

**Subject Areas:**

applied mathematics, fluid mechanics, wave motion

Keywords:

waves in fluids, internal waves, nonlinear stability

Author for correspondence:

Francisco de Melo Viríssimo

e-mail:

f.de.melo.virissimo@noc.ac.uk

Nonlinear stability of two-layer shallow water flows with a free surface

Francisco de Melo Viríssimo^{1,2} and Paul A. Milewski¹¹Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK²Now at: National Oceanography Centre, European Way, Southampton, SO14 3HZ, UK

The problem of two layers of immiscible fluid, bordered above by an unbounded layer of passive fluid and below by a flat bed, is formulated and discussed. The resulting equations are given by a first order, four-dimensional system of PDEs of mixed-type. The relevant physical parameters in the problem are presented and used to write the equations in a non-dimensional form. The conservation laws for the problem, which are known to be only six, are explicitly written and discussed in both non-Boussinesq and Boussinesq cases. Both dynamics and nonlinear stability of the Cauchy problem are discussed, with focus on the case where the upper unbounded passive layer has zero density, also called the free surface case. We prove that the stability of a solution depends only on two “baroclinic” parameters (the shear and the difference of layer thickness, the former being the most important one) and give a precise criterion for the system to be well-posed. It is also numerically shown that the system is nonlinearly unstable, as hyperbolic initial data evolves into the elliptic region before the formation of shocks. We also discuss the use of simple waves as a tool to bound solutions and preventing a hyperbolic initial data to become elliptic and use this idea to give a mathematical proof for the nonlinear instability.

1. Introduction

Internal waves are a major topic of scientific interest [1], playing a key role in climate and weather studies [2], [3], and mathematical models and their analysis play an important part in their understanding [4], [5]. Internal waves can, in certain cases be modelled as interfacial

© The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

waves, which are waves propagating due to the difference in density between layers of fluid. This approach simplifies the problem substantially and yet may be used to model well the horizontal propagation of disturbances.

The correspondence between interfacial (or ‘layered’) models and the more geophysically relevant continuous stratification case is clear when the stratification has sharp pycnoclines but is also used more broadly based on the idea of an “effective layer depth”. This is when the layer depths are chosen to match the wave-speeds of the first few modes arising from the physical stratification. In general, layered flow models in the geophysical context are considered in the long wave (or ‘shallow water’) limit, where horizontal length scales are much larger than layer depths. The resulting models can be either weakly or strongly nonlinear, and dispersive or not. KdV equations arise in the balance of weak dispersion and weak nonlinearity [6], [7], whereas MCC type equations arise in the fully nonlinear, weakly dispersive case [8], [9]. Strongly nonlinear, non-dispersive models yield first order PDEs of mixed-type which allow for the study of breaking waves and instabilities arising from large shear. Such phenomena are difficult to study in MCC systems where it is not known whether waves can break, and where stability is difficult to study and instabilities tend to be “filtered” for numerical simulations.

In this paper we focus on the strongly nonlinear, non-dispersive case, and we study the dynamics of two layers of immiscible fluids, bounded below by a horizontal bottom and above by an unbounded layer of fluid, which is dynamically passive but has density, contributing hydrostatically to the pressure. This is sometimes called a ‘two-and-a-half-layer’ configuration in contrast with the two-layer free surface case which we also study and for which the upper fluid has zero density [10]. The PDEs resulting from this model are of mixed-type: hyperbolic or elliptic, depending on the local layer depths and shear. In such a mixed-type model, a crucial mathematical question is whether initially wave-like (hyperbolic) solutions remain inside the hyperbolic domain. It is known that this is always the case for two-layer flows with a rigid lid in the Boussinesq approximation [11], but that it fails in the non-Boussinesq case [12], where more restricted initial data needs to be imposed to ensure hyperbolicity.

For two-layer flows with a free surface, the situation is similar: a wave-like initial condition might also leave the hyperbolic region before it breaks [13], meaning that the Cauchy problem is ill-posed in the long-time (this is not to be confused with local well-posedness, which has been studied in [23]). This fact contrasts to one-layer flows with a free surface, as the latter is known to be always well-posed and connected to the two-layer Boussinesq equations through a map [14]. The stability of these equations has been an active topic of research for decades [15], [16], [17], but most progress has been made towards numerical studies [18], [19] rather than analytical ones [20], [21], [22]. In particular, no explicit criteria that prevents a solution to leave the hyperbolic region is known. In this paper, we shed some light into the problem by presenting a way of visualising the phase plane in two dimensions, and use simple waves [24], [25] to study the stability of the system.

The paper is organised as follows. In Section 2, we first state the problem and derive the governing equations for the general (*i.e.* non-Boussinesq) case resulting in a semilinear system of first-order PDEs. The relevant physical parameters for the problem are introduced and used to write the system in a non-dimensional form. Some limit cases are discussed, in particular the Boussinesq case (in which the density differences contribute only in the buoyancy terms) and the free surface case (in which the density in the upper unbounded and passive layer of fluid is assumed to be zero), being the latter a very common model for internal ocean waves. Section 2 concludes with a discussion of the conservation laws for this problem in both non-Boussinesq and Boussinesq cases. In Section 3, we turn our attention to the dynamics and the nonlinear stability properties of the model. We first present a numerical example of a wave-like solution that enters the elliptic region before the breaking time, showing that the system is nonlinear unstable. It is then shown that in the free surface case, the stability of the system depends only on two physical (“baroclinic”) quantities: the shear and the difference of layer thickness. This result allows a characterisation of the phase space and its hyperbolic region in a plane, and to see the transition

to ellipticity without the need for a four-dimensional space (as the PDE system is 4×4) and in this context, a numerical example of transition is shown. This result is used to provide a precise and explicit criterion for nonlinear stability, which refines the ones provided by [13] and [20]. Finally, we discuss the use of simple waves as a way to look into individual waves of the system, as well as a dynamical system tool to investigate and prevent the transition from the hyperbolic to the elliptic region. In Section 4 we summarise our findings and discuss further extensions to the present work.

2. Two-and-a-half-layer shallow water flows

Consider a two-dimensional, irrotational flow of ideal, incompressible and immiscible fluids in two layers, under the action of gravity, bounded below by a horizontal rigid bottom and above by an unbounded layer of fluid, as shown in Figure 1. The dynamics in the upper layer of fluid is neglected and all its effects in the dynamics of the flow is due to the hydrostatic pressure exerted on the lower layers.

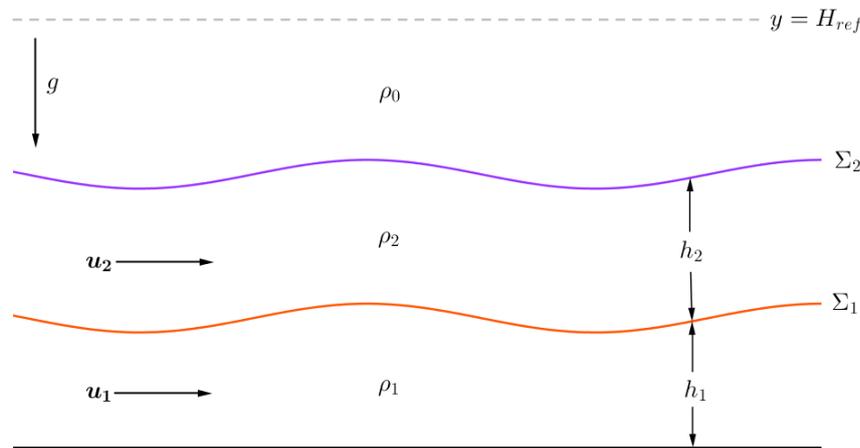


Figure 1: Schematic illustration for the two-and-a-half-layer problem.

The fluid density, pressure and velocity fields in each layer are given by ρ_j , $p_j(x, y, t)$ and $\mathbf{u}_j(x, y, t) = (u_j(x, y, t), v_j(x, y, t))$ respectively, with $j = 1$ representing the lower layer and $j = 2$ representing the upper layer. The passive unbounded fluid above, represented by $j = 0$, has no velocity field and the pressure due to it is assumed to be hydrostatic below a reference level height H_{ref} at which it is constant (of course, the resulting equations are independent of H_{ref}). Usually this represents a layer of much lower density such as the atmosphere above a stratified ocean. The height of each of the active layers is given by $h_j(x, t)$ and the interface between the layers are given by $\Sigma_1 = \{y = h_1\}$ and $\Sigma_2 = \{y = h_1 + h_2\}$, as schematically indicated in Figure 1. In this paper, we only consider the stable case where $\rho_1 \geq \rho_2 \geq \rho_0$. In the real world (e.g. the ocean), ρ_1 usually varies from $1021 - 1070 \text{ kg/m}^3$ (very salty water), ρ_2 ranges from $999 - 1020 \text{ kg/m}^3$ (fresh to slightly salty water) and ρ_0 (the density of air) is considerably smaller, being around 1.2 kg/m^3 at the sea level.

The mathematical model [26], [10] for the dynamics in each layer is given by the incompressible Euler equations

$$\rho_j \frac{D\mathbf{u}_j}{Dt} = -\nabla p_j - \mathbf{F}_j, \quad (2.1a)$$

$$\nabla \cdot \mathbf{u}_j = 0, \quad (2.1b)$$

for $j = 1, 2$, with \mathbf{F}_j being the force field due to external sources. In this model, only gravitational forces play a significant role and so we will consider $\mathbf{F}_j = (0, \rho_j g)$.

