Water Waves
https://doi.org/10.1007/s42286-020-00041-x

ORIGINAL ARTICLE

Normal Mode Decomposition and Dispersive and Nonlinear Mixing in Stratified Fluids

Benoît Desjardins[1] · **David Lannes**[2] · **Jean-Claude Saut**[3]

Received: 18 December 2019 / Accepted: 8 July 2020

Abstract

Motivated by the analysis of the propagation of internal waves in a stratified ocean, we consider in this article the incompressible Euler equations with variable density in a flat strip, and we study the evolution of perturbations of the hydrostatic equilibrium corresponding to a stable vertical stratification of the density. We show the local well-posedness of the equations in this configuration and provide a detailed study of their linear approximation. Performing a modal decomposition according to a Sturm–Liouville problem associated with the background stratification, we show that the linear approximation can be described by a series of dispersive perturbations of linear wave equations. When the so-called Brunt–Vaisälä frequency is not constant, we show that these equations are coupled, hereby exhibiting a phenomenon of dispersive mixing. We then consider more specifically shallow water configurations (when the horizontal scale is much larger than the depth); under the Boussinesq approximation (i.e., neglecting the density variations in the momentum equation), we provide a well-posedness theorem for which we are able to control the existence time in terms of the relevant physical scales. We can then extend the modal decomposition to the nonlinear case and exhibit a nonlinear mixing of different nature than the dispersive mixing mentioned above. Finally, we discuss some perspectives such as the sharp stratification limit that is expected to converge towards two-fluid systems.

Keywords Internal waves · Modal decomposition · Dispersive mixing · Sharp stratification limit

To the memory of our friend Walter Craig.

✉ David Lannes
david.lannes@math.u-bordeaux.fr

1 CMLA ENS Cachan, 61 Avenue du Président Wilson, 94230 Cachan, France

2 Université de Bordeaux, IMB et CNRS UMR 5251, 33405 Talence, France

3 Laboratoire de Mathématiques, UMR 8628, Université Paris-Sud et CNRS, Bat. 307, 91405 Orsay, France

Published online: 21 July 2020

1 Introduction

1.1 General Setting

The aim of this paper is to analyze the propagation of internal waves in a continuously stratified fluid. In oceans, this variation of density (pycnocline) may be due to a difference of salinity (halocline) or temperature (thermocline); note also that such waves also appear in applications to atmospheric studies (e.g., [19,20,31]).

The modeling of such waves has a long history, starting with the pioneering mathematical study of periodic waves by Dubreil-Jacotin [16]. We refer to [4,9,15,39,40,46,47] for theoretical and experimental studies of solitary waves and to [25] for a survey on oceanic internal waves.

While many classical equations such as the Benjamin–Ono equation [5,43] (see also [15]) or the Intermediate Long-Wave equation [32] have been formally derived in this context, no rigorous derivation seems to be available. This is in contrast with the two-layer formulation where the internal waves propagate at the interface of two layers of (incompressible) fluids of different densities. In this setting, one can generalize the classical approach for surface waves (see, e.g., [35]) and derive rigorously (in the sense of consistency), in various regimes, a plethora of asymptotic models including all the classical models of internal waves. Roughly speaking, this is achieved by expanding with respect to suitable small parameters two nonlocal operators that appear when expressing the (free boundary) two-layer system as an equation on a fixed domain. Together with the delicate analysis of the Cauchy problem for the full two-layers system (see [36]), which involves Kelvin–Helmholtz-type instabilities (see also [2,3, 38] for the persistence of these instabilities in shallow water asymptotic models), this leads to the complete justification of some internal wave asymptotic systems.

The situation is quite different for a continuously stratified fluid. There is no more free boundary and one starts from the nonhomogeneous Euler system for which the local well-posedness of the Cauchy problem is well known (see, e.g., [14] and the references therein), although it does not seem to have been considered in the present setting. Another difficulty, addressed here, is to establish a time of existence for the solutions which is relevant with respect to the different physical scales involved.

From a more qualitative viewpoint, it is common in oceanography or atmospheric studies to decompose the various quantities of interest on a well-chosen basis of vertical modes related to the background stratification (e.g., [20,22]). This approach has not been fully justified so far; we provide such a rigorous justification here, exhibiting additional conditions that need to be satisfied if one wants this decomposition to converge properly. Moreover, in most of the studies dealing with continuous stratification, the nonhydrostatic component of the pressure is neglected (this is an important difference with the case a discontinuous stratification, for which many multilayer nonhydrostatic models have been derived, e.g., [6,12]); we show here how to take it into account and that, at the linear level, its contribution is of dispersive nature. We also show that in some cases (when the so-called Brunt–Vaisälä frequency is not constant), this dispersive term induces some mixing between the different modes of the decomposition; this type of mixing is a linear effect. In the case of a constant Brunt–Vaisälä frequency where such mixing does not occur at leading (linear) order, we show that such mixing occurs in shallow water at the next order due to nonlinearities, and we derive a sequence of coupled Boussinesq-like systems.

We now make more precise the physical context of our study.

We assume that the fluid domain is infinite in the horizontal direction $X \in \mathbb{R}^d$ ($d = 1, 2$) delimited by a flat bottom located at $z = -H$ and a rigid lid at $z = 0$. The velocity field at time t and at the point (X, z) of the fluid domain is denoted by $U(t, X, z) \in \mathbb{R}^{d+1}$, and its horizontal and vertical components are, respectively, denoted by $V(t, X, z) \in \mathbb{R}^d$, $w(t, X, z) \in \mathbb{R}$. We also denote by $P(t, X, z) \in \mathbb{R}$ the pressure field and by $\mathbf{g} = -g\mathbf{e}_z$ the (constant) acceleration of gravity. The Euler equations governing the fluid motion are therefore:

$$\begin{cases} \rho(\partial_t U + U \cdot \nabla_{X,z} U) = -\nabla_{X,z} P + \rho \mathbf{g}, \\ \partial_t \rho + U \cdot \nabla_{X,z} \rho = 0, \\ \nabla_{X,z} \cdot U = 0 \end{cases} \qquad (t \geq 0, X \in \mathbb{R}^d, z \in (-H, 0)), \tag{1}$$

with the boundary conditions:

$$w_{|z=-H} = w_{|z=0} = 0, \tag{2}$$

expressing the impermeability of the rigid bottom and lid.

These equations possess equilibrium solutions $(U_{\mathrm{eq}}, \rho_{\mathrm{eq}}, P_{\mathrm{eq}})$ depending only on the vertical variable z of the form:

$$U_{\mathrm{eq}} = 0, \quad \rho_{\mathrm{eq}} = \rho_{\mathrm{eq}}(z), \quad \frac{\mathrm{d}}{\mathrm{d}z} P_{\mathrm{eq}} = -\rho_{\mathrm{eq}} g. \tag{3}$$

Perturbation of such equilibrium solutions gives rise in the oceans to the propagation of waves called “internal waves”. These internal waves are, therefore, exact solutions $(U_{\mathrm{ex}}, \rho_{\mathrm{ex}}, P_{\mathrm{ex}})$ to (1)–(2) of the form:

$$U_{\mathrm{ex}} = \varepsilon U, \quad \rho_{\mathrm{ex}} = \rho_{\mathrm{eq}} + \varepsilon \rho, \quad P_{\mathrm{ex}} = P_{\mathrm{eq}} + \varepsilon P,$$

where $\varepsilon > 0$ is a parameter measuring the amplitude of the perturbation, and with (U, ρ, P) solving:

$$\begin{cases} (\partial_t U + \varepsilon U \cdot \nabla_{X,z} U) = -\dfrac{1}{\rho_{\mathrm{eq}} + \varepsilon \rho} \nabla_{X,z} P - \dfrac{\rho}{\rho_{\mathrm{eq}} + \varepsilon \rho} g \mathbf{e_z}, \\ \partial_t \rho + \varepsilon U \cdot \nabla_{X,z} \rho + w \dfrac{\mathrm{d}}{\mathrm{d}z} \rho_{\mathrm{eq}} = 0, \\ \nabla_{X,z} \cdot U = 0 \end{cases} \tag{4}$$

with the boundary conditions:

$$w_{|z=-H} = w_{|z=0} = 0. \tag{5}$$

Since some readers might only be interested in the physical aspects of our results, we chose to postpone the most involved mathematical proofs to the end of the paper. In Sect. 2, we state the local well-posedness for the Euler system in the configuration (4)–(5) considered here. We then consider in Sect. 3 the linear approximation to the full system of Eqs. (4)–(5). This approximation is justified in Sect. 3.1 and the normal mode decomposition of the solutions to this linear system is performed in Sect. 3.2. When the so-called Brunt–Vaisälä frequency is constant, we show in Sect. 3.3 that the evolution of the coefficients of this decomposition is governed by a sequence of uncoupled dispersive perturbations of wave equations. When the Brunt–Vaisälä frequency is not constant, we exhibit in Sect. 3.4 the mechanism of dispersive mixing. A particular attention is also paid to the derivation of additional conditions ensuring a proper convergence of the modal decomposition.

Finally, in Sect. 4, we consider the case of shallow water configurations, when the horizontal scale of the perturbations is much larger than the depth of the ocean. Under the additional *strong Boussinesq assumption* under which the density is assumed to be constant in Euler's equations, we are able to derive nonlinear models. The first step is to state in Sect. 4.2 a local existence theorem for the nondimensionalized system that ensures that the existence time is relevant with respect to the different physical scales of the problem. We then extend in Sect. 4.3 the modal representation introduced in Sect. 3 to take into account the nonlinear effects. It is shown that they induce another kind of mixing between the modes. Finally, some perspectives are considered in Sect. 5, such as the sharp stratification limit towards two-fluid models. The main mathematical proofs of the paper are provided in Sect. 6 and technical results on commutator estimates are given in Appendix A.

1.2 Notations

- $X = (x, y) \in \mathbb{R}^2$ denotes the horizontal variables. We also denote by z the vertical variable.
- ∇ is the gradient with respect to the horizontal variables; $\nabla_{X,z}$ is the full three-dimensional gradient operator.
- We denote by $d = 1, 2$ the horizontal dimension. When $d = 1$, we often identify functions on $\mathbb{R}$ as functions on $\mathbb{R}^2$ independent of the y variable. In particular, when $d = 1$, the gradient operator takes the form:

$$\nabla_{X,z} f = \begin{pmatrix} \partial_x f \\ 0 \\ \partial_z f \end{pmatrix}.$$

- $\mathcal{S}$ is the flat strip $\mathbb{R}^d \times (-H, 0)$ (or $\mathbb{R}^d \times (-1, 0)$ when working with dimensionless variables in Sect. 4).
- We write U the velocity field; its horizontal component is written V, and its vertical component w.

- We always use simple bars to denote functional norms on $\mathbb{R}^d$ and double bars to denote functional norms on the $d+1$-dimensional domain $\mathcal{S}$; for instance:

$$|f|_p = |f|_{L^p(\mathbb{R}^d)}, \quad |f|_{H^s} = |f|_{H^s(\mathbb{R}^d)}, \quad \|f\|_p = \|f\|_{L^p(\mathcal{S})}, \text{ etc.}$$

- For $f, g \in L^2(\mathcal{S})$, we denote by (f, g) the standard $L^2(\mathcal{S})$ scalar product.
- We use the Fourier multiplier notation:

$$f(D)u = \mathcal{F}^{-1}(\xi \mapsto f(\xi)\widehat{u}(\xi)),$$

  and denote by $\Lambda = (1-\Delta)^{1/2} = (1+|D|^2)^{1/2}$ the fractional derivative operator.
- For all $\nu \in \mathbb{N}$, we denote by $W^{\nu,\infty}$ the space:

$$W^{\nu,\infty} = \{f \in L^\infty(\mathcal{S}), \quad \forall \alpha \in \mathbb{N}^3, \quad \alpha_1+\alpha_2+\alpha_3 \leq \nu, \quad \partial_x^{\alpha_1}\partial_y^{\alpha_2}\partial_z^{\alpha_3} f \in L^\infty(\mathcal{S})\}.$$

- We define, for all $s \in \mathbb{R}$, $k \in \mathbb{N}$ the space $H^{s,k} = H^{s,k}(\mathcal{S})$ by:

$$H^{s,k} = \bigcap_{j=0}^{k} H^j((-H,0); H^{s-j}(\mathbb{R}^d)), \quad \text{with} \quad \|u\|_{H^{s,k}} = \sum_{j=0}^{k} \left\| \Lambda^{s-j}\partial_z^j u \right\|_2 . \tag{6}$$

- If ω is a positive scalar function, we denote by L^2_ω the weighted L^2-space on $\mathcal{S}$ with associated norm:

$$\|f\|^2_{L^2_\omega} = \int_{\mathcal{S}} |f(X,z)|^2 \omega(X,z) \mathrm{d}X\mathrm{d}z.$$

- We generically denote by $C(\cdot)$ some positive function that has a nondecreasing dependence on its arguments.

