{\Phi}_2 G_2[\zeta]\widetilde{\Phi}_2 + g(\rho_2-\rho_1)\zeta^2\right)\mathrm{d}x\,\mathrm{d}y\,, \tag{C.9}$$

where $G_j[\zeta]$ are the Dirichlet to Neumann operators defined via the relations

$$G_1[\zeta]\widetilde{\Phi}_1 = -\partial_n \Phi_{1|_{z=\zeta}}, \quad G_2[\zeta]\widetilde{\Phi}_2 = \partial_n \Phi_{2|_{z=\zeta}}, \tag{C.10}$$

the minus sign coming from the orientation of the interface.

To fully describe the problem in terms of free canonical variables, one needs to find the explicit relation between $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$. Namely, in the approach of [17], one has to consider the Laplace problems associated with both fluids (C.7) and (C.8) and formulate the problem in terms of the variables ζ and $\xi = \rho_2\widetilde{\Phi}_2 - \rho_1\widetilde{\Phi}_1$ of [4].

Using the boundary condition (C.5) one arrives at finding the constraint

$$\nabla\widetilde{\Phi}_2 = -\nabla\big(G_2[\zeta]^{-1} \circ G_1[\zeta](\widetilde{\Phi}_1)\big). \tag{C.11}$$

As it customary, one introduces the small parameters α, ϵ as in section 3.1 in order to find asymptotic expansions of the DN operators, and, in particular, of the constraint (C.11).

In this asymptotics, the system (C.7) becomes

$$\begin{cases} \epsilon^2\Delta\Phi_1 + \partial_z^2\Phi_1 = 0 & \text{in} \quad \Omega_1 \\ \Phi_{1z} = 0 & \text{at} \quad z = \frac{h_1}{h} \\ \Phi_1 = \widetilde{\Phi}_1 & \text{at} \quad z = \alpha\zeta, \end{cases} \tag{C.12}$$

a similar formulation being given for (C.8).

As proven in [16, 18], $G_1[\zeta]$ and $G_2[\zeta]$ admit a series expansions in ζ, so that, taking our scalings into account, we have

$$G_i[\alpha\zeta] = \sum_{j=0}^{\infty} \alpha^j G_i^j[\zeta]. \tag{C.13}$$

By using the Fourier transform on the first equation of (C.12) and on the first equation of the corresponding system to (C.8), working in the WNL asymptotics we get

$$G_i[\zeta] = (-\epsilon^2\frac{h_i}{h})\left(\Delta + \frac{\epsilon^2}{3}\frac{h_i^2}{h^2}\Delta^2 + (-1)^i\alpha\frac{h}{h_i}\nabla\cdot(\zeta\nabla) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2)\right), \tag{C.14}$$

where it might be useful to remark that the operator $\nabla\cdot(\zeta\nabla)$ acting on a function F, stands for the divergence of the pointwise product of ζ with the gradient of F.

The inversion of $G_2[\zeta]$, even if in the WNL asymptotics, introduces a non-local term in the picture. Indeed, we can formally write

$$G_2[\zeta]^{-1} = \left(Id - \frac{\epsilon^2}{3}\frac{h_2^2}{h^2}\Delta - \alpha\frac{h}{h_2}\Delta^{-1}\nabla\cdot(\zeta\nabla) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2)\right)(-\frac{1}{\epsilon^2}\frac{h}{h_2}\Delta^{-1}), \tag{C.15}$$

whence

$$\begin{aligned} G_2[\zeta]^{-1} \circ G_1[\zeta] &= \frac{h_1}{h_2}\left(Id - \frac{\epsilon^2}{3}\frac{h_2^2}{h^2}\Delta - \alpha\frac{h}{h_2}\Delta^{-1}\nabla\cdot(\zeta\nabla) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2)\right)\circ \\ &\circ\left(Id + \frac{\epsilon^2}{3}\frac{h_1^2}{h^2}\Delta - \alpha\frac{h}{h_1}\Delta^{-1}\nabla\cdot(\zeta\nabla) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2)\right) \\ &= \frac{h_1}{h_2} + \frac{\epsilon^2}{3}\frac{h_1}{h_2}(\frac{h_1^2}{h} - \frac{h_2^2}{h})\Delta - \alpha(\frac{h}{h_1} + \frac{h}{h_2})\Delta^{-1}\nabla\cdot(\zeta\nabla) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2). \end{aligned} \tag{C.16}$$