The boundary conditions are the impermeability condition at the bottom wall

$$v_1 = 0 \quad \text{on} \quad y = 0, \quad (2.2)$$

the kinematic conditions (KBC) and the dynamic condition (DBC) on Σ_1 respectively given by

$$h_{1,t} + u_1 h_{1,x} = v_1, \quad (2.3a)$$

$$h_{1,t} + u_2 h_{1,x} = v_2, \quad (2.3b)$$

$$p_1 = p_2, \quad (2.3c)$$

while on Σ_2 those conditions are

$$(h_1 + h_2)_t + u_2 (h_1 + h_2)_x = v_2, \quad (2.4a)$$

$$p_2 = p_0. \quad (2.4b)$$

Considering the normal vector on Σ_1 , $\mathbf{n} = (-h_{1,x}, 1)$ and combining (2.3a) and (2.3b), gives

$$\mathbf{n} \cdot \mathbf{u}_1 = \mathbf{n} \cdot \mathbf{u}_2 \quad (2.5)$$

on the interface Σ_1 , which implies the continuity of normal velocity condition between the active fluid. The tangential velocity, however, can be discontinuous at the interface.

This gives us a set of 6 first order equations with 6 boundary conditions for the 6 unknowns u_j, v_j, p_j , $j = 1, 2$. This is a *free boundary problem*, as the domain is also unknown and appears through h_1 and h_2 in the boundary conditions. This configuration is called the ‘two-and-a-half-layer’ model.

(a) Nondimensionalisation and relevant parameters

We introduce the following non-dimensional variables

$$\begin{aligned} \tilde{x} &= x/L, & \tilde{y} &= y/H, & \tilde{t} &= t(\sqrt{gH}/L), & \tilde{\rho}_j &= \rho_j/\bar{\rho}, \\ \tilde{h}_j &= h_j/H, & \tilde{u}_j &= u_j/\sqrt{gH}, & \tilde{v}_j &= v_j/\mu\sqrt{gH}, & \tilde{p}_j &= p_j/\bar{\rho}gH. \end{aligned}$$

Here $\bar{\rho} = (\rho_1 + \rho_2)/2$ is the average density in the layers, H and L are respectively the typical reference height and length and

$$\mu = \frac{H}{L} \quad (2.6)$$

is the *long wave* (or ‘*shallow water*’) *parameter*, which is small in the shallow water limit. Physically, $H = H_1 + H_2$, where H_j is the undisturbed height for the j -th layer of fluid. From now on, we will drop the tildes and work only with the rescaled non-dimensional variables. A direct consequence of the rescaling is that

$$\bar{\rho} \doteq \frac{\rho_1 + \rho_2}{2} = 1. \quad (2.7)$$

(b) Shallow water limit and the governing equations

The *shallow water* (or *long wave*) limit corresponds to the case $\mu \ll 1$, which implies that the horizontal velocity is well represented by its average and that the pressure is hydrostatic. The vertical mean of u_j , for $j = 1, 2$ is defined by

$$\overline{u_j}(x, t) \doteq \frac{1}{h_j(x, t)} \int_{y_j(x, t)}^{y_j(x, t) + h_j(x, t)} u_j(x, y, t) dy,$$

where $y_j(x, t)$ is the coordinate of the lower interface of the j -th layer. The averaged quantities $\overline{v_j}$ and $\overline{p_j}$ can be defined in a similar fashion. Consequences of the shallow water limit are that the

pressure is hydrostatic, and that, if the flow is irrotational in each layer,

$$u_j(x, y, t) = \bar{u}_j(x, t) + \mathcal{O}(\mu^2). \quad (2.8)$$

This means the equations for the mean horizontal velocities can be decoupled from the vertical components [27], resulting in a closed system for \bar{u}_j and h_j . For simplicity we drop the bars and in the shallow water limit $\mu \rightarrow 0$, the governing equations become

$$h_{1,t} + (h_1 u_1)_x = 0, \quad (2.9)$$

$$h_{2,t} + (h_2 u_2)_x = 0, \quad (2.10)$$

$$\rho_1 \left((h_1 u_1)_t + (h_1 u_1^2)_x \right) + (\rho_1 - \rho_0) \left(\frac{h_1^2}{2} \right)_x + (\rho_2 - \rho_0) h_1 h_{2,x} = 0, \quad (2.11)$$

$$\rho_2 \left((h_2 u_2)_t + (h_2 u_2^2)_x \right) + (\rho_2 - \rho_0) \left(\frac{h_2^2}{2} \right)_x + (\rho_2 - \rho_0) h_2 h_{1,x} = 0. \quad (2.12)$$

For more details, the reader is referred to [28], where a similar derivation for the three-layer rigid-lid shallow water equations is carried out.

Note that these constitute 4 equations for the 4 variables u_1, u_2, h_1 and h_2 , which makes the system mathematically closed. The (hydrostatic) pressure do not appear as an additional unknown in these equations, meaning that they are local evolution equations even in the non-Boussinesq case and for any boundary conditions in the horizontal direction. This is a major difference from the layered problems with a rigid lid [12], [28], where the pressure appears in the momentum equations in the non-Boussinesq case as an additional unknown, generally requiring a nonlocal equation to close the problem. This difficulty appears because the pressure at the lid (or another reference level) is unknown whereas in the present case the pressure is known in a reference location.

We note that of the three density ratios appearing in the problem

$$\frac{\rho_1 - \rho_0}{\rho_1}, \quad \frac{\rho_2 - \rho_0}{\rho_1} \quad \text{and} \quad \frac{\rho_2 - \rho_0}{\rho_2},$$

result from only 2 independent parameters (say ρ_0/ρ_1 and ρ_2/ρ_1).

(c) Limiting cases

There are generally several further approximations possible in this problem. One can make the *free surface approximation* where $\rho_0 = 0$. This is applicable, for example when modelling interfacial waves in the ocean, with the atmosphere represented in the model by the upper unbounded layer having much lower density. In this approximation, the number of independent parameters reduces from 2 to 1. Both the dynamics and stability for this case will be discussed in detail in Section 3.

A second possible approximation is the *Boussinesq approximation* [27]. This is usually an approximation imposed *ab initio* by considering that the density differences only affect the buoyancy terms and not the acceleration terms. Under this approximation, Equations (2.11) and (2.12) then become

$$(h_1 u_1)_t + (h_1 u_1^2)_x + (\rho_1 - \rho_0) \left(\frac{h_1^2}{2} \right)_x + (\rho_2 - \rho_0) h_1 h_{2,x} = 0, \quad (2.13)$$

$$(h_2 u_2)_t + (h_2 u_2^2)_x + (\rho_2 - \rho_0) \left(\frac{h_2^2}{2} \right)_x + (\rho_2 - \rho_0) h_2 h_{1,x} = 0, \quad (2.14)$$

were the densities multiplying the acceleration terms were replaced by the average density $\bar{\rho} = 1$. In contrast with the rigid lid case, the Boussinesq approximation in the free surface case does not change the nature of the equations. It only changes the numerical value of the coefficients of the buoyancy terms, and the physical interpretation of the resulting conservation laws. On the

other hand, the Boussinesq approximation in the rigid lid case changes the nature of equations by making them local [12], [28].

From the two-and-a-half-layer Boussinesq equations, one can then derive the two-layer rigid-lid Boussinesq equations as a special limiting case, when the ratio of the speed of the internal mode to the surface mode tends to zero [29].

(d) Conservation laws

A *conservation law* is an expression of the form

$$(q(\mathbf{u}(x, t)))_t + (F(\mathbf{u}(x, t)))_x = 0 \quad (2.15)$$

where q is the quantity conserved and F is the flux of q . Conservation laws reflect physical principles and can be used to provide realistic physics when waves break and discontinuous solutions result [30]. For example, appropriate choices of conservation laws can be used as part of a model in stratified flows to understand the macroscopic consequences of the interfacial wave breaking [29] with or without entrainment between layers. They can also be useful to verify the accuracy of certain numerical methods and devise conservation law based ones [31].

It is well known that the systems discussed here have exactly 6 linearly independent conservation laws [32]: 2 conservations of mass, 2 conservations of circulation, conservation of total momentum and conservation of total energy.

(i) Conservation of mass

The conservation of mass in each layer is stated by equations (2.9), (2.10) obtained previously. Note that, since the flow is incompressible and ρ_j is constant, the conservation of mass can also be seen as *conservation of volume*.

(ii) Conservation of circulation

Expanding the derivatives in the momentum equations (2.11), (2.12) and using the conservation of mass above leads to conservation laws for the layer velocities:

$$u_{1,t} + \left(\frac{u_1^2}{2} + \tilde{\rho}_{11} h_1 + \tilde{\rho}_{21} h_2 \right)_x = 0, \quad (2.16a)$$

$$u_{2,t} + \left(\frac{u_2^2}{2} + \tilde{\rho}_{22} h_1 + \tilde{\rho}_{22} h_2 \right)_x = 0. \quad (2.16b)$$

with the new ‘reduced’ densities given by $\tilde{\rho}_{i,j} = (\rho_i - \rho_0)/\rho_j$, for $j = 1, 2$.