2 Local Well-Posedness

The Cauchy problem for the nonhomogeneous Euler equations has been considered in various settings: whole space or bounded and unbounded domains, L^2 or L^p-based spaces, etc (see, for instance, [14,27,28]). It seems, however, that the Cauchy problem for the configuration considered here (unbounded domain with a density whose gradient is not in $L^2(\mathcal{S})$) is not included in the existing results. We, therefore, provide below a local well-posedness result. Contrary to the above references, we do not seek in this result to be sharp in the regularity requirements for the initial conditions; our main concern is rather to distinguish as much as possible the vertical and horizontal derivatives in the proof, since they play drastically different roles in the qualitative descriptions of the solutions addressed in this paper. For the sake of clarity, the proof of this theorem is postponed to Sect. 6.1.

Theorem 1 *Let* $0 < \varepsilon \leq 1$*,* $\nu > d+2$*,* $\rho_{\rm eq} \in W^{\nu,\infty}(-H,0)$*, and* $\rho^0, U^0 \in H^\nu(\mathcal{S})$ *be such that* $\nabla_{X,z} \cdot U^0 = 0$ *and:*

$$\exists \rho_{\rm min} > 0, \qquad \inf_{z\in[-H,0]} \rho_{\rm eq}(z) \geq \rho_{\rm min} \quad \textit{and} \quad \inf_{\mathcal{S}} (\rho_{\rm eq} + \varepsilon\rho^0) \geq \rho_{\rm min},$$

Then, there exists $T > 0$*, such that for all* $\varepsilon \in (0,1]$*, there is a unique solution* $(U,\rho) \in C([0,T]; H^\nu(\mathcal{S})^{d+2})$ *to* (4)–(5) *with initial data* (U^0,ρ^0)*; moreover:*

$$\frac{1}{T} = c_1, \qquad \sup_{t\in[0,T]} (\|U\|_{H^\nu} + \|\rho\|_{H^\nu}) \leq c_2,$$

with $c_j = C\left(\frac{1}{\rho_{\rm min}}, \|\rho_{\rm eq}\|_{W^{\nu,\infty}}, H, \|\rho^0\|_{H^\nu}, \|U^0\|_{H^\nu}\right)$*,* $j = 1,2$.

Remark 1 The existence time provided by the theorem is independent of ε. Note, however, that without the linear terms, one would obtain an existence time of order $O(1/\varepsilon)$ instead of $O(1)$. The issue of large time existence (as well as shallow water stability) is addressed in Theorem 2 below.

3 A Linear Approximation

We consider here the linear approximation to the full system of Eqs. (4)–(5), formally obtained by setting $\varepsilon = 0$ in the equations. We show in Sect. 3.1 that this approximation is as expected of precision $O(\varepsilon)$ and then turn to analyze its behavior. As often in oceanography (e.g., [22]), it is convenient to describe the vertical dependence of the functions involved in the problem by decomposing them on a Sturm–Liouville basis. General facts on Sturm–Liouville decompositions are recalled in Sect. 3.2. We then show in Sect. 3.3 that, when the so-called Brunt–Vaisälä frequency N is independent of z, the coefficients of these decompositions are found by solving *uncoupled* dispersive perturbations of wave systems (whose speed is related to the eigenvalues of the Sturm–Liouville problem satisfied by the vertical velocity). We then turn to study in Sect. 3.4 the general case where N is not constant. We show that for such a configuration, the dispersive terms induce a coupling between the different modes of the decomposition. In both cases (constant and nonconstant N), we give sufficient conditions to improve the speed of convergence of the modal decomposition.

3.1 Error Estimate for the Linear Approximation

We consider here the linearized system obtained by taking $\varepsilon = 0$ in (4)–(5). Decomposing the equation on the velocity into its horizontal and vertical components, this yields:

$$
\begin{cases}
\partial_t V + \dfrac{1}{\rho_{\mathrm{eq}}} \nabla P = 0 \\
\partial_t w + \dfrac{1}{\rho_{\mathrm{eq}}} \partial_z P + \dfrac{\rho}{\rho_{\mathrm{eq}}} g = 0, \\
\partial_t \rho + w \dfrac{\mathrm{d}}{\mathrm{d}z} \rho_{\mathrm{eq}} = 0, \\
\nabla_{X,z} \cdot U = 0
\end{cases}
\tag{7}
$$

with the boundary conditions:

$$
w_{|_{z=-H}} = w_{|_{z=0}} = 0. \tag{8}
$$

As shown in the following proposition, the solutions of this linear model provide an $O(\varepsilon)$ approximation to the exact solution of the full nonlinear problem (4)–(5). The proof of this result is postponed to Sect. 6.2.

Proposition 1 *Let the assumptions of Theorem 1 be satisfied and denote by $T > 0$ the existence time of the solution (U, ρ) of the nonlinear problem (4)–(5) provided by this theorem. There exists also a unique solution $(U^{\mathrm{lin}}, \rho^{\mathrm{lin}}) \in C([0, T]; H^\nu(\mathcal{S})^{d+2})$ to the linear problem (7)–(8) with the same initial data, and the following error estimate holds:*

$$
\|(\rho - \rho^{\mathrm{lin}}, U - U^{\mathrm{lin}})\|_{L^\infty([0,T]\times H^{\nu-1})} \leq c_3 \varepsilon
$$

with $c_3 = C\left(T, \frac{1}{\rho_{\min}}, \|\rho_{\mathrm{eq}}\|_{W^{\nu,\infty}}, H, \|\rho^0\|_{H^\nu}, \|U^0\|_{H^\nu}\right)$.

3.2 Modal Decomposition

Simple manipulations of the linear equations (7)–(8) show that the vertical velocity w must satisfy:

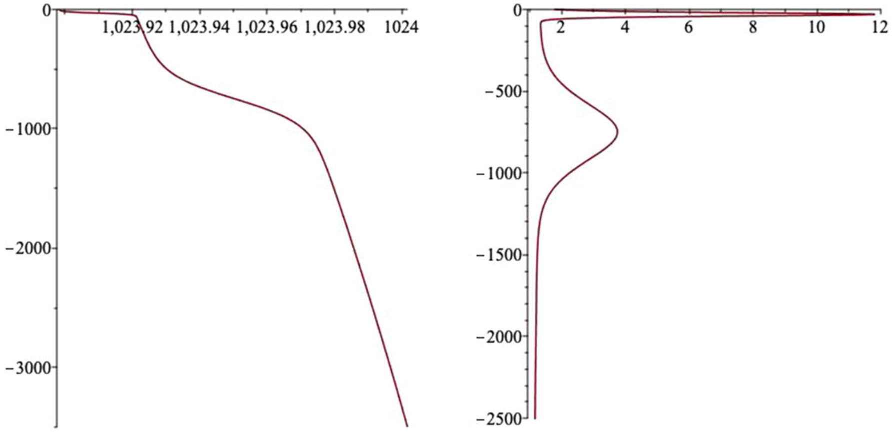
$$
\partial_t^2 \left[\left(\Delta + \frac{1}{\rho_{\mathrm{eq}}} \partial_z \left(\rho_{\mathrm{eq}} \partial_z \cdot \right) \right) w \right] + N^2 \Delta w = 0, \tag{9}
$$

where $N = N(z)$ is the Brünt–Vaisälä frequency:

$$
N^2 = -\frac{\rho_{\mathrm{eq}}'}{\rho_{\mathrm{eq}}} g.
$$

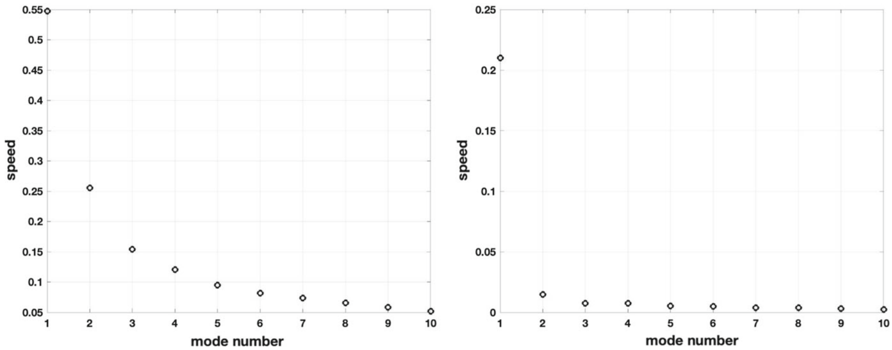
To solve (9), it is, therefore, quite natural to decompose w on the orthonormal basis $(\mathbf{f}_n)_{n\in\mathbb{N}}$ of $L^2([-H, 0], \rho_{\mathrm{eq}} N^2 dz)$ formed by the eigenfunctions of the following Sturm–Liouville problem:

$$
\begin{cases}
\dfrac{\mathrm{d}}{\mathrm{d}z} \left(\rho_{\mathrm{eq}} \dfrac{\mathrm{d}}{\mathrm{d}z} \mathbf{f}_n \right) + \dfrac{\rho_{\mathrm{eq}} N^2}{c_n^2} \mathbf{f}_n = 0, \\
\mathbf{f}_n(-H) = \mathbf{f}_n(0) = 0,
\end{cases}
\tag{10}
$$



(a) Density (kg.m^{-3}) **(b)** N (cycles/hour)

Fig. 1 Example of the vertical dependence of the density and Brunt–Vaisälä frequency in an ocean



(a) For the ocean stratification of Fig. 1a **(b)** For the sharp stratification (25)

Fig. 2 The speeds associated with the first modes

with corresponding eigenvalues $(\frac{1}{c_n^2})_{n\in\mathbb{N}}$ satisfying $c_1 > c_2 > \cdots > c_n > \ldots$ and $c_j > 0$ for all $j \in \mathbb{N}^*$; we also have the asymptotic behavior $c_n = O(1/n)$ as $n \to \infty$.

Example 1 We show in Fig. 1 a typical stratification for an ocean of depth $H = 3500$m (see [11]) and the corresponding Brunt–Vaisälä frequency. The speeds associated with the first modes in the Sturm–Liouville problem (10) are represented in Fig. 2a.

Finally, the first six eigenfunctions of the Sturm–Liouville basis $(\mathbf{f}_n)$ associated with (10) are represented in Fig. 3. For the numerical simulations, we performed a classical discretization on Chebyshev nodes of the differential operator of the Sturm–Liouville operator, and computed the eigenvalues and eigenvectors of the resulting matrix. We refined until mesh convergence of the first eigenvalues and eigenvectors. The eigenfunctions of the Sturm–Liouville problem were then approximated by representing the eigenvectors on the Chebyshev nodes.

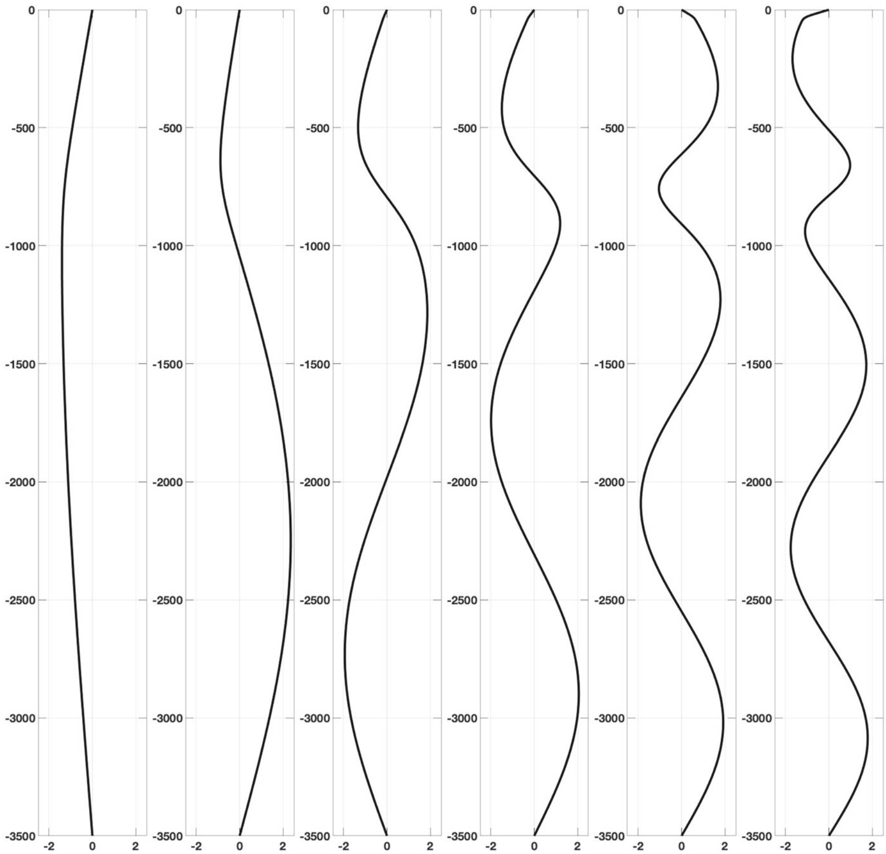


Fig. 3 The first six vertical modes of the Sturm–Liouville decomposition (10) (from left to right)

We, therefore, look for w under the form:

$$w(t, X, z) = \sum_{n=1}^{\infty} w_n(t, X)\mathbf{f}_n(z). \tag{11}$$

Since the coefficients of the system (7) depend on z, the velocity, pressure, and density fields must be decomposed on other orthonormal basis that are introduced in the following lemma.