Substituting in the constraint equation (C.11) we get, in the notations of [5], the final expression of the constraint as

$$\nabla\widetilde{\Phi}_2 = -\frac{h_1}{h_2}\nabla\widetilde{\Phi}_1 - \frac{\epsilon^2}{3}\frac{h_1}{h_2}\frac{h_1^2 - h_2^2}{h^2}\Delta\nabla\widetilde{\Phi}_1 + \alpha h\frac{h_1}{h_2}(\frac{1}{h_1} + \frac{1}{h_2})\Pi(\zeta\nabla\widetilde{\Phi}_1) + O(\alpha^2, \epsilon^4, \alpha\epsilon^2), \tag{C.17}$$

where Π is the non-local operator

$$\Pi := \Delta^{-1}(\nabla(\nabla\cdot\,)). \tag{C.18}$$

We can explicitly recover relation (3.31) of section 3.3 expressing the interface velocities of the upper fluid in terms of those of the lower fluid as follows. First we recall that

$$\widetilde{\boldsymbol{u}}_i = \nabla\widetilde{\Phi}_i + O(\alpha\epsilon^2)\,, \tag{C.19}$$

so that (C.17) translates into

$$\widetilde{\boldsymbol{u}}_2 = -\frac{h_1}{h_2}\widetilde{\boldsymbol{u}}_1 - \frac{\epsilon^2}{3}\frac{h_1}{h_2}\frac{h_1^2-h_2^2}{h^2}\Delta\widetilde{\boldsymbol{u}}_1 + \alpha h\frac{h_1}{h_2}\left(\frac{1}{h_1}+\frac{1}{h_2}\right)\Pi(\zeta\widetilde{\boldsymbol{u}}_1) + O(\alpha^2,\epsilon^4,\alpha\epsilon^2)\,. \tag{C.20}$$

Then we use the identity

$$\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot(\zeta\widetilde{\boldsymbol{u}}_1)) = \Delta(\zeta\widetilde{\boldsymbol{u}}_1) + \boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\zeta\widetilde{\boldsymbol{u}}_1) = \Delta(\zeta\widetilde{\boldsymbol{u}}_1) + \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\zeta\times\widetilde{\boldsymbol{u}}_1) + O(\alpha\epsilon^2) \tag{C.21}$$

(the second one following from (C.19)). Since Δ^{-1} is a scalar operator and only the first two components of the curl need to be considered for the definition of Π, we can write at the WNL order

$$\Pi(\zeta\widetilde{\boldsymbol{u}}_1) = \zeta\widetilde{\boldsymbol{u}}_1 + \boldsymbol{\nabla}\times\Delta^{-1}(\boldsymbol{\nabla}\zeta\times\widetilde{\boldsymbol{u}}_1)\,, \tag{C.22}$$

yielding the desired equivalence between (C.17) and (3.31).

ORCID iDs

R Camassa 0000-0002-1607-5950
G Falqui 0000-0002-4893-9186
G Ortenzi 0000-0003-2192-6737
M Pedroni 0000-0002-7358-0945
E Sforza 0009-0008-4215-1540