Physically, these are better interpreted as conservation of circulation by taking their linear combinations. Define $w_1 = \rho_1 u_1 - \rho_2 u_2$. Then,

$$w_{1,t} + \left(\frac{1}{2} (\rho_1 u_1^2 - \rho_2 u_2^2) + (\rho_1 - \rho_2) h_1 \right)_x = 0 \quad (2.17)$$

which states the conservation of circulation around Σ_1 . Defining $w_2 = \rho_2 u_2 - \rho_0 u_0$ (recall $u_0 = 0$), we get

$$w_{2,t} + \left(\frac{u_2^2}{2} + (\rho_2 - \rho_0) (h_1 + h_2) \right)_x = 0, \quad (2.18)$$

which states the conservation of circulation around Σ_2 .

(iii) Volume flux and conservation of total momentum

Defining $M = \rho_1 u_1 h_1 + \rho_2 u_2 h_2$ and adding the momentum equations gives

$$M_t + \left(\rho_1 u_1^2 h_1 + \rho_2 u_2^2 h_2 + (\rho_1 - \rho_0) \frac{h_1^2}{2} + (\rho_2 - \rho_0) \frac{h_2^2}{2} + (\rho_2 - \rho_0) h_1 h_2 \right)_x = 0 \quad (2.19)$$

A related quantity to momentum is the total volume flux (or discharge) defined by $Q = u_1 h_1 + u_2 h_2$. This quantity satisfies

$$Q_t + \left(u_1^2 h_1 + u_2^2 h_2 + \tilde{\rho}_{11} \frac{h_1^2}{2} + \tilde{\rho}_{22} \frac{h_2^2}{2} \right)_x + (\tilde{\rho}_{21} h_1 h_{2,x} + \tilde{\rho}_{22} h_2 h_{1,x}) = 0 \quad (2.20)$$

which is not a conservation law since $\tilde{\rho}_{21} \neq \tilde{\rho}_{22}$ in general. However, in the Boussinesq case, from equations (2.13), (2.14) we have

$$Q_t + \left(u_1^2 h_1 + u_2^2 h_2 + (\rho_1 - \rho_0) \frac{h_1^2}{2} + (\rho_2 - \rho_0) \frac{h_2^2}{2} + (\rho_2 - \rho_0) h_1 h_2 \right)_x = 0 \quad (2.21)$$

Thus, the Boussinesq approximation conserves discharge (which is indistinguishable from its ‘‘momentum’’).

These conservation laws should be contrasted with the rigid lid case for 2 (or more) layers. In that case, the discharge can be shown not to depend on x and in the Boussinesq approximation, to be a constant. Further, in the non-Boussinesq case, the evolution equation for Q is useful in calculating the nonlocal equation for the pressure [12] and the momentum is not generally conserved [33].

(iv) Conservation of total energy

The total energy of the system can be defined as $E = K + P$, where K is the *kinetic energy*

$$K = \frac{\rho_1}{2} h_1 u_1^2 + \frac{\rho_2}{2} h_2 u_2^2$$

and P is the *potential energy*

$$\begin{aligned} P &= \int_0^{h_1} \rho_1 z dz + \int_{h_1}^{h_1+h_2} \rho_2 z dz + \int_{h_1+h_2}^{H_{\text{ref}}} \rho_0 z dz \\ &= \frac{(\rho_1 - \rho_0)}{2} h_1^2 + \frac{(\rho_2 - \rho_0)}{2} (h_1 + h_2)^2 + \frac{\rho_0}{2} g H_{\text{ref}}^2. \end{aligned}$$

The conservation law for E is therefore given by

$$E_t + \left(\frac{\rho_1 h_1 u_1^3 + \rho_2 h_2 u_2^3}{2} + (\rho_1 - \rho_2) h_1^2 u_1 + (\rho_2 - \rho_0) (h_1 u_1 + h_2 u_2) (h_1 + h_2) \right)_x = 0. \quad (2.22)$$

As mentioned before, a problem of interest which involves conservation laws arises when studying shock solutions. In this context, one main question is whether the full system of governing equations can be replaced by a system of its conserved laws, which provides uniqueness for such a solution after the shock [29]. We note that, in this case, it is essential for $q(\mathbf{u})$ to be invertible and written as $\mathbf{u}(q)$ in order to have F as a function of q in the right-hand side of Equation (2.15). This question is discussed in more detail in [34], where this fact is used to study the mixing and entrainment in two-layer shallow water flows with and without a rigid wall on the top

3. Dynamics and stability

The two-and-a-half-layer non-Boussinesq model discussed in this paper is a quasi-linear 4×4 system of PDEs which, by equations (2.9) to (2.12), can be written as

$$\mathbf{U}_t + A(\mathbf{U})\mathbf{U}_x = 0, \quad (3.1)$$

where

$$\mathbf{U} = (h_1, h_2, u_1, u_2)^T \quad (3.2)$$

and

$$A(\mathbf{U}) = \begin{pmatrix} u_1 & 0 & h_1 & 0 \\ 0 & u_2 & 0 & h_2 \\ \tilde{\rho}_{11} & \tilde{\rho}_{21} & u_1 & 0 \\ \tilde{\rho}_{22} & \tilde{\rho}_{22} & 0 & u_2 \end{pmatrix}, \quad (3.3)$$

with $\tilde{\rho}_{i,j}$ being the ‘reduced’ densities introduced in Section 2.4.2.

Physical solutions for (3.1) must satisfy $h_j \geq 0$, with no *a priori* restrictions on u_j . We note that $h_j = 0$ are invariant subspaces of codimension 1 in the four-dimensional phase space, which follows from Equations (3.1) to (3.3) by making $h_j = 0$. Moreover, we note that a solution that initially satisfies $h_j \geq 0$ will never cross $h_j = 0$ and leave this region: since $h_j = 0$ has codimension 1, such a solution should cross this subspace tangentially in order to reach $h_j < 0$. However, at each point of $h_j = 0$, the matrix (3.1) has only 3 eigenvectors, lying in the subspace itself and since $\mathbf{U}_t = -A(\mathbf{U})\mathbf{U}_x$, a solution that is tangent to one (or both) of these hyperplanes at a point will remain tangent all throughout its evolution. This is numerically illustrated in Figure 2 for a solution initially satisfying $h_1(x_1, 0) = 0$ and $h_2(x_2, 0) = 0$. An alternative explanation for that will be presented in Subsection (d).

(a) Linear waves and fundamental modes

Consider the general situation showed in Figure 1 and described by equations (2.1a) to (2.5). Disturbing the uniform quiescent state of constant $h_j = H_j$ and zero u_j, v_j as in a standard Kelvin-Helmholtz linear stability analysis, one obtains travelling wave modes proportional to $e^{i(kx - \omega t)}$, with ω given by the following dispersion relation:

$$\begin{aligned} & \left(\rho_1 \coth(|k|H_1)(\rho_0 \sinh(|k|H_2) + \rho_2 \cosh(|k|H_2)) \right. \\ & \left. + \rho_2(\rho_2 \sinh(|k|H_2) + \rho_0 \cosh(|k|H_2)) \right) \omega^4 \\ & + g|k| \left((\rho_2 - \rho_1)(\rho_0 \sinh(|k|H_2) + \rho_2 \cosh(|k|H_2)) \right. \\ & \left. - (\rho_2 - \rho_0)(\rho_1 \sinh(|k|H_2) \coth(|k|H_1) + \rho_2 \cosh(|k|H_2)) \right) \omega^2 \\ & + g^2 |k|^2 (\rho_2 - \rho_0)(\rho_1 - \rho_2) \sinh(|k|H_2) = 0. \end{aligned}$$

By rescaling the variables and taking the long wave limit, where $|k|H_j \ll 1$, one gets the equation for the wave-speeds $\lambda = \omega/|k|$:

$$\lambda^4 - (\tilde{\rho}_{11}R_1 + \tilde{\rho}_{22}R_2)\lambda^2 + \tilde{\rho}_{22}\epsilon R_1R_2 = 0, \quad (3.4)$$

where $R_j = H_j/(H_1 + H_2)$ and $\epsilon = (\rho_1 - \rho_2)/\rho_1$.

Equivalently, one could linearize the system (3.1) around $\mathbf{U}_0 = (R_1, R_2, 0, 0)$. Substituting $\mathbf{U} = \mathbf{U}_0 + \tilde{\mathbf{U}}$ and dropping the nonlinear terms gives

$$\tilde{\mathbf{U}}_t + A(\mathbf{U}_0)\tilde{\mathbf{U}}_x = 0, \quad (3.5)$$

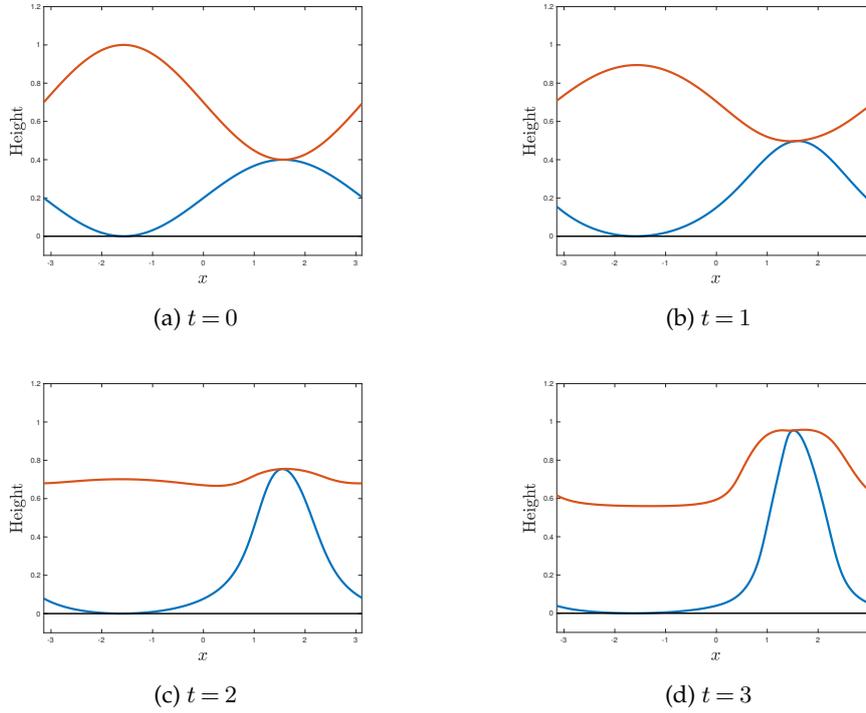


Figure 2: Numerical example of tangency to the hypersurfaces defined by $h_1 = 0$ and $h_2 = 0$. The profile in blue corresponds to h_1 , while $h_1 + h_2$ is plotted in red. This solution has a initial condition satisfying $h_1(x_1, 0) = 0$ and $h_2(x_2, 0) = 0$, for $x_1 = -\pi/2$ and $x_2 = \pi/2$.