Lemma 1 *Let* $\rho_{\text{eq}} \in W^{1,\infty}(-H, 0)$ *and* $\alpha = \left(\int_{-H}^{0} \rho_{\text{eq}}^{-1}\right)^{-1/2}$*, and assume that:*

$$\inf_{z\in(-H,0)} \rho_{\text{eq}}(z) > 0 \quad \textit{and} \quad \inf_{z\in(-H,0)} N(z) > 0,$$

and let $(\boldsymbol{f}_n)_{n\geq 1}$ *be the orthonormal basis of* $L^2((-H, 0), \rho_{\text{eq}} N^2 dz)$ *formed by the eigenfunctions of* (10)*. Then:*

(i) *The sequence* $(\rho_{\mathrm{eq}} N^2 \boldsymbol{f}_n)_{n\geq 1}$ *forms an orthonormal basis of the weighted space* $L^2((-H,0), (\rho_{\mathrm{eq}} N^2)^{-1} dz)$.
(ii) *The sequence* $(\boldsymbol{g}_n)_{n\geq 0}$ *with* $\boldsymbol{g}_0 = \alpha \rho_{\mathrm{eq}}^{-1}$ *(with* α *as above), and* $\boldsymbol{g}_n = c_n \boldsymbol{f}'_n$ *for* $n \geq 1$ *forms an orthonormal basis of* $L^2((-H,0), \rho_{\mathrm{eq}} dz)$.
(iii) *The sequence* $(\rho_{\mathrm{eq}} \boldsymbol{g}_n)_{n\geq 0}$ *forms an orthonormal basis of* $L^2((-H,0), \rho_{\mathrm{eq}}^{-1} dz)$.

Proof The first point is straightforward.

For the second point, remark first that integrating by parts and using (10), one gets:

$$\begin{aligned}
\forall m, n \geq 1, \qquad (\mathbf{g}_m, \mathbf{g}_n)_{L^2_{\rho_{\mathrm{eq}}}} &= c_m c_n (\rho_{\mathrm{eq}} \mathbf{f}'_m, \mathbf{f}'_n) \\
&= -c_m c_n ((\rho_{\mathrm{eq}} \mathbf{f}'_m)', \mathbf{f}_n) \\
&= \frac{c_n}{c_m} (\mathbf{f}_m, \mathbf{f}_n)_{L^2_{\rho_{\mathrm{eq}} N^2}} \\
&= \delta_{mn},
\end{aligned}$$

which proves that $(\mathbf{g}_n)_{n\geq 1}$ is an orthonormal family for $L^2((-H,0), \rho_{\mathrm{eq}} dz)$. We, therefore, need to prove that it forms an orthonormal basis if we complement it with $\mathbf{g}_0$. The scalar α is chosen, so that $\mathbf{g}_0$ is a unit vector. To check that it is orthogonal to $\mathbf{g}_n$ for $n \geq 1$, just remark that:

$$\begin{aligned}
(\mathbf{g}_0, \mathbf{g}_n)_{L^2_{\rho_{\mathrm{eq}}}} &= \alpha c_n \int_{-H}^{0} \mathbf{f}'_n \\
&= 0,
\end{aligned}$$

since $\mathbf{f}_n$ vanishes at the boundaries. The family $(\mathbf{g}_n)_{n\geq 0}$ is, therefore, orthonormal and we just have to prove that its orthogonal reduces to $\{0\}$. Let, therefore, $f \in L^2((-H,0), \rho_{\mathrm{eq}} dz)$ be such that:

$$\forall n \in \mathbb{N}, \qquad (f, \mathbf{g}_n)_{L^2_{\rho_{\mathrm{eq}}}} = 0,$$

and let us prove that $f = 0$. From the case $n = 0$, one deduces that $\int_{-H}^{0} f = 0$. Denoting $F(z) = \int_{-H}^{z} f$, one also deduces that, for all $n \geq 1$, one has:

$$\begin{aligned}
0 &= c_n (f, \mathbf{f}'_n)_{L^2_{\rho_{\mathrm{eq}}}} \\
&= -c_n (F, (\rho_{\mathrm{eq}} \mathbf{f}'_n)') \\
&= c_n^{-1} (F, \mathbf{f}_n)_{L^2_{\rho_{\mathrm{eq}} N^2}},
\end{aligned}$$

where we used the fact that F vanishes at the boundaries (owing to the fact that f has zero mean) to derive the second equality. Since $(\mathbf{f}_n)_{n\geq 1}$ is an orthonormal basis of $L^2((-H,0), \rho_{\mathrm{eq}} N^2 dz)$, we deduce that $F = 0$ and, therefore, that $f = 0$.

The third point is then directly deduced from the second one. □

Owing to the lemma, we can look for V, P and ρ under the following expansions:

$$
\begin{aligned}
V(t,X,z) &= \sum_{n=0}^{\infty} V_n(t,X)\mathbf{g}_n(z),\\
P(t,X,z) &= \sum_{n=0}^{\infty} P_n(t,X)\rho_{\mathrm{eq}}(z)\mathbf{g}_n(z),\\
\rho(t,X,z) &= \sum_{n=1}^{\infty} \frac{1}{g}\rho_n(t,X)\rho_{\mathrm{eq}}(z)N^2(z)\mathbf{f}_n(z).
\end{aligned}
\tag{12}
$$

Remark 2 Note, in particular, that:

$$
\|V\|^2_{L^2_{\rho_{\mathrm{eq}}}} = \sum_{n=0}^{\infty} |V_n|_2^2, \quad \|w\|^2_{L^2_{\rho_{\mathrm{eq}}N^2}} = \sum_{n=0}^{\infty} |w_n|_2^2, \quad \|g\rho\|^2_{L^2_{\frac{1}{\rho_{\mathrm{eq}}N^2}}} = \sum_{n=0}^{\infty} |\rho_n|_2^2;
$$

more generally, one has for all $s \in \mathbb{R}$:

$$
\|V\|^2_{H^{s,0}_{\rho_{\mathrm{eq}}}} = \sum_{n=0}^{\infty} |V_n|^2_{H^s}, \quad \|w\|^2_{H^{s,0}_{\rho_{\mathrm{eq}}N^2}} = \sum_{n=0}^{\infty} |w_n|^2_{H^s}, \quad \|g\rho\|^2_{H^{s,0}_{\frac{1}{\rho_{\mathrm{eq}}N^2}}} = \sum_{n=0}^{\infty} |\rho_n|^2_{H^s}.
$$

3.3 The Case of Constant *N*: No Dispersive Mixing

The following proposition restates the system of equations (7) as a set of equations on the coefficients of these decompositions. We first assume that the Brünt–Vaisälä frequency N is constant.

Proposition 2 *Let $\rho_{\mathrm{eq}} \in W^{1,\infty}(-H,0)$ be such that $N^2 = -\frac{\rho'_{\mathrm{eq}}}{\rho_{\mathrm{eq}}}$ is constant and strictly positive on $(-H,0)$. Let also $s \geq 0$ and $T > 0$ and $(U^{\mathrm{lin}}, \rho^{\mathrm{lin}}) \in C([0,T]; H^{s,0})$ solve the linear equations (7)–(8), and assume that the vertical component of the vorticity is zero; that is, $\nabla^\perp \cdot V^{\mathrm{lin}}_{|t=0} = 0$.*

The coefficients (V_n, w_n, ρ_n, P_n) provided by the decompositions (11) and (12) are found by solving, and for all $n \geq 1$, the equations:

$$
\begin{cases}
\left(1 - \dfrac{c_n^2}{N^2}\Delta\right)\partial_t V_n + c_n \nabla \rho_n = 0,\\
\partial_t \rho_n + c_n \nabla \cdot V_n = 0,
\end{cases}
$$

with the initial data given by the coefficients of the decomposition of and V^0 and ρ^0, and w_n and P_n are given by:

$$
w_n = -c_n \nabla \cdot V_n, \quad \text{and} \quad P_n = c_n \rho_n + \frac{c_n}{N^2}\partial_t w_n,
$$

while for $n = 0$, one must have $P_0 = 0$ and $V_0 = 0$. Moreover:

$$\sum_{n=1}^{\infty}\Big[|\rho_n|_{H^s}^2 + |V_n|_{H^s}^2 + |w_n|_{H^s}^2\Big] < \infty.$$

Remark 3 The assumption that $\nabla^\perp \cdot V^{\mathrm{lin}}_{|t=0} = 0$ is not restrictive. Indeed, one deduces from the first equation of (7) that $\nabla^\perp \cdot V^{\mathrm{lin}}$ is constant in time. We can, therefore, always decompose V^{in} as $V^{\mathrm{in}}(t, X, z) = V_{\mathrm{I}}(X, z) + V_{\mathrm{II}}(t, X, z)$ with $\nabla \cdot V_{\mathrm{I}} = \nabla^\perp \cdot V_{\mathrm{II}} = 0$ and apply the proposition to $V_{\mathrm{II}}(t, X, z)$.

Remark 4 The system solved by the (V_n, ρ_n) is a linear Boussinesq-type system similar to those arising in the study of shallow water waves (e.g., [35]). It is striking that for water waves, the dispersive perturbation appears as an approximation of a nonlocal operator (related to a Dirichlet–Neumann operator) which is valid only in shallow water. Here, the simple, differential, form of the dispersive term is valid outside the shallow water regime (see Sect. 4 for an investigation of this latter).

Proof The proof consists in decomposing the equations of (7) on several basis related to the Sturm–Liouville basis $(\mathbf{f}_n)_{1\leq n}$. We have, therefore, to decompose time and space derivatives of V, w, ρ, and P on such basis. This is straightforward for time and horizontal derivatives; for instance:

$$\partial_t V = \sum_{n=0}^{\infty}(\partial_t V_n)\mathbf{g}_n, \qquad \nabla P = \sum_{n=0}^{\infty}(\nabla P_n)\rho_{\mathrm{eq}}\mathbf{g}_n,$$

but this is less obvious for vertical derivatives, which are considered in the following lemma.

Lemma 2 *Let $s \in \mathbb{R}$.*

(i) *The following two assertions are equivalent:*

1. *One has $w \in H^{s,1}(\mathcal{S})$ and $w_{|z=-H,0} = 0$.*
2. *One has $w = \sum_{n=1}^{\infty} w_n \boldsymbol{f}_n$ and $\sum_{n=1}^{\infty}|w_n|_{H^s}^2 + \sum_{n=1}^{\infty} c_n^{-2}|w_n|_{H^{s-1}}^2 < \infty$.*

Moreover, one has:

$$\partial_z w = \sum_{n=1}^{\infty}\frac{1}{c_n}w_n\boldsymbol{g}_n.$$

(ii) *The following two assertions are equivalent:*

1. *One has $\nabla P \in H^{s,0}$ and $\partial_z P \in H^{s+1,0}$.*
2. *One has $P = \sum_{n=0}^{\infty} P_n \boldsymbol{\rho}_{\mathrm{eq}}\boldsymbol{g}_n$ and $\sum_{n=1}^{\infty} c_n^{-2}|P_n|_{H^{s+1}}^2 < \infty$.*

Moreover, one has:

$$\partial_z P = -\sum_{n=1}^{\infty}\frac{1}{c_n}P_n\rho_{\mathrm{eq}}N^2\boldsymbol{f}_n.$$

Proof of the lemma (i) Let us first prove the direct implication. By assumption, $\partial_z w \in L^2_z H^{s-1}(\mathbb{R}^d)$ and can, therefore, be decomposed on the basis $(\mathbf{g}_n)_{0\leq n}$, which, we recall, is orthonormal for the $L^2_{\rho_{\mathrm{eq}}}$ scalar product.

$$\partial_z w = \sum_{n=0}^{\infty} (\partial_z w, \mathbf{g}_n)_{L^2_{\rho_{\mathrm{eq}}}} \mathbf{g}_n.$$

Since w vanishes at the boundaries, we have:

$$\begin{aligned}(\partial_z w, \mathbf{g}_n)_{L^2_{\rho_{\mathrm{eq}}}} &= -\int_{-H}^{0} w c_n \frac{\mathrm{d}}{\mathrm{d}z}\left(\rho_{\mathrm{eq}} \frac{\mathrm{d}}{\mathrm{d}z}\mathbf{f}_n\right) \\ &= \frac{1}{c_n} w_n,\end{aligned}$$

using the definition of $\mathbf{f}_n$ and $\mathbf{g}_n$ (when $n = 0$, the last expression should be replaced by 0). The result follows, therefore, from Parseval identity.

For the reverse inequality, Parseval identity directly yields $w \in H^{s,1}(\mathcal{S})$, but it remains to show that the trace vanishes at the boundary. This is the case, because w is an infinite sum of functions that vanish at the boundary and that, under our assumptions, and recalling that the eigenfunctions are uniformly bounded [21], this sum is absolutely convergent.