References

[1] Barros R and Choi W 2013 On regularizing the strongly nonlinear model for two-dimensional internal waves *Physica D* **264** 27–34
[2] Barros R, Gavrilyuk S L and Teshukov V M 2007 Dispersive nonlinear waves in two-layer flows with free surface. I. Model derivation and general properties *Stud. Appl. Math.* **119** 191–211
[3] Benjamin T B 1986 On the Boussinesq model for two-dimensional wave motions in heterogeneous fluids *J. Fluid Mech.* **165** 445–74
[4] Benjamin T B and Bridges T J 1997 Reappraisal of the Kelvin-Helmholtz problem. Part 1. Hamiltonian structure *J. Fluid Mech.* **333** 301–25
[5] Bona J, Lannes D and Saut J-C 2008 Asymptotic models for internal waves *J. Math. Pures Appl.* **89** 538–66
[6] Bleck R 2002 An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates *Ocean Mod.* **4** 55–88
[7] Bowman S 1987 A note on Hamiltonian structure *Math. Proc. Camb. Phil. Soc.* **102** 173–89
[8] Camassa R, Chen S, Falqui G, Ortenzi G and Pedroni M 2014 Topological selection in stratified fluids: an example from air-water systems *J. Fluid Mech.* **743** 534–53
[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] Camassa R, Falqui G, Ortenzi G, Pedroni M and Vu Ho T T 2023 Simple two-layer dispersive models in the Hamiltonian reduction formalism *Nonlinearity* **36** 4523–52
[11] Camassa R and Levermore C D 1997 Layer-mean quantities, local conservation laws and vorticity *Phys. Rev. Lett.* **78** 650–3
[12] Choi W and Camassa R 1996 Weakly nonlinear internal waves in a two-fluid system *J. Fluid Mech.* **313** 83–103
[13] Choi W and Camassa R 1999 Fully nonlinear internal waves in a two-fluid system *J. Fluid Mech.* **396** 1–36
[14] Chumakova L, Mezanque 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
[15] Craig W and Groves M D 1994 Hamiltonian long-wave approximations to the water-wave problem *Wave Motion* **19** 367–89
[16] Coifman R R and Meyer Y 1985 Nonlinear Harmonic Analysis and Analytic Dependence *Pseudodifferential operators and applications* Proc. Symp. Pure Math ed Tréves François vol 43 pp 71–78 (AMS, Providence RI)
[17] Craig W, Guyenne P and Kalish H 2005 Hamiltonian long wave expansions for free surfaces and interfaces *Commun. Pure Appl. Math.* **58** 1587–641
[18] Craig W, Schanz U and Sulem C 1997 The modulational regime of three-dimensional water waves and the Davey-Stewartson system *Ann. Inst. Henri Poincaré C* **14** 615–67
[19] Craig W, Sulem C and Sulem P L 1992 Nonlinear modulation of gravity waves: a rigorous approach *Nonlinearity* **5** 497–522
[20] Dirac P A M 1950 Generalized Hamiltonian dynamics *Canad. J. Math.* **2** 129–48
[21] Falqui G and Sforza E Hamiltonian reductions, scalings, and effective wave models in stratified fluids (arXiv:2505.15225)
[22] Fokas A S and Santini P M 1988 Recursion operators and bi-Hamiltonian structures in multidimensions I *Commun. Math. Phys.* **115** 375–419
[23] Homsy G M 1987 Viscous fingering in porous media *Ann. Rev. Fluid Mech.* **19** 271–311
[24] Klein C and Saut J-C 2025 On the Kaup-Broer-Kupershmidt systems *EMS Surv. Math. Sci.* **12** 215–42

[25] Lannes D 2013 The water waves problem. Mathematical analysis and asymptotics *Math. Surveys Monogr.* vol 188 (American Mathematical Society)

[26] Lucas D, Beyer M, Pietruske H and Szalinski L 2017 Counter-current flow limitation for air-water and steam-water flows in a PWR hot leg geometry *Nucl. Eng. Des.* **323** 56–67

[27] Marsden J E and Ratiu T S 1986 Reduction of Poisson manifolds *Lett. Math. Phys.* **11** 161–9

[28] Miles J and Salmon R 1985 Weakly dispersive nonlinear gravity waves *J. Fluid Mech.* **157** 519–31

[29] Pedlosky J 1986 *Geophysical Fluid Dynamics* (Springer)

[30] Percival J R, Cotter C J and Holm D D 2008 A Euler–Poincaré framework for the multilayer Green–Nagdhi equations *J. Phys. A: Math. Theor.* **41** 344018

[31] Salmon R 1988 Hamiltonian fluid mechanics *Ann. Rev. Fluid Mech.* **20** 225–56

[32] Whitham G B 2000 *Linear and Nonlinear Waves* (Wiley)

[33] Wu T Y 1999 On modeling unsteady fully nonlinear dispersive interfacial waves *Fluid Dynamics at Interfaces* ed W Shyy and R Narayanan (Cambridge University Press) pp 171–8

[34] Zakharov V E 1968 Stability of periodic waves of finite amplitude on the surface of a deep fluid *J. Appl. Mech. Tech. Phys.* **9** 86–94

[35] Zakharov V E and Kuznetsov E A 1997 Hamiltonian formalism for nonlinear waves *Phys.Usp.* **40** 1087–116

[36] 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