where

$$A(\mathbf{U}_0) = \begin{pmatrix} 0 & 0 & R_1 & 0 \\ 0 & 0 & 0 & R_2 \\ \tilde{\rho}_{11} & \tilde{\rho}_{21} & 0 & 0 \\ \tilde{\rho}_{22} & \tilde{\rho}_{22} & 0 & 0 \end{pmatrix}, \quad (3.6)$$

and the characteristic polynomial for (3.6) is given exactly by (3.4). The solutions for (3.4) are

$$\lambda_+^2 = \frac{1}{2} \left((\tilde{\rho}_{11} R_1 + \tilde{\rho}_{22} R_2) + \sqrt{(\tilde{\rho}_{11} R_1 - \tilde{\rho}_{22} R_2)^2 + 4\tilde{\rho}_{21}\tilde{\rho}_{22} R_1 R_2} \right) > \tilde{\rho}_{11} R_1,$$

$$\lambda_-^2 = \frac{1}{2} \left((\tilde{\rho}_{11} R_1 + \tilde{\rho}_{22} R_2) - \sqrt{(\tilde{\rho}_{11} R_1 - \tilde{\rho}_{22} R_2)^2 + 4\tilde{\rho}_{21}\tilde{\rho}_{22} R_1 R_2} \right) < \tilde{\rho}_{11} R_1.$$

This result corresponds to two modes in each direction. The two faster ones, whose interfaces are travelling in phase (see below), are called the *first baroclinic* (or *barotropic*) mode, and the two slower ones, whose interfaces are travelling out of phase (see below), are called the *second baroclinic* (or simply *baroclinic*) mode, respectively depending on the geophysical context [26].

To illustrate this, we can solve the linear system (3.5) for λ_+ to get

$$h_1(x) = R_1 + \tilde{h}(x),$$

$$h_2(x) = R_1 + R_2 + \frac{1}{2R_1\tilde{\rho}_{21}} \left(\lambda_+^2 - \tilde{\rho}_{11} R_1 \right) \tilde{h}(x).$$

Figure 3 (left) shows how this solution looks like when $\tilde{h}(x) = \cos(x)$. We note that as $\lambda_+^2 > \tilde{\rho}_{11} R_1$, this solution travels in phase. This is the case of a barotropic wave. On the other hand, by

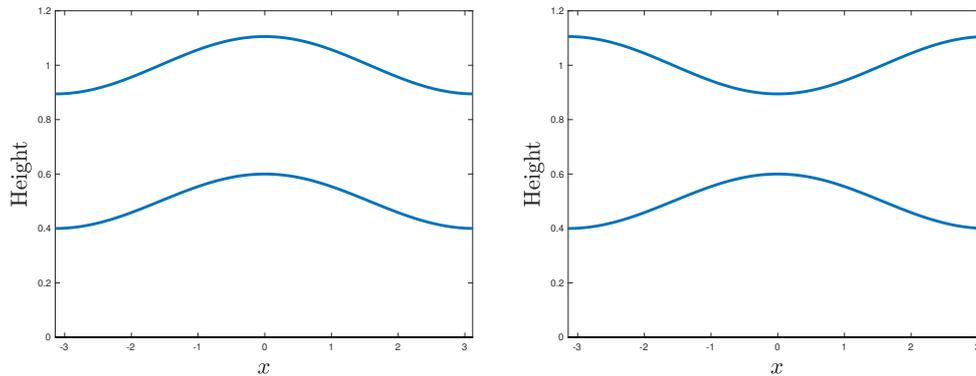


Figure 3: Numerical examples of a barotropic wave (left) and a baroclinic wave (right), with $\tilde{h}(x) = \cos(x)$.

solving (3.5) for λ_- we get

$$h_1(x) = R_1 + \tilde{h}(x)$$

$$h_2(x) = R_1 + R_2 + \frac{1}{2R_1\tilde{\rho}_{21}} \left(\lambda_-^2 - \tilde{\rho}_{11}R_1 \right) \tilde{h}(x).$$

This solution is shown in Figure 3 (right) and is an example of a baroclinic wave. Note that $\lambda_-^2 < \tilde{\rho}_{11}R_1$ and therefore the interfaces are out of phase.

(b) Nonlinear stability

The system (3.1) is *hyperbolic* [35], [36] if all its eigenvalues are real and its eigenvectors span \mathbb{R}^4 . The eigenvalues, denoted by λ , are given by the roots of the characteristic polynomial

$$\begin{aligned}
 P(\lambda) = & \lambda^4 - 2(u_1 + u_2)\lambda^3 + ((u_1 + u_2)^2 + 2u_1u_2 - (\tilde{\rho}_{11}h_1 + \tilde{\rho}_{22}h_2))\lambda^2 \\
 & - 2(u_1u_2(u_1 + u_2) - (\tilde{\rho}_{11}h_1u_2 + \tilde{\rho}_{22}h_2u_1))\lambda \\
 & + (u_1u_2)^2 - (\tilde{\rho}_{11}h_1u_2^2 + \tilde{\rho}_{22}h_2u_1^2) + \epsilon(\tilde{\rho}_{22}h_1h_2).
 \end{aligned} \tag{3.7}$$

Note that in general, the polynomial above cannot be written as a bi-quadratic equation, which implies that the characteristic velocities are different in each direction (left and right) due to the nonzero shear $|u_2 - u_1|$.

An important question related to stability is whether an initially hyperbolic flow remains hyperbolic throughout its evolution. This is called *nonlinear stability* and is a form of long-time *well-posedness* for a Cauchy problem [37].

For the system (3.1), (3.2), (3.3), contrary to the two-layer rigid lid case, and consistent with the remarks of [14], a solution evolving from initially hyperbolic data might cross into the elliptic region of the phase space. This is numerically shown in Figure 4 for $\rho_0 = 0$, $\rho_1 = 1$ and $\rho_2 = 0.9$ and provides a counter example to the conjecture posed by Barros and Choi [20]. A mathematical justification for this fact will be presented later on.

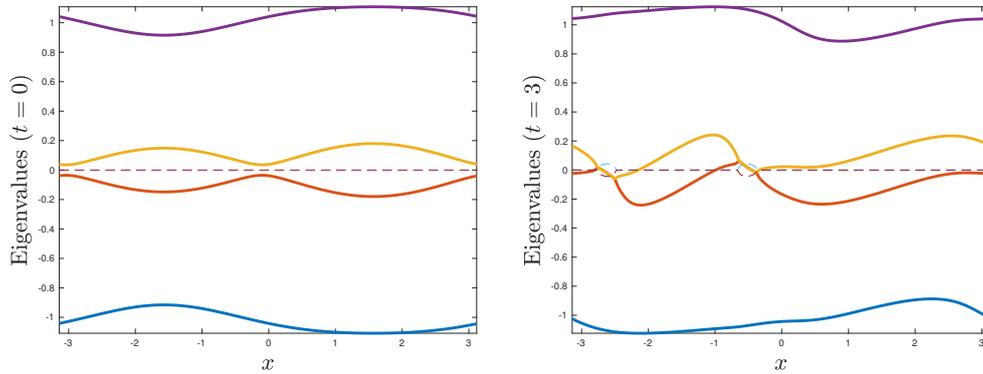


Figure 4: Numerical example of nonlinear instability in the two-and-a-half layer shallow water model. The real part and imaginary parts of the eigenvalues are shown in solid and dashed lines, respectively. On the left are shown the eigenvalues at initial stage for a hyperbolic initial condition. On the right one sees that transition occurs as the baroclinic speeds become complex conjugates, indicating a loss of hyperbolicity and ill-posedness.