(ii) Since $\partial_z P \in H^{s+1,0}$ and $\nabla P \in H^{s,0}$, there are functions $q_n \in H^{s+1}(\mathbb{R}^d)$ and $Q_n \in H^s(\mathbb{R}^d)^d$, such that:

$$\partial_z P = \sum_{n=1}^{\infty} q_n \rho_{\mathrm{eq}} N^2 \mathbf{f}_n \quad \text{and} \quad \nabla P = \sum_{n=0}^{\infty} Q_n \rho_{\mathrm{eq}} \mathbf{g}_n.$$

Since $\nabla\partial_z P \in H^{s,0}$, we can decompose it on the basis $(\rho_{\mathrm{eq}} N^2 \mathbf{f}_n)$ which is orthonormal for the $L^2((-H, 0), (\rho_{\mathrm{eq}} N^2)^{-1}\mathrm{d}z)$ scalar product, and we have:

$$(\partial_z \nabla P, \rho_{\mathrm{eq}} N^2 \mathbf{f}_n)_{L^2_{(\rho_{\mathrm{eq}} N^2)^{-1}}} = -\left(\nabla P, \frac{1}{c_n}\mathbf{g}_n\right)_{L^2} = -\frac{1}{c_n} Q_n$$

(note that contrary to the first point, it is not required that P vanishes at the boundary, because the eigenfunctions $\mathbf{f}_n$ satisfy such a cancelation). This yields that $Q_n = -c_n \nabla q_n$ and the result follows by setting $P_n = -c_n q_n$.

Using Lemmas 1 and 2, the system of equations (7) is equivalent to the relations obtained by taking the $L^2((-H, 0), dz)$ scalar product of these equations with $\rho_{\mathrm{eq}}\mathbf{g}_n$ ($n \geq 0$), $\rho_{\mathrm{eq}}\mathbf{f}_n$ ($n \geq 1$), $\mathbf{f}_n$ ($n \geq 1$) and $\rho_{\mathrm{eq}}\mathbf{g}_n$ ($n \geq 0$), respectively. These relations are:

$$\begin{cases} \partial_t V_0 + \nabla P_0 = 0, \\ \nabla \cdot V_0 = 0 \end{cases} \tag{13}$$

and

$$\forall n \geq 1, \qquad \begin{cases} \partial_t V_n + \nabla P_n = 0, \\ \dfrac{1}{N^2} \partial_t w_n + \rho_n - \dfrac{P_n}{c_n} = 0, \\ \partial_t \rho_n - w_n = 0, \\ \dfrac{1}{c_n} w_n = -\nabla \cdot V_n, \end{cases} \tag{14}$$

where the assumption that N is independent of z has been used to derive the second equation in (14). The equations on the coefficients given in the proposition follow easily. Finally, the convergence of the summation given in the last point of the proposition is a consequence of Remark 2. □

Remark 5 The second equation in (14) is actually:

$$\frac{P_n}{c_n} - \rho_n = (\partial_t w, \mathbf{f}_n)_{L^2_{\rho_{\mathrm{eq}}}},$$

and mixes all the modes if N^2 is not constant. Since we assumed that N^2 is constant in the statement of the proposition, we have $(\partial_t w, \mathbf{f}_n)_{L^2_{\rho_{\mathrm{eq}}}} = \frac{1}{N^2} \partial_t w_n$. The general case is considered in Proposition 4 below.

Remark 6 One directly deduces from (14) that:

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(|V_n|_2^2 + \frac{c_n^2}{N^2} |\nabla \cdot V_n|_2^2 + |\rho_n|_2^2 \right) = \frac{\mathrm{d}}{\mathrm{d}t} \left(|V_n|_2^2 + \frac{1}{N^2} |w_n|_2^2 + |\rho_n|_2^2 \right) = 0.$$

Owing to Remark 2, we get after summing over $n \in \mathbb{N}$:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d} \left(\frac{1}{2} \rho_{\mathrm{eq}} |U|^2 + \frac{1}{2} \rho^2 \frac{g^2}{N^2 \rho_{\mathrm{eq}}} \right) = 0,$$

which is the energy identity satisfied by the solution to the linear equations (7)–(8).

Remark 7 In the physical literature, the dynamics of the small amplitude adjustments is generally described in terms of the vertical velocity w through (9), or equivalently, after decomposing on the basis $(\mathbf{f}_n)_{n \geq 1}$, by:

$$\forall n \in \mathbb{N}^*, \qquad \partial_t^2 w_n - c_n^2 \left(1 - \frac{c_n^2}{N^2} \Delta \right)^{-1} \Delta w_n = 0.$$

This equation can, of course, be easily deduced from (14). This configuration will be considered in Sect. 4 below where the more complex nonlinear case is considered.

The representation of the solution to the linear equations (7)–(8) based on the modal decomposition (11) and (12) brings some interesting qualitative insight on the behavior of the waves. It should be used with caution however; indeed, these decompositions involve infinite sums that without further assumptions converge slowly, and in general no better than in the L^2-sense. For instance, the modal decomposition for the density is:

$$\rho(t, X, z) = \sum_{n=1}^{\infty} \frac{1}{g} \rho_n(t, X) \rho_{\text{eq}}(z) N^2 \mathbf{f}_n(z);$$

all the terms in the summation vanish at the boundary, so that, when ρ does not vanish at the boundary, this summation *cannot* converge uniformly in, say, $L^\infty([-H, 0]; H^s(\mathbb{R}^d))$. Representing numerically the solution by a finite sum of modes with coefficients computed as in Proposition 2 is, therefore, inaccurate. This is due to the lack of regularity with respect to z of the solution considered in Proposition 2. Indeed, as it appears in the Sturm–Liouville problem (10), differentiation with respect to z is roughly equivalent to the multiplication of the coefficients of the modal expansion by a factor of size $O(1/c_n) = O(n)$. If we consider a function $F \in H^m(-H, 0)$, represented under the form:

$$F = \sum_{n=1}^{\infty} F_n \mathbf{f}_n,$$

it is, therefore, expected that $\sum n^{2m} |F_n|^2 < \infty$. As seen above, this is, however, false without additional assumptions on the behavior of F at the boundaries, because one cannot identify $\partial_z^m F$ with $\sum_{n=1}^{\infty} F_n \frac{d^m}{dz^m} \mathbf{f}_n$.

The following proposition provides such additional conditions on the initial data under which the modal decomposition is more strongly (and in particular uniformly) convergent. To state these conditions, we need to define the second-order differential operators T_1, T_2 as:

$$T_1 = \partial_z \left(\frac{1}{-\rho'_{\text{eq}}} \partial_z \left(\rho_{\text{eq}} \cdot \right) \right) \quad \text{and} \quad T_2 = \partial_z \left(\rho_{\text{eq}} \partial_z \left(\frac{1}{-\rho'_{\text{eq}}} \cdot \right) \right);$$

in the statement below, the condition on V^0 must be removed in the case $l = 0$ and we denote with a star the adjoint operator with the standard $L^2((-H, 0), \mathrm{d}z)$ scalar product.

Proposition 3 *Under the assumptions of Proposition 2, assume, moreover, that for some $k \in \mathbb{N}^*$ and $\nu = 2k$ or $2k + 1$, one has $\rho_{\text{eq}} \in W^{\nu,\infty}$, and that the solution $(U^{\text{lin}}, \rho^{\text{lin}})$ to the linear equation belongs to $C([0, T], H^{s,\nu})$, and also that its initial data (U^0, ρ^0) satisfy the additional conditions:*

$$\partial_z (T_1)^{l-1} V^0 = 0, \qquad \left(T_2^*\right)^l w^0 = 0, \quad (T_2)^l \rho^0 = 0,$$

for all $0 \le l \le k$. *Then, the following convergence holds:*

$$\sum_{n=1}^{\infty}\left[\frac{1}{c_n^{2\nu}}|\rho_n|^2_{H^{s-\nu}}+\frac{1}{c_n^{2\nu}}|V_n|^2_{H^{s-\nu}}+\frac{1}{c_n^{2\nu}}|w_n|^2_{H^{s-\nu}}\right]<\infty.$$

Proof The following key lemma shows the importance of the differential operators T_1 and T_2 introduced above.

Lemma 3 *Assume that* $N^2 = g\frac{-\rho'_{\text{eq}}}{\rho_{\text{eq}}}$ *does not depend on* z *and that* (U, ρ) *be a smooth enough solution of the linear system:*

$$\begin{cases} \partial_t V + \frac{1}{\rho_{\text{eq}}}\nabla P & = 0 \\ \partial_t w + \frac{1}{\rho_{\text{eq}}}(\partial_z P + \rho g) & = 0 \\ \partial_t \rho + \rho'_{\text{eq}} w & = 0, \\ \nabla\cdot V + \partial_z w & = 0. \end{cases} \tag{15}$$

Defining $V^\sharp$, $w^\sharp$, $\rho^\sharp$, *and* $P^\sharp$ *through:*

$$V^\sharp = T_1 V, \qquad w^\sharp = T_2^* w, \qquad \rho^\sharp = T_2\rho, \qquad P^\sharp = T_1^* P,$$

one has:

$$\begin{cases} \partial_t V^\sharp + \frac{1}{\rho_{\text{eq}}}\nabla P^\sharp & = 0 \\ \partial_t w^\sharp + \frac{1}{\rho_{\text{eq}}}(\partial_z P^\sharp + \rho^\sharp g) & = 0 \\ \partial_t \rho^\sharp + \rho'_{\text{eq}} w^\sharp & = 0, \\ \nabla\cdot V^\sharp + \partial_z w^\sharp & = 0. \end{cases}$$

Proof of the lemma The first, third, and fourth equations follow from simple computations by applying T_1 to the first equation of (15), T_2 to the third one and T_1 to the fourth one. Let us now define:

$$T_3 = \frac{1}{\rho_{\text{eq}}}\left(\rho_{\text{eq}}\partial_z\left(\frac{\rho_{\text{eq}}}{-\rho'_{\text{eq}}}\cdot\right)\right);$$

applying T_3 to the second equation in (15), one gets:

$$T_3\partial_t w + \frac{1}{\rho_{\text{eq}}}(\partial_z P^\sharp + \rho^\sharp g) = 0,$$

and the result follows from the observation that if N^2 does not depend on z, then $T_3 = T_2^*$.

It is then quite easy to check that the conditions made in the statement of the proposition on the initial data are propagated by the equations. And since the system satisfied by $(U^\sharp, \rho^\sharp, P^\sharp)$ has exactly the same structure as the original one, it is enough to prove that result for $\nu = 1$ and $\nu = 2$. The case $\nu = 1$ follows from Lemma 2 for w and P. For V, the results are obtained as for P; finally, the assumption that ρ vanishes at the boundaries is propagated from the initial condition, and the result for ρ can be obtained as for w.

We now prove the case $\nu = 2$. Since, under the assumptions of the proposition, one has $T_2^* w \in L^2([-H, 0]; H^{s-2}(\mathbb{R}^d))$, we can represent it on the basis $\mathbf{f}_n$ with coefficients:

$$(T_2^* w, \mathbf{f}_n)_{L^2_{N^2\rho eq}} = -\frac{1}{c_n}(\partial_z w, \mathbf{g}_n)_{L^2_{\rho_{\rm eq}}} = -\frac{1}{c_n^2} w_n,$$

the last equality stemming from Lemma 2. Since, moreover, $w_n = -c_n \nabla \cdot V_n$ (and that $\nabla^\perp \cdot V_n = 0$), we then deduce that $\sum_{n=1}^\infty \frac{1}{c_n^{4-2}} |V_n|_{H^{s-2k+1}} < \infty$.

Since ρ is now assumed to vanish at the boundary, there is a representation formula for $\partial_z \rho$ similar to the one provided for w in Lemma 2 and one can conclude as above. For V, there is a representation formula similar to the one given for P in Lemma 2 and the fact that $\partial_z V$ vanishes at the boundary provides a representation formula for $T_1 V$ that provides the result through Parseval identity. □

3.4 The Case of Variable *N* and Dispersive Mixing

We have already noted in Remark 5 that it is necessary that N^2 be constant in the derivation of the equations given in Proposition 2 for the coefficients of the modal decomposition. The key point was that when N^2 is constant, then one can write:

$$\begin{aligned}(\mathbf{f}_m, \mathbf{f}_n)_{L^2_{\rho_{\rm eq}}} &= \frac{1}{N^2}(\mathbf{f}_m, \mathbf{f}_n)_{L^2_{\rho_{\rm eq} N^2}}, \\ &= \frac{1}{N^2}\delta_{mn},\end{aligned}$$

where $\delta_{mn} = 1$ if $m = n$ and 0 otherwise. In the general case where N^2 depends on z, this is no longer the case and we are, therefore, led to define the interaction coefficients α_{mn} as:

$$\alpha_{mn} = (\mathbf{f}_m, \mathbf{f}_n)_{L^2_{\rho_{\rm eq}}}, \tag{16}$$

and the second equation in (14) becomes:

$$\begin{aligned}\frac{P_n}{c_n} - \rho_n &= (\partial_t w, \mathbf{f}_n)_{L^2_{\rho_{\rm eq}}}, \\ &= \sum_{m=1}^\infty \alpha_{mn} \partial_t w_m\end{aligned}$$

$$= -\sum_{m=1}^{\infty} c_m \alpha_{mn} \nabla \cdot \partial_t V_m; \tag{17}$$

we, therefore, have the following generalization of Proposition 2.