(c) The free surface case

The free surface case corresponds to zero density of the unbounded upper layer. The characteristic polynomial (3.7) becomes

$$\begin{aligned}
 P(\lambda) &= \lambda^4 - 2(u_1 + u_2)\lambda^3 + ((u_1 + u_2)^2 + 2u_1u_2 - (h_1 + h_2))\lambda^2 \\
 &\quad - 2(u_1u_2(u_1 + u_2) - (h_1u_2 + h_2u_1))\lambda \\
 &\quad + (u_1u_2)^2 - (h_1u_2^2 + h_2u_1^2) + \epsilon h_1h_2.
 \end{aligned} \tag{3.8}$$

In order to understand the stability in the free surface case, we can locally change the variables in this polynomial so that $u_1 + u_2 = 0$. In fact, redefine u_j as

$$\tilde{u}_j = u_j - \left(\frac{u_1 + u_2}{2}\right)$$

for $j = 1, 2$. Then, if we write $u = \tilde{u}_1$, we have $\tilde{u}_2 = -\tilde{u}_1 = -u$. In these new variables we have

$$P(\lambda) = \lambda^4 - (2u^2 + (h_1 + h_2))\lambda^2 - 2u(h_1 - h_2)\lambda + (u^4 - (h_1 + h_2)u^2 + \epsilon h_1h_2). \tag{3.9}$$

This corresponds to a pointwise reference frame transformation in the original PDEs, whose main consequence was to drop the cubic term in the characteristic polynomial.

Next, for every $\gamma \in \mathbb{R} \setminus \{0\}$, the rescaling

$$\tilde{h}_j = \gamma h_j, \quad \tilde{u} = \gamma^{1/2}u, \quad \tilde{\lambda} = \gamma^{1/2}\lambda$$

do not change the roots. Now, by taking $\gamma = (h_1 + h_2)^{-1}$, we have the identity $\tilde{h}_1 + \tilde{h}_2 = 1$. Introducing the variables $w = \gamma^{1/2}\tilde{u}_1 - \gamma^{1/2}\tilde{u}_2 = 2\tilde{u}$ and $h = \tilde{h}_1 - \tilde{h}_2$. Then, the polynomial (3.9) depends *only* on the two variables h, w , instead of four. This surprising result shows that, in the free surface case, we can state the following theorem:

Theorem 3.1. *The stability of the free surface non-Boussinesq equations system (3.1) to (3.3) with $\rho_0 = 0$ depends only on the two dynamic variables $h = \frac{1}{(h_1 + h_2)}(h_1 - h_2)$ (rescaled layer displacement) and*

$w = \frac{1}{(h_1 + h_2)^{1/2}}(u_1 - u_2)$ (rescaled interfacial shear). The characteristic polynomial is

$$P(\lambda) = \lambda^4 - \left(\frac{w^2}{2} + 1\right)\lambda^2 - (wh)\lambda + \left(\left(\frac{w}{2}\right)^4 - \left(\frac{w}{2}\right)^2 + \frac{\epsilon}{4}(1 - h^2)\right). \quad (3.10)$$

While the stability of the system depends on two quantities, the dynamics is still four-dimensional.

Theorem 3.1 is very helpful as it gives a precise picture of the hyperbolic region, as shown numerically in Figure 5 for different values of ϵ . Although these plots represent only a cross section of the phase space, Theorem 3.1 guarantees that such projections do not change along the other two directions and therefore they show precisely how an initial condition evolves in time. This fact allows, for instance, one to see the dynamics in the phase space for the nonlinear unstable example in Figure 4. This is presented in Figure 6, where it is possible to see that the solution enters the elliptic region.

This result can also be better interpreted when combined with the comments made in the beginning of Section 3, where it was shown that a solution satisfying $h_j \geq 0$ cannot leave this region: we note that this region corresponds to the region $-1 \leq h \leq 1$ in the variables of Theorem 3.1 and therefore a solution in the phase space (H, h, G, w) (with H and G introduced in Section iii) can only enter the elliptic region through the w direction, contrary to what the polynomial (3.10) would initially suggest.

We note that the arguments used to prove Theorem 3.1 do not apply to the case $\rho_0 \neq 0$. In fact, as the upper layer is dynamically passive, its governing dynamical equation is $u_0 = 0$, which is not invariant under translations as equations (2.11) and (2.12).

(i) Collision of baroclinic modes: the onset of instability

Theorem 3.1 also sets the basis for further information on stability. First, introduce $\Omega = w/2$ in Equation (3.10), which results in

$$\begin{aligned} P(\lambda; \Omega) &= \lambda^4 - (2\Omega^2 + 1)\lambda^2 - (2\Omega h)\lambda + \left(\Omega^4 - \Omega^2 + \frac{\epsilon}{4}(1 - h^2)\right) \\ &= (\lambda^2 - \Omega^2)^2 - (\lambda - \Omega)^2 - 2(h + 1)\lambda\Omega + \frac{\epsilon}{4}(1 - h^2). \end{aligned}$$

This resulting equation possesses a remarkable symmetry between λ and Ω . This suggests the use of the variables

$$\zeta = \lambda - \Omega, \quad \eta = \lambda + \Omega. \quad (3.11)$$

This results in

$$Q(\zeta, \eta) = \left(\zeta^2 - \left(\frac{1+h}{2}\right)\right) \left(\eta^2 - \left(\frac{1-h}{2}\right)\right) - (1-\epsilon) \left(\frac{1-h^2}{4}\right).$$

The equation above can be simplified through the rescaling

$$\tilde{\zeta} = \zeta / \sqrt{(1+h)/2}, \quad \tilde{\eta} = \eta / \sqrt{(1-h)/2}, \quad (3.12)$$

which in turn produces

$$\tilde{Q}(\tilde{\zeta}, \tilde{\eta}) = (\tilde{\zeta}^2 - 1) (\tilde{\eta}^2 - 1) - (1-\epsilon). \quad (3.13)$$

Note that the equation above has asymptotes $\tilde{\zeta} \rightarrow \pm 1$ and $\tilde{\eta} \rightarrow \pm 1$, which divides the phase space in two regions. One inside the square $[-1, 1] \times [-1, 1]$ and another outside this domain. The solution curve inside the square obeys $|\tilde{\eta}|, |\tilde{\zeta}| \leq \sqrt{1-\epsilon}$ and represents baroclinic modes only, while the hyperbola the barotropic ones. Note that equations (3.11) and (3.12) can be combined to

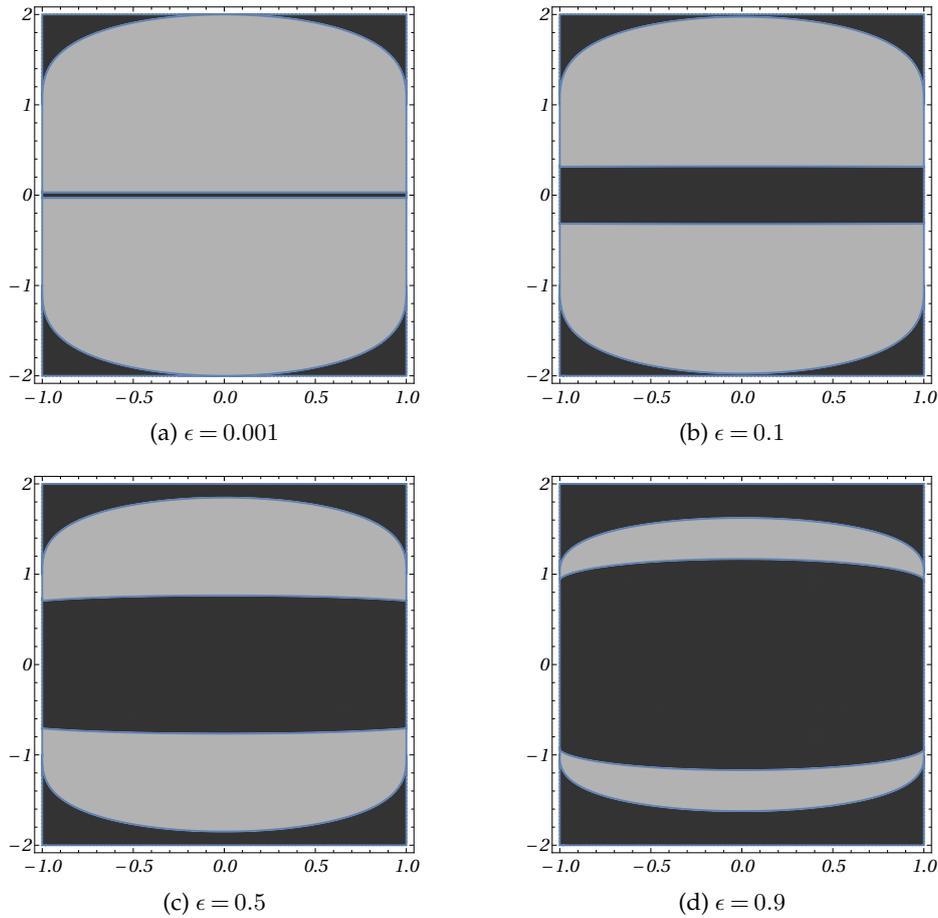


Figure 5: Examples of shear plane (h, w) for different values of ϵ , where the hyperbolic region is in grey and the elliptic region is in black. The physical region is given by $-1 < h < 1$ and $w \in \mathbb{R}$. Note that the solution is hyperbolic for $|w|$ small (weak shear) and become elliptic as the shear increases, indicating the transition to a regime where small scale effects are important and that is not captured by the model. Physically, this indicates the appearance of large-scale Kelvin-Helmholtz instabilities in the interface. The height of the hyperbolic region scales like $\sqrt{\epsilon}$ and when $\epsilon \ll 1$ converges to the rectangle $(h, w) = [-1, 1] \times [-\sqrt{\epsilon}, \sqrt{\epsilon}]$ corresponding to the hyperbolic region in the two-layer rigid-lid Boussinesq equations [11].

give the tangency condition

$$\tilde{\eta} = \sqrt{\frac{1+h}{1-h}} \tilde{\zeta} + \frac{w}{\sqrt{(1-h)/2}}, \quad (3.14)$$

where the angular coefficient $\sqrt{(1+h)/(1-h)}$ can be written in the original variables as $\sqrt{h_2/h_1}$ and $w/(\sqrt{(1-h)/2})$ is nothing more than the Froude number associated to the flow [20]. Equations (3.13) and (3.14) consists in a closed system of two variables for two equations. This system admits 2, 3 or 4 solutions depending on how Equation (3.14) intersects Equation (3.13). This intersection is solely determined by the ‘‘Froude number’’ $w/(\sqrt{(1-h)/2})$, which in turn is directly proportional to the shear w as one could expect. From all these considerations, we conclude that:

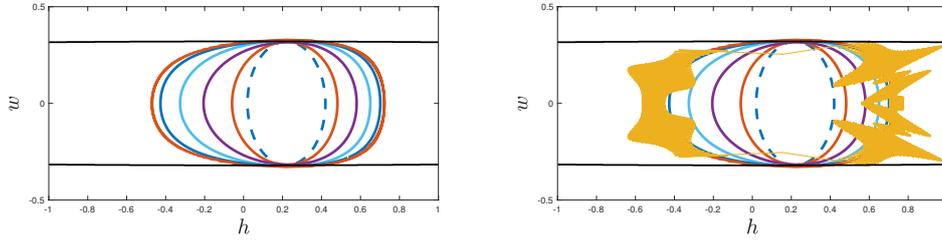


Figure 6: Numerical example of nonlinear instability in the free surface case presented in Figure 4, where $\epsilon = 0.1$. The image on the left shows the solution from an initial stage (blue dashed curve) where it is still inside the hyperbolic domain, up to the point where it crosses the parabolic boundary (red curve crossing the black line). On the right it is possible to see that the solution becomes unstable immediately after it moves into the elliptic region (yellow curve).

Theorem 3.2. *For Froude numbers smaller than 1, the barotropic modes are always stable and do never collide. Baroclinic modes are unstable and their collision is the only route for the system to lose hyperbolicity.*

This result is not new and it was first indicated by Ovsyannikov [13] and reinterpreted later on by Barros and Choi [20]. The latter authors also reinterpreted this result in terms of a criterion for quartic equations by combining equations (3.13) and (3.14) into a single one. An alternative proof can be obtained by rewriting Equation (3.10) (equivalently Equation (3.13)) in polar coordinates, solving it for polar radius and showing that the fast (barotropic) modes do never become complex.

(ii) A criterion for transition

A precise criterion for stability can be obtained from Theorem 3.2. From Equation (3.14), one can isolate the shear w such that

$$\left| \frac{w}{\sqrt{(1-h)/2}} \right| = \left| \tilde{\eta} - \sqrt{\frac{1+h}{1-h}} \tilde{\zeta} \right|. \quad (3.15)$$

The transition occur for values $(\tilde{\eta}, \tilde{\zeta})$ where the line (3.14) is tangent to the curve (3.13). This tangency occur in 4 different points, which by symmetry are $(\tilde{\eta}_{min}, \tilde{\zeta}_{min})$, $(\tilde{\eta}_{max}, \tilde{\zeta}_{max})$ and their antipodes. To find such points, we must solve the equation

$$\frac{d}{d\tilde{\zeta}} \tilde{\eta} = \sqrt{\frac{1+h}{1-h}} \quad (3.16)$$

for $\tilde{\zeta}$ and replace this value in Equation (3.13) to get $\tilde{\eta}$. Note that both $\tilde{\zeta}$ and $\tilde{\eta}$ depend only on h (or equivalently on the angular coefficient in (3.14)) and ϵ .

Applying polar coordinates $\tilde{\eta} = R(\theta) \cos(\theta)$ and $\tilde{\zeta} = R(\theta) \sin(\theta)$ to Equation (3.15), we obtain

$$\left| \frac{w}{\sqrt{(1-h)/2}} \right| = |R(\theta)| \left| \cos(\theta) - \sqrt{\frac{1+h}{1-h}} \sin(\theta) \right| = |R(\theta)| \left| \sqrt{\frac{2}{1-h}} \cos(\theta - \alpha) \right|, \quad (3.17)$$

where $\alpha = \arctan\left(-\sqrt{(1-h)/(1+h)}\right)$. Applying the same polar coordinates in Equation (3.13) gives a bi-quadratic polynomial in $R(\theta)$

$$\left(\frac{\sin^2(2\theta)}{4} \right) R^4(\theta) - R^2(\theta) + \epsilon = 0,$$

which is easy to solve and gives R for the inner contour and for the hyperbolic branches outside the unit box. The resulting solutions are

$$R_{\pm}^2(\theta) = \frac{1 \pm \sqrt{1 - \epsilon \sin^2(2\theta)}}{(\sin^2(2\theta))/2}, \quad (3.18)$$

where clearly $R_-^2 < R_+^2$ for all $\epsilon \in (0, 1)$. In order to find θ for each R , we need to solve Equation (3.16), which in polar coordinates reads

$$\frac{\tilde{\zeta}(\theta)}{(\tilde{\zeta}^2(\theta) - 1)\sqrt{(\tilde{\zeta}^2(\theta) - 1)(\tilde{\zeta}^2(\theta) - \epsilon)}} = \pm \frac{1}{(\epsilon - 1)} \sqrt{\frac{1+h}{1-h}}, \quad (3.19)$$

with the \pm sign depending on whether you are inside or outside the unit box, which we denote by θ_- and θ_+ respectively. The relation above shows that the angle θ depends only on the angular coefficient $\sqrt{(1+h)/(1-h)}$ and the parameter ϵ , as noted geometrically by [13]. In resume, we have the following criterion.

Theorem 3.3. *The system (3.1) to (3.3), with $\rho_0 = 0$, is nonlinearly stable if, and only if*

$$|w| \leq |R_-(\theta_-)| |\cos(\theta_- - \alpha)| \quad \text{or} \quad |w| \geq |R_+(\theta_+)| |\cos(\theta_+ - \alpha)|, \quad (3.20)$$

for all smooth solutions of (3.1) throughout their evolution, where $\alpha = \arctan\left(-\sqrt{(1-h)/(1+h)}\right)$, $R_{\pm}(\theta)$ is given by Equation (3.18) and $\theta_{\pm} = \theta_{\pm}(h, \epsilon)$ is a function of h and ϵ that can be found from Equation (3.19).

This result is not entirely new but it refines the estimates given by both Ovsyannikov [13] and Barros and Choi [20]. In particular, we found expressions for the functions f_1 and f_2 in [13] (equivalently the ‘‘Froude’’ numbers F_{crit}^- and F_{crit}^+ in [20]).

(iii) Governing equations in the new variables

The ideas used to prove Theorem 3.1 suggests a novel way of writing the system (2.9) to (2.12), such that its phase space is precisely the one provided by the theorem. Introduce the ‘‘barotropic’’ variables

$$H = h_1 + h_2 \quad \text{and} \quad G = \frac{u_1 + u_2}{2}.$$

Under this change of variables, the equations (2.9) to (2.12) become, for $\rho_0 = 0$,

$$H_t + \left(\frac{H}{2} (2G + hw\sqrt{H})\right)_x = 0, \quad (3.21a)$$

$$(hH)_t + \left(\frac{H}{2} (2Gh + w\sqrt{H})\right)_x = 0, \quad (3.21b)$$

$$G_t + \left(\frac{4G^2 + w^2H}{8} + H - \epsilon \frac{H(1-h)}{4}\right)_x = 0, \quad (3.21c)$$

$$(w\sqrt{H})_t + \left(Gw\sqrt{H} - \epsilon \frac{H(1-h)}{2}\right)_x = 0, \quad (3.21d)$$

where h and w are the ‘‘baroclinic’’ variables from Theorem 3.1. As before, these equations can be written as a 4×4 system of PDEs of the form (3.1), where $\mathbf{U} = (H, h, G, w)^T$, and $A(\mathbf{U})$ is given by

$$\begin{pmatrix} G + \frac{3}{4}hw\sqrt{H} & \frac{1}{2}wH^{3/2} & H & \frac{1}{2}hH^{3/2} \\ \frac{3}{4}w(1-h^2)H^{-1/2} & G - \frac{1}{2}hw\sqrt{H} & 0 & \frac{1}{2}(1-h^2)\sqrt{H} \\ \frac{1}{8}w^2 + 1 - \frac{1}{4}\epsilon(1-h) & \frac{1}{4}\epsilon H & G & \frac{1}{4}wH \\ -(\epsilon \frac{1}{2}(1-h) + \frac{3}{8}w^2h)H^{-1/2} & \frac{1}{2}(\epsilon - \frac{w^2}{2})\sqrt{H} & \frac{1}{2}w & G - \frac{1}{4}hw\sqrt{H} \end{pmatrix}. \quad (3.22)$$

This system has characteristic polynomial $P(\sqrt{H}(\lambda - G))$, where $P(\lambda)$ is given by Equation (3.10).

The new way of writing the system also shows that, contrary to the Boussinesq case [29], the rigid lid limit in equations (3.21a) to (3.21d) does not provide the decoupling of the baroclinic modes as modelled by the two-layer rigid lid non-Boussinesq system. Physically, the reason is because, while both two-layer rigid-lid and free surface systems are local in the Boussinesq case, the corresponding systems in the non-Boussinesq case are not local, and this pressure term cannot be recasted in the limit.