Proposition 4 *Let $\rho_{\mathrm{eq}} \in W^{1,\infty}(-H,0)$ be such that $N^2 = -\frac{\rho'_{\mathrm{eq}}}{\rho_{\mathrm{eq}}}$ is a strictly positive function on $(-H,0)$. Let also $s \geq 0$ and $T > 0$ and $(U^{\mathrm{lin}}, \rho^{\mathrm{lin}}) \in C([0,T]; H^{s,0})$ solve the linear equations (7)–(8), and assume that the vertical component of the vorticity is zero; that is, $\nabla^{\perp} \cdot V^{\mathrm{lin}} = 0$.*

The coefficients (V_n, w_n, ρ_n, P_n) provided by the decompositions (11) and (12) are found by solving, for all $n \geq 1$, the equations:

$$\begin{cases} \partial_t V_n - \sum_{m=1}^{\infty} c_m c_n \alpha_{mn} \Delta \partial_t V_m + c_n \nabla \rho_n = 0, \\ \partial_t \rho_n + c_n \nabla \cdot V_n = 0 \end{cases}$$

with initial data given by the coefficients of the decomposition of and V^0 and ρ^0, and w_n and P_n are given by:

$$w_n = -c_n \nabla \cdot V_n, \quad \text{and} \quad P_n = c_n \rho_n + c_n \sum_{m=1}^{\infty} \alpha_{mn} \partial_t w_m,$$

while for $n = 0$, one must have $P_0 = 0$ and $V_0 = 0$.

Remark 8 As in Proposition 2, one gets that the sequences $(|V_n|_{H^s})_n$ and $(|\rho_n|_{H^s})_n$ are in $l^2(\mathbb{N})$, but it is not possible to improve this convergence as in Proposition 3. Indeed, the proof of this proposition relied on Lemma 3 which requires that N^2 is constant. The best one can get is to extend to nonconstant N the case $\nu = 1$ of Proposition 3; this ensures that if $(U^{\mathrm{lin}}, \rho^{\mathrm{lin}}) \in C([0,T]; H^{s,1})$, then one has convergence of $(c_n^{-1}|V_n|_{H^{s-1}})_n$ and $(c_n^{-1}|\rho_n|_{H^{s-1}})_n$ in $l^2(\mathbb{N})$. This is fortunately enough to get uniform convergence of the modal decomposition.

Remark 9 Though mode mixing due to topography [23] or nonlinear terms [41] is known to occur, such a linear dispersive mixing does not seem to have been noticed before; the reason is that in the physical literature (e.g, [20,22]), only the hydrostatic component of the pressure is taken into account. This means that the term $\partial_t w$ is usually neglected in the second equation of (7), so that the right-hand side is zero in (17). The linear dispersive mode mixing comes, therefore, from the contribution of the time derivative of the vertical velocity to the pressure field.

Example 2 We represent in Table 1 the mixing coefficients α_{mn} for $1 \leq m, n \leq 6$ for the example of stratified ocean considered in Example 1 (the table being symmetric, we just represent the upper half). These coefficients are numerically computed from the representation of the eigenfunctions on the Chebyshev nodes described in Example 1 using the Clenshaw–Curtis quadrature for the computation of the integral in (16).

Table 1 Mixing coefficients α_{mn} for $1 \leq m, n \leq 6$ (in hour^{-2})

0.26	−0.22	−0.02	0.07	−0.05	−0.01
*	0.57	−0.13	0.01	−0.05	0.02
*	*	0.45	−0.20	−0.03	0.08
*	*	*	0.47	0.14	0.06
*	*	*	*	0.45	0.17
*	*	*	*	*	0.38

For this example, the mixing coefficients between two neighboring coefficients (n and $n \pm 1$) are smaller but of same order as the corresponding diagonal terms, while the other interactions are at least one order smaller. It follows that for perturbations for which dispersion is significant (i.e., perturbations that do not have a too large wavelength), it can trigger neighboring modes.

4 The Shallow Water Regime

We consider here some configurations of interest in oceanography, when the horizontal scale is much larger than the depth (shallow water). Under the additional *strong Boussinesq assumption* under which the density is assumed to be constant in Euler's equations, we are able to derive nonlinear models. We first derive in Sect. 4.1 the dimensionless equations, that involve two parameters: ε (nonlinearity) and μ (shallowness). The local well-posedness of these equations is granted by a straightforward adaptation of Theorem 1, but the resulting time of existence cannot be controlled as μ gets smaller, which is, by definition, the case in the shallow water regime. This problem is addressed in Sect. 4.2. We then propose in Sect. 4.3 a nonlinear extension of the modal representation of the solutions used above in the linear case.

4.1 Dimensionless Equations Under the Strong Boussinesq Assumption

We are interested here in describing the behavior of the solutions to (4)–(5) for configurations arising in oceanography with the propagation of internal waves. For such applications, the equilibrium density is of the form:

$$\rho_{\mathrm{eq}}(z) = \rho_0 + \tilde{\rho}_{\mathrm{eq}}(z),$$

with ρ_0 a constant and $\tilde{\rho}_{\mathrm{eq}} \ll \rho_0$. Therefore, the so-called *strong Boussinesq assumption* is generally made; it consists in neglecting the variation of the density everywhere except in the density equation and in buoyancy forces, so that the equations under considerations are:

$$\begin{cases} (\partial_t U + \varepsilon U \cdot \nabla_{X,z} U) = -\dfrac{1}{\rho_0} \nabla_{X,z} P - \dfrac{\rho}{\rho_0} g \mathbf{e_z}, \\ \partial_t \rho + \varepsilon U \cdot \nabla_{X,z} \rho + w \dfrac{\mathrm{d}}{\mathrm{d}z} \rho_{\mathrm{eq}} = 0, \\ \nabla_{X,z} \cdot U = 0 \end{cases} \tag{18}$$

with the boundary conditions:

$$w_{|_{z=-H}} = w_{|_{z=0}} = 0. \tag{19}$$

In most cases, the configurations under consideration are typically of "shallow water" type, meaning that horizontal scales are much larger than vertical ones. (see, for instance, [11,22,23]). This particular setting has strong consequences on the behavior of the solutions that are more easily captured by working with a dimensionless version of the equations for which the scales are adapted to the physical phenomenon under consideration.

We, therefore, define the following dimensionless variables and unknown (denoted with a tilde):

$$\begin{aligned} &X = L\widetilde{X}, \quad z = H\widetilde{z}, \quad t = \frac{L}{\sqrt{gH}}\widetilde{t}, \\ &V = \sqrt{gH}\widetilde{V}, \quad w = \sqrt{\mu}\sqrt{gH}\widetilde{w}, \quad \rho = \rho_0\widetilde{\rho}, \quad P = \rho_0 g H \widetilde{P}, \end{aligned}$$

where L is a typical horizontal scale, H is the total depth, and ρ_0 is the average volumic mass of sea-water, while $\sqrt{gH}$ correspond to the speed of the barotropic (surface) mode neglected here under our rigid lid assumption, and $\mu > 0$ is the so-called "shallowness parameter":

$$\mu = \frac{H^2}{L^2}.$$

We also introduce the quantities $\underline{\rho}$ and N^2 defined as:

$$\underline{\rho}(z) = \frac{\rho_{\mathrm{eq}}(zH)}{\rho_0}, \qquad N^2 = N^2(z) = -\frac{\mathrm{d}}{\mathrm{d}z}\underline{\rho},$$

where it is implicitly assumed that ρ_{eq} is a nonincreasing function of z (which is an obvious and classical condition for the linear stability of the equilibrium); the (depth depending) quantity N is the nondimensionalized *Brunt–Väisälä*—or buoyancy—frequency.

With these new variables and unknowns, Eqs. (4)–(5) become after dropping the tildes and separating the equations for the horizontal and vertical velocities, and in the nondimensionalized fluid domain $X \in \mathbb{R}^d$, $z \in (-1, 0)$:

$$\begin{cases} (\partial_t V + \varepsilon U \cdot \nabla_{X,z} V) = -\nabla P, \\ \mu(\partial_t w + \varepsilon U \cdot \nabla_{X,z} w) = -(\partial_z P + \rho), \\ \partial_t \rho + \varepsilon U \cdot \nabla_{X,z} \rho - N^2 w = 0, \\ \nabla_{X,z} \cdot U = 0, \end{cases} \tag{20}$$

where ∇ stands for the gradient operator taken with respect to the horizontal variables alone. These equations are complemented by the boundary conditions:

$$w_{|_{z=-1}} = w_{|_{z=0}} = 0. \tag{21}$$

4.2 Uniform Well-Posedness

One can without difficulty adapt the proof of Theorem 1 to get an existence result for the nondimensionalized system (20)–(21); however, such a proof does not provide an existence time that is uniform with respect to μ in the shallow water limit, that is, when $\mu \to 0$. The following theorem provides an existence time of order $O(\varepsilon/\sqrt{\mu})$ and which is, therefore, uniformly of size $O(1)$ when the perturbation is of *medium amplitude* ($\varepsilon = O(\sqrt{\mu})$ and uniformly of size $O(1/\sqrt{\varepsilon})$ in the *weakly nonlinear regime* $\varepsilon = O(\mu)$. Its proof is postponed to Sect. 6.3

Theorem 2 *Assume that the Brünt–Vaisälä frequency $N > 0$ is independent of z and let $k_0 \in \mathbb{N}$, $\nu = 2k_0 > d + 2$. Then, for all $\rho^0, U^0 \in H^\nu(\mathcal{S})$, such that $\nabla_{X,z} \cdot U^0 = 0$ and:*

$$\begin{cases} \forall 0 \le k \le k_0 - 1, & \partial_z^{2k} \rho^0 = \partial_z^{2k+1} V^0 = 0 \\ \forall 0 \le k \le k_0, & \partial_z^{2k} w^0 = 0 \end{cases} \quad \text{at } z = -1, 0,$$

there exists $T > 0$, such that for all $0 < \varepsilon, \mu \le 1$, there is a unique solution $(U, \rho) \in C([0, \frac{T}{\varepsilon/\sqrt{\mu}}]; H^n(\mathcal{S})^{d+2})$ to (20)–(21) with the initial data (U^0, ρ^0).

Remark 10 For surface waves (water waves), the existence time one obtains in the weakly nonlinear regime in shallow water—and for the asymptotic models derived under these assumptions (Boussinesq systems or the KdV and BBM equations, etc.)—the existence time is $O(1/\varepsilon)$, uniformly with respect to $\mu \in (0, 1)$ (see, for instance, [37] and references therein). This is also true for interfacial waves between two layers of fluids of different densities [36] and corresponding asymptotic models (see for instance [6,13] and the review [44] and references therein). The reason why the existence time is only $O(\varepsilon/\sqrt{\mu})$ is partly due to the fact that irrotationality cannot be used. Indeed, if the vertical vorticity was $O(\sqrt{\mu})$, i.e., if we had $\partial_z V = \mu\nabla w + O(\sqrt{\mu})$, then the term $[\Lambda^\nu, w]\partial_z V$ would not be singular with respect to μ anymore. This is consistent with [10] where it is shown that in the presence of vorticity, the water waves equations are well-posed on a time scale $O(\varepsilon/\sqrt{\mu})$ under the assumption that the vertical vorticity is $O(\sqrt{\mu})$.

4.3 Modal Decomposition and Nonlinear Mixing

Proceeding as for the derivation of (9), one can derive from the linear version of (20)–(21) the following equation for w:

$$\partial_t^2 \left[\left(\mu\Delta + \partial_z^2 \right) w \right] + N^2 \Delta w = 0, \tag{22}$$

where N is the Brünt–Vaisälä frequency that is now assumed to be independent of z. Instead of (10), the relevant Sturm–Liouville problem for a modal decomposition is now:

$$\begin{cases} \dfrac{d^2}{dz^2}\mathbf{f}_n + \dfrac{N^2}{c_n^2}\mathbf{f}_n = 0, \\ \mathbf{f}_n(-1) = \mathbf{f}_n(0) = 0; \end{cases} \tag{23}$$

for this Sturm–Liouville problem, the eigenvalues and the orthonormal (for the $L^2(N^2 \mathrm{d}z)$ scalar product) basis of eigenfunctions are explicit, as well as the functions $\mathbf{g}_n$ defined in Lemma 1 (which are now orthonormal for the L^2 scalar product):

$$c_n = \frac{N}{n\pi}, \qquad \mathbf{f}_n = \frac{\sqrt{2}}{N}\sin(n\pi z), \qquad \mathbf{g}_n = \sqrt{2}\cos(n\pi z).$$

The modal decompositions (11) and (12) are consequently replaced by:

$$\begin{aligned} V(t,X,z) &= \sum_{n=0}^{\infty} V_n(t,X)\mathbf{g}_n(z), \quad P(t,X,z) = \sum_{n=0}^{\infty} P_n(t,X)\mathbf{g}_n(z), \\ w(t,X,z) &= \sum_{n=0}^{\infty} w_n(t,X)\mathbf{f}_n(z), \ \ \rho(t,X,z) = \sum_{n=1}^{\infty} \rho_n(t,X)N^2\mathbf{f}_n(z). \end{aligned} \tag{24}$$

Proposition 5 *The smooth solutions of* (20)–(21) *admit a modal decomposition of the form* (24) *that satisfy the following coupled system of evolution equations up to terms of size* $O(\varepsilon\mu)$*:*

$$\begin{cases} \left(1 - \mu\dfrac{1}{\pi^2 n^2}\Delta\right)\partial_t V_n + \dfrac{N}{n\pi}\nabla\rho_n = -\varepsilon\dfrac{1}{\sqrt{2}} \displaystyle\sum_{(p\pm q)^2 = n^2} \left[V_p \cdot \nabla V_q \pm \frac{q}{p}\nabla\cdot V_p V_q \right], \\ \partial_t \rho_n + \dfrac{N}{n\pi}\nabla\cdot V_n = \varepsilon\dfrac{1}{N^2}\dfrac{1}{\sqrt{2}} \displaystyle\sum_{(p\pm q)^2 = n^2} \pm\left[V_p\cdot\nabla\rho_q + \frac{q}{p}\nabla\cdot V_p\rho_q \right]. \end{cases}$$

Remark 11 The nonlinear terms induce a mixing of the different modes of a different nature than the dispersive mixing exhibited in the linear case when N is not constant. It can be expected that after some time, the coupling becomes less and less efficient. Indeed, at first order, V_p and V_q $(p \neq q)$ travel at a different speed, and if they are localized enough, the product $V_p \cdot \nabla V_q$ (as well as the other coupling terms) becomes very small for large times. Such an asymptotic has been used for instance in diffractive optics [33].

Proof Decomposing V, w, ρ, and P as in (24) and taking the $L^2(-1,0)$ scalar product of the first equation of (20) with $\mathbf{g}_n$ gives:

$$\partial_t V_n + \nabla P_n = -\varepsilon \sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\left[\beta_{pqn} V_p \cdot \nabla V_q + \gamma_{pqn} w_p \frac{V_q}{c_q}\right]$$

with

$$\beta_{pqn} = (\mathbf{g}_p \mathbf{g}_q, \mathbf{g}_n), \qquad \gamma_{pqn} = -N^2(\mathbf{f}_p \mathbf{f}_q, \mathbf{g}_n).$$

Taking the $L^2(-1,0)$ scalar product of the second equation of (20) $\mathbf{f}_n$ gives similarly:

$$\frac{1}{N^2}\mu \partial_t w_n + \rho_n - \frac{1}{c_n} P_n = O(\varepsilon\mu).$$

For the equation on ρ, we take the scalar product with $\mathbf{f}_n$ to obtain:

$$\partial_t \rho_n - w_n = \varepsilon \frac{1}{N^2} \sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\left[\gamma_{pqn} V_p \cdot \nabla \rho_q + \gamma_{pnq} w_p \frac{\rho_q}{c_q}\right],$$

while the divergence free condition yields:

$$w_n = -c_n \nabla \cdot V_n.$$

Up to terms of order $O(\varepsilon\mu)$, we, therefore, obtain the following system on (V_n, ρ_n):

$$\begin{cases} \left(1 - \mu \frac{c_n^2}{N^2}\Delta\right) \partial_t V_n + c_n \nabla \rho_n = -\varepsilon \sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\left[\beta_{pqn} V_p \cdot \nabla V_q + \gamma_{pqn} \frac{c_p}{c_q} \nabla \cdot V_p V_q\right], \\ \partial_t \rho_n + c_n \nabla \cdot V_n = \varepsilon \frac{1}{N^2} \sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\left[\gamma_{pqn} V_p \cdot \nabla \rho_q + \gamma_{pnq} \frac{c_p}{c_q} \nabla \cdot V_p \rho_q\right]. \end{cases}$$

To get the result of the proposition, we just need to compute the coefficients β_{pqn} and γ_{pqn} using the explicit expression of the eigenfunctions derived above. One readily checks that $\sqrt{2}\mathbf{g}_p \mathbf{g}_q \mathbf{g}_n$ is equal to:

$$\begin{aligned} &\cos((p+q+n)\pi z) + \cos((p+q-n)\pi z) \\ &\quad + \cos((p-q+n)\pi z) + \cos((p-q-n)\pi z). \end{aligned}$$

Therefore, one has $\beta_{pqn} = 0$ except if $(p \pm q)^2 = n^2$ in which case $\beta_{pqn} = \frac{1}{\sqrt{2}}$. Similarly, $\gamma_{pqn} = 0$ unless $(p \pm q)^2 = n^2$, in which case $\gamma_{pqn} = \pm\frac{1}{\sqrt{2}}$. The proposition follows. □

5 Perspectives

Let us briefly mention here some interesting perspective for further works. An obvious one would be to take into account Coriolis effects as they become relevant at large oceanic scales. In particular, the dispersive effects due to the Earth rotation are known to be important (see, for instance, [11]) and it would be interesting to see how they compare to the dispersive effects investigated in this paper. In the same vein, taking into account thermal effects, salinity, etc. are important for many physical applications. For atmospheric studies, it may also be relevant to relax the incompressibility assumption; the analysis of the Sturm–Liouville decomposition is then complicated by the interaction with acoustic modes [7].

Another interesting generalization is to consider a nonzero background current, i.e., to take a nonzero U_{eq} in (3). In this context, and in the shallow water limit, Maslowe and Redekopp [42] showed that it is possible to construct approximate solutions concentrated on the first eigenmode and that the corresponding coefficient satisfies a nonlinear KdV equation. This is in sharp contrast with the nonlinear Boussinesq type system derived in Proposition 5. Indeed, in this latter, a mode n can only interact nonlinearly with modes p and q, such that $(p \pm q)^2 = n^2$. In particular, self-quadratic interaction of the coefficient of the nth mode cannot occur. This does not contradict [42], since one can check in that reference that the coefficient in front of the nonlinear term in the KdV equation vanishes when the background current is taken equal to 0. This suggests, however, that including background currents leads to additional interesting mathematical and physical phenomenons.

Let us mention a last important perspective. As said in the introduction, when the stratification is continuous, but varies very rapidly in a thin layer, two-layer models are used to describe internal waves: instead of a single continuously stratified fluid, one considers two fluids of different densities separated by an interface. The propagation of internal waves, therefore, reduces to the study of the evolution of this interface. The convergence of internal waves with a sharp continuous background stratification towards the solution of the corresponding two-fluid models is, therefore, a natural question. It was answered by the affirmative in [29] in the particular case of solitary waves. The general case of nonsteady waves is much more complex because of the presence of Kelvin–Helmholtz instabilities created by the discontinuity of the tangential velocity at the interface. In particular, two-fluid Euler equations are ill posed [18,26,30]; including surface tension effects, well-posedness is restored and a generalized Rayleigh–Taylor criterion governing well-posedness can be derived [36]. The interest of this last result is that it shows that a very small enough of surface tension is enough to stabilize interfacial waves, but its drawback is that, in the context of internal waves, there is no natural definition of surface tension. It is natural to conjecture that an alternative mechanism for the control of Kelvin–Helmholtz instabilities could come from the sharp but continuous variation of the density at the "interface" (indeed, as shown here, continuously stratified models are locally well-posed). Indeed, in the presence of a sharp stratification, the Brünt–Vaisäla is not constant and, as exhibited in Sect. 3.4, a linear mixing phenomenon occurs which, as surface tension, is of dispersive nature and could maybe control Kelvin–Helmholtz instabilities in a similar fashion. To understand this mechanism, a first step is to study the behavior of the

Sturm–Liouville modal decomposition as the background stratification ρ_{eq} converges to a discontinuous stratification.

To be more precise, let $\alpha \in C^\infty(\mathbb{R})$ be a positive function, compactly supported in $(0, 1)$ and such that $\int_{\mathbb{R}} \alpha = 1$, and define the smoothed jump function χ as $\chi(x) = \int_{-\infty}^{x} \alpha$. We consider here a strip with height $H = 1$ and a family of continuous stratifications $(\rho_{\mathrm{eq},\delta})_{0<\delta<1}$ that converges as $\delta \to 0$ to a discontinuous stratification with jump located at $z = z_0 \in (-1, 0)$ with density ρ_+ in the lower layer and ρ_- in the upper one:

$$\rho_{\mathrm{eq},\delta}(z) = \rho_+ \exp(-\delta z) - (\rho_+ - \rho_-)\chi\left(\frac{z - z_0}{\delta}\right) \qquad (\rho_+ > \rho_-), \tag{25}$$

for all $z \in [-1, 0]$, and we consider the associated Sturm–Liouville problem:

$$\frac{1}{\rho_{\mathrm{eq},\delta}} \frac{\mathrm{d}}{\mathrm{d}z}\left(\rho_{\mathrm{eq},\delta} \frac{\mathrm{d}}{\mathrm{d}z} f_{n,\delta}\right) + \frac{N^2}{c_{n,\delta}^2} f_{n,\delta} = 0, \qquad (-1 < z < 0), \tag{26}$$

with the usual boundary conditions $f_{n,\delta}(-1) = f_{n,\delta}(0) = 0$. The following proposition provides the asymptotic behavior of the first eigenvalue and of the associated eigenmode as $\delta \to 0$.

Proposition 6 *Let $\underline{c}$ and $\underline{f}$ be defined as:*

$$\underline{c}^2 = (\rho_+ - \rho_-)g\left(\frac{\rho_+}{(z_0 + 1)} + \frac{\rho_-}{(-z_0)}\right)^{-1}$$

and

$$\underline{f}(z) = \begin{cases} a(z+1) & \text{if } -1 \le z \le z_0, \\ a\frac{1+z_0}{z_0} z & \text{if } z_0 \le z \le 0, \end{cases} \qquad \text{with} \quad a = \frac{1}{\sqrt{(\rho_+ - \rho_-)g(z_0 + 1)}}.$$

For all $\delta \in (0, 1)$, let also $c_{1,\delta}^{-2}$ be the smallest eigenvalue of the Sturm–Liouville problem (26) and denote by $f_{1,\delta}$ the associated unit eigenfunction, such that $f'_{1,\delta}(-1) > 0$. Then, as $\delta \to 0$, the following approximations hold:

$$c_{1,\delta}^2 = \underline{c}^2 + O(\delta) \quad \text{and} \quad |f_{1,\delta} - \underline{f}|_{L^\infty(-1,0)} = O(\delta).$$

Proof **Step 1**. Let us first prove that $\frac{1}{c_{1,\delta}^2} \le \frac{1}{\underline{c}^2} + O(\delta)$. We recall that $c_{1,\delta}^2$ is given by the Rayleigh quotient:

$$\frac{1}{c_{1,\delta}^2} = \inf_{f \in C^1_{\mathrm{pw}}([-1,0]),\, f(-1)=f(0)=0} \frac{\int_{-1}^0 \rho_{\mathrm{eq},\delta}(f')^2}{\int_{-1}^0 \rho_{\mathrm{eq},\delta} N_\delta^2 f^2},$$

where $C^1_{\mathrm{pw}}([-1,0])$ denotes the space of continuous and piecewise C^1 functions on $[-1,0]$ and $N_\delta^2 = -g\rho'_{\mathrm{eq},\delta}/\rho_{\mathrm{eq},\delta}$. An upper bound for $c_{1,\delta}^2$ is, therefore, obtained by evaluating the quotient $R_\delta(f)$ in the right-hand-side with $f = \underline{f}$ given as in the statement of the proposition (the amplitude coefficient a is chosen, so that $\int_{-1}^0 \rho_{\mathrm{eq},\delta} N_\delta^2 \underline{f}^2 \to 1$ as $\delta \to 0$). Using the fact that $\frac{1}{\varepsilon}\chi'(\frac{z-z_0}{\varepsilon})$ is an approximation of unity converging to the Dirac distribution centered at $z = z_0$, one readily checks that $R(\underline{f}) = \underline{c}^{-2} + O(\delta)$, which proves the result.