(d) Simple waves

In a two-and-a-half-layer shallow water flow, the decoupling of characteristic modes is not possible as in the three-layer shallow water flow [28]. However, the study of pure baroclinic and barotropic waves is possible through special solutions called *simple waves*. These are special solutions of the form

$$\mathbf{U}(x, t) = \mathbf{V}(\theta(x, t)). \quad (3.23)$$

Replacing equation (3.23) into (3.1) yields

$$\mathbf{V}_\theta \theta_t + A(\mathbf{V})\mathbf{V}_\theta \theta_x = 0, \quad (3.24)$$

which has a solution if and only if $A(\mathbf{V})\mathbf{V}_\theta$ is proportional to \mathbf{V}_θ , leading to the eigenvalue problem

$$[A(\mathbf{V}(\theta)) - \lambda(\theta)I]\mathbf{V}_\theta(\theta) = 0, \quad (3.25)$$

and where $\theta(x, t)$ must obey the evolution PDE

$$\theta_t + \lambda(\theta)\theta_x = 0, \quad (3.26)$$

which is hyperbolic if $\lambda(\theta)$ is real, for all θ . The eigenvectors \mathbf{V}_θ from equation (3.25) yield, for each eigenvalue family, a vector field in the phase space whose integral curves are the simple waves, meaning that \mathbf{V}_θ is tangent to these curves. For regions in phase space where our system is strictly hyperbolic, this implies the existence of 4 curves through each point. Each of these curves is a simple wave and is invariant under the evolution of the PDE: solutions starting on these curves remain on them, only the parametrisation $\theta(x, t)$ changes with time. Thus the 4 eigenvectors at each point yield a local basis of the phase space providing a decomposition based on the wave speeds $\lambda(\theta)$, or, physically speaking, in terms of the two baroclinic waves and the two barotropic waves. Examples of numerically computed evolution of simple waves in the physical system are shown in Figures 7 and 8.

Simple waves are of crucial importance in the study of nonlinear first-order hyperbolic PDEs. In two-dimensional systems, they define invariant regions [38], [12] due to the property that simple waves do not allow a general solution to cross it transversally [24]. That is the case discussed in the beginning of Section 3: the hyperplanes $h_j = 0$ are not only invariant subspaces, but they are also generated by simple waves, so that any solution that is tangent to one (or both) of these hyperplanes is actually tangent to a three-dimensional surface of simple waves, which prevents the solution to go into the elliptic region.

Furthermore, for mixed-type first order PDE systems, if an initial condition can be bounded by simple waves that do not themselves reach the boundary of the hyperbolic region, then the solution will remain hyperbolic until breaking. Therefore, using simple waves one can build the largest such region, which can be seen as a sharp bound to on hyperbolic initial data that prevents the solution straying into the elliptic region and therefore rendering the problem ill-posed [12].

We use this condition to show that the system (3.1) to (3.3), with $\rho_0 = 0$ (or equivalently, the system (3.21a) to (3.21d)), is indeed nonlinearly unstable. At the point $U = (H, h, G, w = w(h, \epsilon))$, which lies on the parabolic manifold that separates the hyperbolic and elliptic regions, we have three eigenvalues: two corresponding to the barotropic modes and one corresponding to the baroclinic modes that are the same on the surface. The associated eigenvectors, which can be computed from the matrix (3.22), have nonzero w -component so that the vector points outside the hyperbolic region. Hence, the hyperbolic domain cannot be bounded by simple waves.

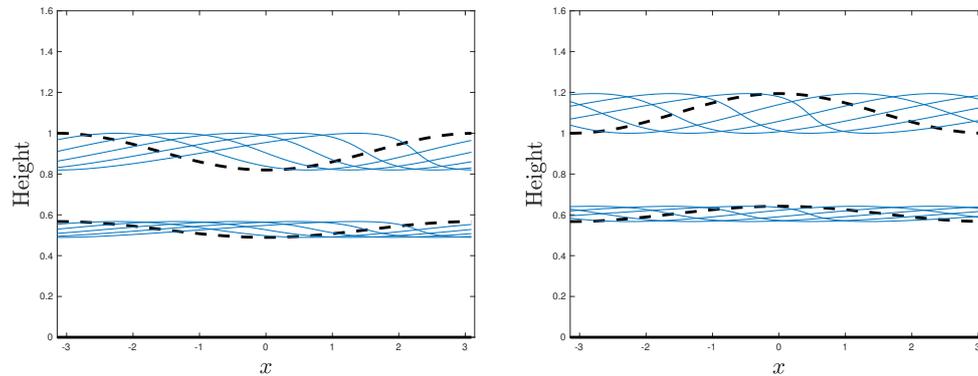


Figure 7: Two numerical examples of pure barotropic waves in the free surface case, computed along a particular simple wave curve, for $\epsilon = 0.1$. The initial wave is in black dashed line and the blue lines show the solution at instant $t_j = j\Delta t$, for $j = 1, \dots, 5$, where $\Delta t = 0.9$. Note the nonlinear steepening of each wave.

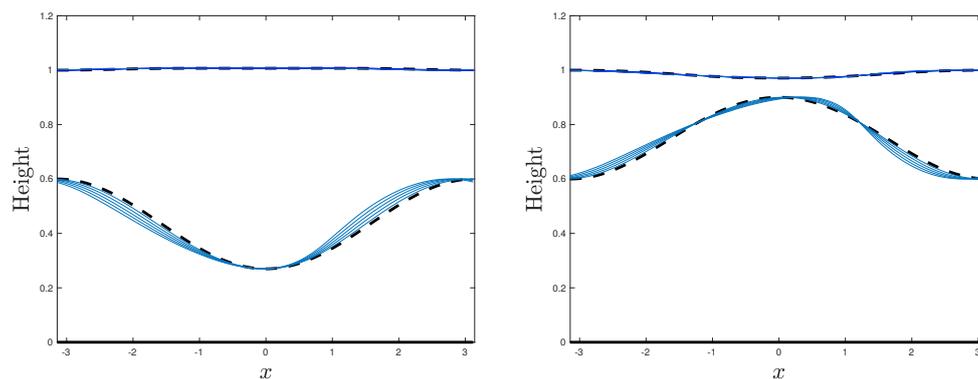


Figure 8: Two numerical examples of pure baroclinic waves in the free surface case, computed along a particular simple wave curve, for $\epsilon = 0.1$. The initial wave is in black dashed line and the blue lines show the solution at instant $t_j = j\Delta t$, for $j = 1, \dots, 5$, where $\Delta t = 0.9$. Note the nonlinear steepening of each wave.

Furthermore, a family of wave-like baroclinic simple waves travelling in this direction will eventually leave the domain. The same analysis can be done to show that a solution in the hyperbolic domains of large shear may also leave this domain.

The examples above show that, in systems of dimensions higher than two, simple waves still provide a construction of “pure” wave solutions, but are less useful for bounding solutions, except in particular cases, for example when there is an invariant subspace of dimension two as in a three-layer shallow water Boussinesq problem [28].

Ideally, we would like to find manifolds generated by families of two (or even three) simple waves. These manifolds would contain families of both simple waves that exist for each mode of motion and would allow one to construct initial data that has waves propagating in both directions in a single mode of the system.

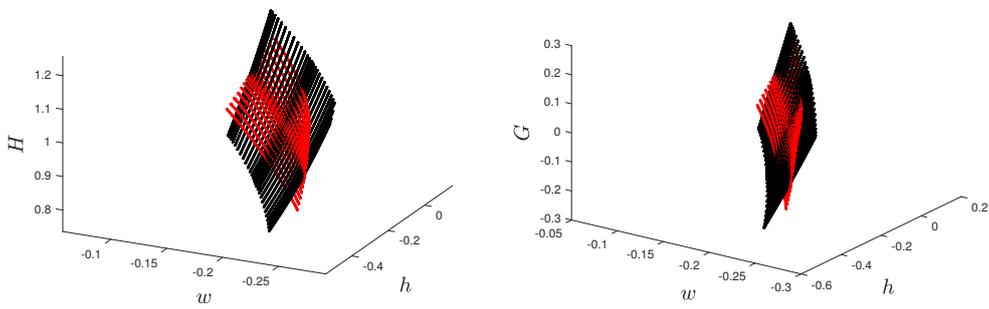


Figure 9: Numerical attempts to construct invariant manifolds for $H = 1, G = 0$.

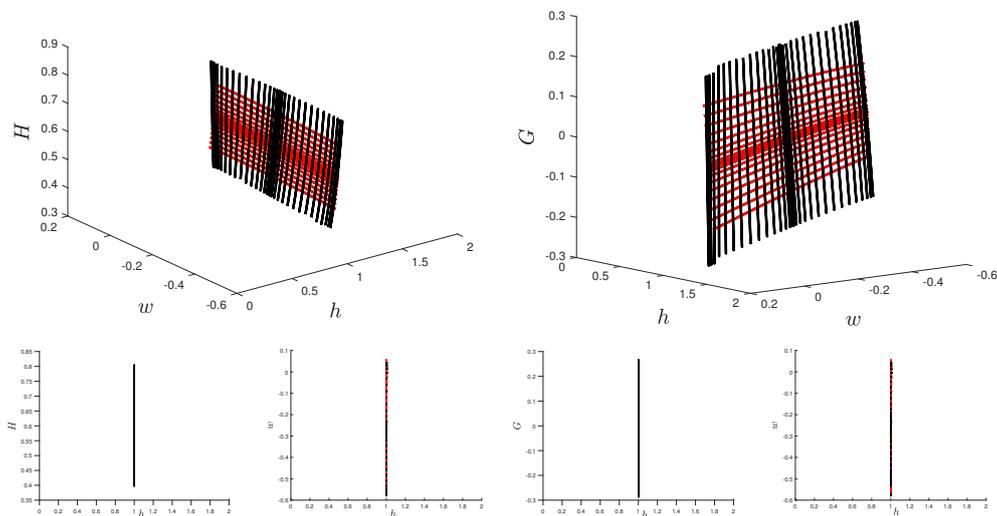


Figure 10: Above: Numerical simulation of the invariant manifolds for $h = 1, w = 0$. Below: Two-dimensional projections of plots above.

However, such manifolds do not exist for general systems of PDEs. The reason is due to the non-existence of an integrating factor for general differential forms in dimensions greater than two. This implies that Riemann invariants, which would allow us to construct such manifolds, do not exist generically, as discussed in [28].

An attempt to construct such a manifold is shown below for $H = 1, G = 0$. This corresponds to the rigid lid (baroclinic) approximation, for which sharp bounds of simple waves have been proven. The resulting images are shown in Figure 9. Note that the contours do not intersect at the nodes and they therefore do not lie on the same surface.

For a comparison effect, we also construct manifolds for the particular case of $h = \pm 1, w = 0$, which is a hypersurface made of simple waves as proven earlier. These plots are shown in Figure 10, where it is possible to see that they in fact form a surface grid.

4. Conclusions

We have derived and studied the long wave dynamics of a two-layer flow with a third unbounded and dynamically passive layer above, in both Boussinesq and non-Boussinesq cases. Some limit cases were discussed, particularly the situation when the density of the upper unbounded layer is zero. In this case, physically motivated changes and rescaling in the dependent variables show that the stability depends only on two dynamical quantities: the shear and the difference in the layer thickness. We use that feature to visualize the phase plane and to show, numerically, that a nonlinear unstable solution indeed leaves the hyperbolic region. The instability is shown to be a consequence of the collision of the baroclinic modes, and a precise criterion to determine whether a solution is stable or not is provided. This result shows the existence of another hyperbolic region for large shears. Preliminary numerical investigations showed that the solutions starting in that region are extremely unstable and move towards the elliptic region at a very fast pace, but this is a matter that deserves further investigation. We use the same change of variables to rewrite the equations in a more appealing way and study the simple waves in this new system. All these results were proven for the free surface case and, although that is the only relevant case in most applications (*e.g.* ocean), it is of mathematical interest to investigate the case in which the density in the upper unbounded layer is weak, but nonzero.

Data Accessibility. The source codes for the numerical simulations are available as supplementary material.

Authors' Contributions. Both authors contributed equally to the presentation of the content. The bulk of calculations and the numerical simulations were performed by F.d.M.V..

Competing Interests. The authors have no competing interests in the context of the current work.

Funding. F.d.M.V. was supported by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brasil), under the grant 249770/2013-0. P.A.M. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo, under the FAPESP-SPRINT University of Bath Scheme.

Acknowledgements. F.d.M.V. also acknowledges the 'QJMAM Fund for Applied Mathematics' for partially funding the presentation of this work at the 12th European Fluid Mechanics Conference, held in Vienna, 2018; and Ricardo Barros for his help with the software Mathematica®

References

1. Helfrich KR, Melville WK. 2005 Long nonlinear internal waves. *Annu. Rev. Fluid Mech.* **38**, 395–425.
2. Stanton TP, Ostrovsky L. 1998 Observations of highly nonlinear internal solitons over the continental shelf. *Geophys. Res. Lett.* **25**, 2695–2698.
3. Ritcher JH, Sassi F, Garcia RR. 2010 Toward a physically based gravity wave source parametrization in a general circulation model. *Journal of the Atmospheric Sciences* **67**, 136–156.
4. Majda A. 2003 *Introduction to PDEs and Waves for the Atmosphere and Ocean*. Society for Industrial and Applied Mathematics 1st edition.
5. Pedlosky J. 1982 *Geophysical Fluid Dynamics*. Springer-Verlag, New York and Berlin 1st edition.
6. Benney DJ. 1966 Long non-linear waves in fluid flows. *Journal of Mathematics and Physics* **45**, 52–63.
7. Grimshaw R, Pelinovsky E, Talipova T. 1997 The modified Korteweg-de Vries equation in the theory of large-amplitude internal waves. *Nonlinear Processes in Geophysics* **4**, 237–250.
8. Miyata M. 1988 Long internal waves of large amplitude. In *Nonlinear water waves* pp. 399–406. Springer.
9. Choi W, Camassa R. 1999 Fully nonlinear internal waves in a two fluid system. *J. Fluid Mech.* **396**, 1–36.
10. Baines PG. 1995 *Topographic Effects in Stratified Flows*. Cambridge University Press, Cambridge 1st edition.
11. Milewski PA, Tabak EG, Turner CV, Rosales RR, Menzague FA. 2004 Nonlinear stability of two-layer flows. *Comm. Math. Sci.* **2**, 427–442.
12. Boonkasame A, Milewski PA. 2011 The stability of large-amplitude shallow interfacial non-Boussinesq flows. *Stud. in Appl. Math.* **128**, 40–58.

13. Ovsyannikov LV. 1979 Two-layer “shallow” water model. *Journal of Applied Mechanics and Technical Physics* **20**, 127–135.
14. Chumakova L, Menzaque FA, Milewski PA, Rosales RR, Tabak EG, Turner CV. 2009 Stability properties and nonlinear mappings of two and three layer stratified flows. *Stud. in Appl. Math.* **122**, 123–137.
15. Schijf JB, Schonfeld JC. 1953 Theoretical considerations on the motion of salt and fresh water. *Proceedings Minnesota International Hydraulic Convention*.
16. Armi L. 1986 The hydraulics of two flowing layers with different densities. *J. Fluid Mech.* **163**, 27–58.
17. Lawrence GA. 1990 On the hydraulics of Boussinesq and non-Boussinesq two-layer flows. *J. Fluid Mech.* **215**, 457–480.
18. Castro M, Macías J, Parés C. 2001 A Q-scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system. *Mathematical Modelling and Numerical Analysis* **35**, 107–127.
19. Castro-Días MJ, Fernández-Nieto ED, González-Vida JM, Parés-Madronal C. 2011 Numerical treatment of the loss of hyperbolicity of the two-layer shallow-water system. *J. Sci. Comput.* **48**, 16–40.
20. Barros R, Choi W. 2008 pp. 95–103. In *On the hyperbolicity of two-layer flows*, pp. 95–103. World Scientific.
21. Gavriluk S, Kazakova M. 2014 Hydraulic jumps in two-layer flows with a free surface. *J. Appl. Mech. Tech. Phy.* **55**, 209–219.
22. Jo TC, Choi YK. 2014 Dynamics of strongly nonlinear internal long waves in a three-layer fluid system. *Ocean Sci. J.* **49**, 357–366.
23. Monjarret R. 2015 Local well-posedness of the two-layer shallow water model with free-surface. *SIAM J. Appl. Math.* **75**, 2311–2332.
24. Chumakova L, Tabak EG. 2010 Simple waves do not avoid eigenvalue crossings. *Comm. on Pure and Applied Math* **63**, 119–132.
25. Mailybaev AA, Marchesin D. 2008 Hyperbolicity singularities in Rarefaction Waves. *J. Dynam. Differ. Eq.* **20**, 1–29.
26. Cushman-Roisin B, Beckers JM. 2011 *Introduction to Geophysical Fluid Dynamics - Physical and Numerical Aspects*. Academic Press, Waltham 2nd edition.
27. de Melo Viríssimo F. 2018 *Dynamical System Methods for Waves in Fluids: Stability, Breaking and Mixing*. PhD thesis Department of Mathematical Sciences, University of Bath, United Kingdom.
28. de Melo Viríssimo F, Milewski PA. 2019 Three-layer flows in the shallow water limit. *Stud. in Appl. Math.* **142**, 487–512.
29. Milewski PA, Tabak EG. 2015 Conservation law modelling of entrainment in layered hydrostatic flows. *J. Fluid Mech.* **772**, 272–294.
30. Smoller J. 1994 *Shock Waves and Reaction-Diffusion Equations*. Springer-Verlag, New York 1st edition.
31. LeVeque RJ. 2002 *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, Cambridge 1st edition.
32. Barros R. 2006 Conservation laws for one-dimensional shallow water models for one and two-layer flows. *Math. Mod. and Meth. in Appl. Sci.* **16**, 119–137.
33. Camassa R, Chen S, Falqui G, Ortenzi G, Pedroni M. 2012 An inertia ‘paradox’ for incompressible stratified Euler fluids. *J. Fluid Mech.* **695**, 330–340.
34. de Melo Viríssimo F, Milewski PA. 2019 Conservation law modelling of mixing and entrainment in stratified multilayer shallow water flows. *In preparation*.
35. John F. 1978 *Partial Differential Equations*. Springer-Verlag, New York 3rd edition.
36. Whitham GB. 1974 *Linear and Nonlinear Waves*. Wiley-Interscience 1st edition.
37. Hadamard JS. 1902 Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin* pp. 49–52.
38. Chumakova L, Menzaque FA, Milewski PA, Rosales RR, Tabak EG. 2009 Shear instability for stratified hydrostatic flows. *Comm. on Pure and Applied Math* **62**, 183–197.