Step 2. Let $f_{1,\delta}$ be a unit eigenfunction associated with the eigenvalue $1/c_{1,\delta}^2$ of the Sturm–Liouville problem. From the definition of $\rho_{\mathrm{eq},\delta}$, $f_{1,\delta}$ solves the following ODE on $(-1, z_0-\delta)$ and $(z_0+\delta, 1)$:

$$f''_{1,\delta} - \delta f'_{1,\delta} + g\delta f_{1,\delta} = 0,$$

from which we can deduce that:

$$f_{1,\delta}(z) = a_\delta(z+1) + O(\delta) \text{ on } [-1, z_0-\delta] \quad \text{and}$$
$$f_{1,\delta}(z) = b_\delta z + O(\delta) \text{ on } [z_0+\delta, 1]$$

with $a_\delta = O(1)$ and $b_\delta = O(1)$ as $\delta \to 0$. On the segment $[z_0-\delta, z_0+\delta]$, we can write:

$$f_{1,\delta}(z) = f_{1,\delta}(z_\pm) + (z-z_\pm) f'_{1,\delta}(z_\pm) + \int_{z_\pm}^{z} \frac{1}{\rho_{\mathrm{eq},\delta}} \int_{z_\pm}^{z'} \left(\rho_{\mathrm{eq},\delta} f'_{1,\delta}(z'')\right)' \mathrm{d}z'' \mathrm{d}z',$$

with $z_\pm = z_0 \pm \delta$. Using the Sturm–Liouville equation to simplify the last integral, one gets:

$$f_{1,\delta}(z) = f_{1,\delta}(z_\pm) + (z-z_\pm) f'_{1,\delta}(z_\pm) + \int_{z_\pm}^{z} \frac{1}{\rho_{\mathrm{eq},\delta}} \int_{z_\pm}^{z'} \frac{g}{c_{1,\delta}^2} \rho'_{\mathrm{eq},\delta}(z'') \mathrm{d}z'' \mathrm{d}z'.$$

Using the definition of $\rho_{\mathrm{eq},\delta}$ and the upper bound in $1/{c_{1,\delta}}^2$ derived in Step 1, we obtain that on $[z_0-\delta, z_0+\delta]$, one has:

$$f_{1,\delta}(z) = a_\delta(1+z_0) + O(\delta) = b_\delta z_0 + O(\delta),$$

so that $a_\delta = a + O(\delta)$, $b_\delta = a\frac{1+z_0}{z_0} + O(\delta)$ and the result follows. □

The behavior predicted by the proposition can be checked on Fig. 2b that represents the speeds associated with the first modes of the modal decomposition associated with (25) (with $\delta = 0.0005$ and $z_0 = -1/3$, in which case the formula for the asymptotic speed gives $\underline{c} \approx 0.209$) and in Fig. 4 where the first six modes are represented. A striking fact is that the formula for the asymptotic speed $\underline{c}$, which is valid even in nonshallow water configurations, coincides with the formula for the propagation of linear waves in two-fluid models in shallow water [6]. Moreover, for such two-layer models, the vertical velocity has a linear dependence in z and is continuous

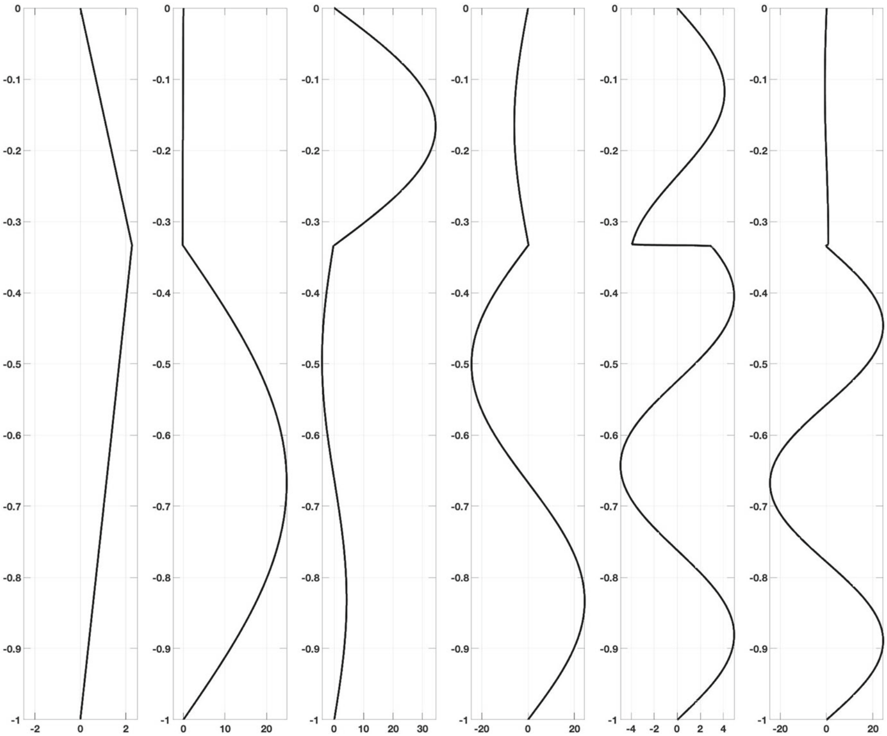


Fig. 4 The first modes associated with the stratification (25) with $\delta = 0.0005$ and $z_0 = -1/3$

at the interface: it is, therefore, a multiple of the asymptotic eigenmode $\underline{f}$. It is, therefore, natural to conjecture that one could find the two-layer shallow water model as a double shallow water/sharp stratification limit by focusing on the first mode of the Sturm–Liouville decomposition and showing that the contribution of the other modes is negligible. This is consistent with the fact that, for surface water waves, the shallow water limit is somehow a “low frequency” limit, and this would also allow one to bypass the two-fluid Euler equations and the corresponding difficulties raised by Kelvin–Helmholtz instabilities. Indeed, such instabilities do not appear for the two-fluid shallow water equations if the discontinuity of the tangential velocity at the interface is small enough [8,24].

6 Proofs of the Main Mathematical Results

We gather here the proof of several results stated in the previous sections.

6.1 Proof of Theorem 1

We prove in this section the local existence resulted of Theorem 1, reproduced below for the sake of clarity.

Theorem 1 *Let* $0 < \varepsilon \leq 1$, $\nu > d+2$, $\rho_{\rm eq} \in W^{\nu,\infty}(-H,0)$, *and* $\rho^0, U^0 \in H^\nu(\mathcal{S})$ *be such that* $\nabla_{X,z}\cdot U^0 = 0$ *and:*

$$\exists \rho_{\rm min} > 0, \qquad \inf_{z\in[-H,0]} \rho_{\rm eq}(z) \geq \rho_{\rm min} \quad \textit{and} \quad \inf_{\mathcal{S}}(\rho_{\rm eq} + \varepsilon\rho^0) \geq \rho_{\rm min},$$

Then, there exists $T > 0$, *such that for all* $\varepsilon \in (0,1]$, *there is a unique solution* $(U,\rho) \in C([0,T]; H^\nu(\mathcal{S})^{d+2})$ *to* (4)–(5) *with the initial condition* (U^0, ρ^0)*; moreover:*

$$\frac{1}{T} = c_1, \qquad \sup_{t\in[0,T]} (\|U\|_{H^\nu} + \|\rho\|_{H^\nu}) \leq c_2,$$

with $c_j = C\left(\frac{1}{\rho_{\rm min}}, \|\rho_{\rm eq}\|_{W^{\nu,\infty}}, H, \|\rho^0\|_{H^\nu}, \|U^0\|_{H^\nu}\right)$, $j = 1, 2$.

Proof *We refer to* Sect. 1.2 *for the definition of the functional spaces used throughout this proof.* We just derive here a priori estimates on solutions to (4)–(5); the construction of solutions from these energy estimates being obtained with classical means (defining the pressure in terms of ρ and U through the elliptic problem (27) below, the equations on ρ and U take the form of a quasilinear system and the a priori estimates can be used to construct solutions as in [1,45] for instance). As for the standard Euler equation, the key point is to control the pressure term which, owing to the fact that U is divergence free, is given by the resolution of the following boundary value problem:

$$\begin{cases} -\nabla_{X,z}\cdot \dfrac{1}{\rho_{\rm eq}+\varepsilon\rho}\nabla_{X,z}P = \varepsilon\nabla_{X,z}\cdot\big(U\cdot\nabla_{X,z}U\big) + \partial_z\left(\dfrac{\rho g}{\rho_{\rm eq}+\varepsilon\rho}\right), \\ -\dfrac{1}{\rho_{\rm eq}+\varepsilon\rho}\partial_z P_{|_{z=-H,0}} = \left(\dfrac{\rho g}{\rho_{\rm eq}+\varepsilon\rho}\right)_{|_{z=-H,0}}. \end{cases} \tag{27}$$

Existence of solutions to (27) follows classically from Lax–Milgram's theorem. We provide below the H^ν estimates on $\nabla_{X,z}P$ that we shall need to establish the a priori estimates on (4)–(5).

Lemma 4 *Let* $\nu > d+2$, *and* $\rho, U \in H^\nu(\mathcal{S})$. *If, moreover,* $\nabla_{X,z}\cdot U = 0$ *and if:*

$$\exists \rho_{\rm min} > 0, \qquad \inf_{z\in(-H,0)} \rho_{\rm eq}(z) \geq \rho_{\rm min} \quad \textit{and} \quad \inf_{\mathcal{S}}(\rho_{\rm eq} + \varepsilon\rho) \geq \rho_{\rm min},$$

then the solution to (27) satisfies the estimate:

$$\|\nabla_{X,z}P\|_{H^\nu} \leq C\left(\frac{1}{\rho_{\rm min}}, \|\rho_{\rm eq}\|_{W^{\nu,\infty}}, \|\rho\|_{H^\nu}\right)\left(\|\rho\|_{H^\nu} + \varepsilon\|U\|_{H^\nu}^2\right).$$

Proof *Since we work in a domain with boundaries, we need to distinguish horizontal and vertical derivatives; to this purpose, we use here the* $H^{s,k}$ *spaces introduced in* (6), *and we repeatedly use the continuous embedding* $H^{s+1/2,1} \subset L^\infty((-H,0); H^s(\mathbb{R}^d))$ *(see, for instance, Proposition 2.12 in* [35]*).*

Let us consider first the following more general boundary value problem, with $\mathbf{f} \in C([-H,0]; L^2(\mathcal{S})^3)$, such that $\mathbf{f}_{|z=-H,0} = 0$, $g \in L^2(\mathcal{S})$ and $h \in C([-H,0]; L^2(\mathbb{R}^d))$:

$$\begin{cases} -\nabla_{X,z} \cdot \frac{1}{\rho_{\rm eq} + \varepsilon\rho} \nabla_{X,z} Q = \varepsilon \nabla_{X,z} \cdot \mathbf{f} + \varepsilon |D| g + \partial_z h, \\ -\frac{1}{\rho_{\rm eq} + \varepsilon\rho} \partial_z Q_{|z=-H,0} = h_{|z=-H,0}. \end{cases} \tag{28}$$

Multiplying by Q and integrating by parts in both sides of the equation (note that the boundary terms cancel each other), we get:

$$\left(\frac{1}{\rho_{\rm eq} + \varepsilon\rho} \nabla_{X,z} Q, \nabla_{X,z} Q \right) = -\varepsilon \left(\mathbf{f}, \nabla_{X,z} Q \right) + \varepsilon (g, |D| Q) - (h, \partial_z Q),$$

from which one readily gets, with $\rho_{\max} = \left\| \rho_{\rm eq} + \varepsilon\rho \right\|_\infty$:

$$\frac{1}{\rho_{\max}} \left\| \nabla_{X,z} Q \right\|_2 \leq \varepsilon \left\| \mathbf{f} \right\|_2 + \varepsilon \left\| g \right\|_2 + \left\| h \right\|_2 . \tag{29}$$

Applying this estimate with $Q = P$, $\mathbf{f} = U \cdot \nabla_{X,z} U$, $g = 0$ and $h = \frac{\rho g}{\rho_{\rm eq} + \varepsilon\rho}$, we get the following L^2-control on the solution $\nabla_{X,z} P$ to (27):

$$\frac{1}{\rho_{\max}} \left\| \nabla_{X,z} P \right\|_2 \leq \varepsilon \left\| U \cdot \nabla_{X,z} U \right\|_2 + \left\| \frac{\rho g}{\rho_{\rm eq} + \varepsilon\rho} \right\|_2 . \tag{30}$$

Horizontal derivatives of $\nabla_{X,z} P$ can also be controlled in L^2 by applying $\Lambda^{r-1}|D|$ ($0 \leq r \leq s$ and $s > d/2 + 3/2$) to (27) and using (28) with $Q = \Lambda^{r-1}|D|P$, $\mathbf{f} = \frac{1}{\varepsilon}[\Lambda^{r-1}|D|, \frac{1}{\rho_{\rm eq}+\varepsilon\rho}]\nabla_{X,z} Q$, $g = \Lambda^{r-1}\nabla_{X,z} \cdot (U \cdot \nabla_{X,z} U)$, and $h = \Lambda^{r-1}|D| \left(\frac{\rho g}{\rho_{\rm eq}+\varepsilon\rho} \right)$. The estimate (29) then yields:

$$\begin{aligned} \frac{1}{\rho_{\max}} \left\| \Lambda^{r-1} |D| \nabla_{X,z} P \right\|_2 \leq \varepsilon & \left\| \frac{1}{\varepsilon} \left[\Lambda^{r-1}|D|, \frac{1}{\rho_{\rm eq} + \varepsilon\rho} \right] \nabla_{X,z} P \right\|_2 \\ & + \varepsilon \left\| \Lambda^{r-1} \nabla_{X,z} \cdot (U \cdot \nabla_{X,z} U) \right\|_2 \\ & + \left\| \Lambda^{r-1} |D| \left(\frac{\rho g}{\rho_{\rm eq} + \varepsilon\rho} \right) \right\|_2 ; \end{aligned} \tag{31}$$

we now turn to control the three terms in the right-hand side:

Control of $\left\| \frac{1}{\varepsilon}[\Lambda^{r-1}|D|, \frac{1}{\rho_{\rm eq}+\varepsilon\rho}]\nabla_{X,z} P \right\|_2$. Remarking that $\frac{1}{\rho_{\rm eq}+\varepsilon\rho} = \frac{1}{\rho_{\rm eq}} - \frac{\varepsilon\rho}{\rho_{\rm eq}(\rho_{\rm eq}+\varepsilon\rho)}$ and that $[\Lambda^{r-1}|D|, \frac{1}{\rho_{\rm eq}}] = 0$, we have:

$$\left\| \frac{1}{\varepsilon} \left[\Lambda^{r-1}|D|, \frac{1}{\rho_{\rm eq} + \varepsilon\rho} \right] \nabla_{X,z} P \right\|_2 = \left\| \left[\Lambda^{r-1}|D|, \frac{\rho}{\rho_{\rm eq}(\rho_{\rm eq} + \varepsilon\rho)} \right] \nabla_{X,z} P \right\|_2$$

$$\leq C\left(\frac{1}{\rho_{\min}}, \|\rho\|_{W^{1,\infty}}\right) \|\rho\|_{H^{s,1}} \left\|\nabla_{X,z} P\right\|_{H^{r-1/2,1}},$$

where we used the commutator estimate (39) and the assumption that $s > d/2 + 3/2$ to derive the second inequality.

Control of $\left\|\Lambda^{r-1}\nabla_{X,z}\cdot(U\cdot\nabla_{X,z}U)\right\|_2$. Recalling that $\nabla_{X,z}\cdot U = 0$, we can write classically:

$$\nabla_{X,z}\cdot\left(U\cdot\nabla_{X,z}U\right) = \sum_{i,j=x,y,z}\partial_i U_j \partial_j U_i,$$

and, therefore, using the product estimate (35) and the assumption that $s > d/2+3/2$ (recall also that $0 \leq r \leq s$):

$$\left\|\Lambda^{r-1}\nabla_{X,z}\cdot(U\cdot\nabla_{X,z}U)\right\|_2 \leq \|U\|^2_{H^{s,2}}.$$

Control of $\left\|\Lambda^{r-1}|D|\left(\frac{\rho g}{\rho_{\mathrm{eq}}+\varepsilon\rho}\right)\right\|_2$. Without any difficulty, one gets:

$$\left\|\Lambda^{r-1}|D|\left(\frac{\rho g}{\rho_{\mathrm{eq}}+\varepsilon\rho}\right)\right\|_2 \leq C\left(\frac{1}{\rho_{\min}}, \|\rho\|_{W^{1,\infty}}\right)\|\rho\|_{H^{s,1}}.$$

Gathering these three estimates, together with (30) and (31), we finally get:

$$\begin{aligned}\left\|\nabla_{X,z}P\right\|_{H^{r,0}} \leq & C\left(\frac{1}{\rho_{\min}}, \left\|\rho_{\mathrm{eq}}\right\|_\infty, \|\rho\|_{W^{1,\infty}}\right)\\ &\times\left(\|\rho\|_{H^{s,1}} + \varepsilon\|U\|^2_{H^{s,2}} + \varepsilon\left\|\nabla_{X,z}P\right\|_{H^{r-1/2,1}}\right),\end{aligned}$$

for all $0 \leq r \leq s$. We now prove that it is possible to obtain an estimate on the $H^{r,1}$-norm instead of the $H^{r,0}$ norm in the left-hand side, at the cost of increasing slightly the regularity in z of ρ. Using the equation satisfied by P, one has:

$$\begin{aligned}-\partial_z^2 P =& (\rho_{\mathrm{eq}}+\varepsilon\rho)\partial_z\left(\frac{1}{\rho_{\mathrm{eq}}+\varepsilon\rho}\right)\partial_z P + (\rho_{\mathrm{eq}}+\varepsilon\rho)\nabla\cdot\frac{1}{\rho_{\mathrm{eq}}+\varepsilon\rho}\nabla P\\ &+\varepsilon(\rho_{\mathrm{eq}}+\varepsilon\rho)\sum_{i,j=x,y,z}\partial_i U_j\partial_j U_i, +(\rho_{\mathrm{eq}}+\varepsilon\rho)\partial_z\left(\frac{\rho g}{\rho_{\mathrm{eq}}+\varepsilon\rho}\right);\end{aligned} \tag{32}$$

controlling the right-hand side in $H^{r-1,0}$ through (35) shows that $\left\|\Lambda^{r-1}\partial_z^2 P\right\|_{H^{r,0}}$ is bounded from above by:

$$C\left(\frac{1}{\rho_{\min}}, \left\|\rho_{\mathrm{eq}}\right\|_{W^{1,\infty}}, \|\rho\|_{W^{1,\infty}}\right)\left(\|\rho\|_{H^{s,2}} + \varepsilon\|U\|^2_{H^{s,2}} + \varepsilon\left\|\nabla_{X,z}P\right\|_{H^{r-1/2,1}}\right),$$

and we have, therefore, obtained that:

$$\left\|\nabla_{X,z} P\right\|_{H^{r,1}} \leq C\left(\frac{1}{\rho_{\min}}, \left\|\rho_{\mathrm{eq}}\right\|_{W^{1,\infty}}, \|\rho\|_{W^{1,\infty}}\right) \times \left(\|\rho\|_{H^{s,2}} + \varepsilon \|U\|_{H^{s,2}}^2 + \varepsilon \left\|\nabla_{X,z} P\right\|_{H^{r-1/2,1}}\right);$$

by a simple finite induction on r, we, therefore, get:

$$\left\|\nabla_{X,z} P\right\|_{H^{s,1}} \leq C\left(\frac{1}{\rho_{\min}}, \left\|\rho_{\mathrm{eq}}\right\|_{W^{1,\infty}}, \|\rho\|_{W^{1,\infty}}\right)\left(\|\rho\|_{H^{s,2}} + \varepsilon \|U\|_{H^{s,2}}^2\right). \quad (33)$$

Using the product estimate (38), we also get from (32) that for all $0 \leq k \leq \nu$:

$$\begin{aligned} \left\|\nabla_{X,z} P\right\|_{H^{\nu,k}} &\leq \left\|\nabla_{X,z} P\right\|_{H^{\nu,1}} + \left\|\partial_z^2 P\right\|_{H^{\nu-1,k-1}} \\ &\leq C\left(\frac{1}{\rho_{\min}}, \left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}}, \|\rho\|_{H^{\nu}}\right) \\ &\quad (\|\rho\|_{H^{\nu}} + \varepsilon \|U\|_{H^{\nu}}^2 + \left\|\nabla_{X,z} P\right\|_{H^{\nu,k-1}}) \end{aligned}$$

(we used here the assumption that $\nu > d+2$). The result then follows from an induction on k and (33). □

Applying $\Lambda^{\nu-k}\partial_z^k$ ($k = 0, 1, 2$) to the equations of (4), we get:

$$\begin{cases} (\partial_t \tilde{U}_k + \varepsilon U \cdot \nabla_{X,z} \tilde{U}_k) = F_1 + F_2, \\ \partial_t \tilde{\rho}_k + \varepsilon U \cdot \nabla_{X,z} \tilde{\rho}_k = f_1 + f_2, \\ \nabla_{X,z} \cdot \tilde{U}_k = 0 \end{cases}$$

where we denoted $\tilde{U}_k = \Lambda^{\nu-k}\partial_z^k U$ and $\tilde{\rho}_k = \Lambda^{\nu-k}\partial_z^k \rho$, and where

$$\begin{aligned} F_1 &= -\Lambda^{\nu-k}\partial_z^k\left(\frac{1}{\rho_{\mathrm{eq}} + \varepsilon\rho}\nabla_{X,z} P\right) - \Lambda^{\nu-k}\partial_z^k\left(\frac{\rho}{\rho_{\mathrm{eq}} + \varepsilon\rho} g\mathbf{e_z}\right), \\ F_2 &= -\varepsilon(\Lambda^{\nu-k}\partial_z^k, U) \cdot \nabla_{X,z} U \end{aligned}$$

and

$$f_1 = -\Lambda^{\nu-k}\partial_z^k\left(w\frac{\mathrm{d}}{\mathrm{d}z}\rho_{\mathrm{eq}}\right), \qquad f_2 = -\varepsilon[\Lambda^{\nu-k}\partial_z^k, U] \cdot \nabla_{X,z}\rho.$$

Multiplying the first equation by $\tilde{U}_k$ and the second one by $\tilde{\rho}_k$, and integrating by parts, we get (recall that $\nabla_{X,z} \cdot U = 0$ and that w vanishes at the boundaries):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\left\|\tilde{U}_k\right\|_2^2+\frac{1}{2}\left\|\tilde{\rho}_k\right\|_2^2\right)=(F_1+F_2,\tilde{U})+(f_1+f_2,\tilde{\rho})$$
$$\leq(\|F_1\|_2+\|F_2\|_2)\left\|\tilde{U}_k\right\|_2+(\|f_1\|_2+\|f_2\|_2)\left\|\tilde{\rho}_k\right\|_2.$$

We easily get from the product estimate (38) and Lemma 4 (for F_1 and f_1) and the commutator estimate (41) (for F_2 and f_2) that:

$$\|F_1\|_2+\|F_2\|_2+\|f_1\|_2+\|f_2\|_2\leq C\left(\frac{1}{\rho_{\min}},\left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}},\|\rho\|_{H^\nu},\|U\|_{H^\nu}\right),$$

and therefore:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\tilde{U}_k\right\|_2^2+\left\|\tilde{\rho}_k\right\|_2^2\right)\leq C\left(\frac{1}{\rho_{\min}},\left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}},\|\rho\|_{H^n},\|U\|_{H^\nu}\right)\left(\left\|\tilde{U}_k\right\|_2+\left\|\tilde{\rho}_k\right\|_2\right),$$
$$\leq C\left(\frac{1}{\rho_{\min}},\left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}},\|\rho\|_{H^\nu},\|U\|_{H^\nu}\right).$$

Summing over all $0\leq k\leq\nu$, this yields:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|U\|_{H^\nu}^2+\|\rho\|_{H^\nu}^2\right)\leq C\left(\frac{1}{\rho_{\min}},\left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}},\|\rho\|_{H^\nu},\|U\|_{H^\nu}\right),$$

which is the desired a priori estimate. □

6.2 Proof of Proposition 1

We recall below the statement of Proposition 1 on the validity of the linear approximation and prove it.

Proposition 1 *Let the assumptions of Theorem 1 be satisfied and denote by $T>0$ the existence time of the solution (U,ρ) of the nonlinear problem (4)–(5) provided by this theorem. There exists also a unique solution $(U^{\mathrm{lin}},\rho^{\mathrm{lin}})\in C([0,T];H^\nu(\mathcal{S})^{d+2})$ to the linear problem (7)–(8) with the same initial data, and the following error estimate holds:*

$$\|(\rho-\rho^{\mathrm{lin}},U-U^{\mathrm{lin}})\|_{L^\infty([0,T]\times H^{\nu-1})}\leq c_3\varepsilon$$

with $c_3=C\left(T,\frac{1}{\rho_{\min}},\left\|\rho_{\mathrm{eq}}\right\|_{W^{\nu,\infty}},H,\left\|\rho^0\right\|_{H^\nu},\left\|U^0\right\|_{H^\nu}\right)$.

Proof It is obvious that the solution to the linear system is defined on the same time interval as the solution to the nonlinear system provided by Theorem 1. We need to prove the error estimate. Let us decompose the pressure P in the nonlinear problem under the form $P=P^{\mathrm{L}}+\varepsilon P^{\mathrm{NL}}$, with:

$$\begin{cases} -\nabla_{X,z} \cdot \dfrac{1}{\rho_{\rm eq}} \nabla_{X,z} P^{\rm L} = \partial_z \left(\dfrac{\rho g}{\rho_{\rm eq}} \right), \\ -\dfrac{1}{\rho_{\rm eq}} \partial_z P^{\rm L}_{|z=-H,0} = \left(\dfrac{\rho g}{\rho_{\rm eq}} \right)_{|z=-H,0}, \end{cases}$$

and

$$\begin{cases} -\nabla_{X,z} \cdot \dfrac{1}{\rho_{\rm eq} + \varepsilon\rho} \nabla_{X,z} P^{\rm NL} = \nabla_{X,z} \cdot F, \\ -\dfrac{1}{\rho_{\rm eq} + \varepsilon\rho} \partial_z P^{\rm NL}_{|z=-H,0} = F_{\rm v}, \end{cases}$$

where

$$F = U \cdot \nabla_{X,z} U - \frac{\rho^2}{\rho_{\rm eq}(\rho_{\rm eq} + \varepsilon\rho)} g \mathbf{e_z} - \frac{\rho}{\rho_{\rm eq}(\rho_{\rm eq} + \varepsilon\rho)} \nabla_{X,z} P^{\rm L},$$

and $F_{\rm h}$ and $F_{\rm v}$ denote its horizontal and vertical components. The difference $\widetilde{U} = U - U^{\rm lin}$, $\widetilde{\rho} = \rho - \rho^{\rm lin}$ and $\widetilde{P} = P^{\rm L} - P^{\rm lin}$ satisfies therefore:

$$\begin{cases} \partial_t \widetilde{V} + \dfrac{1}{\rho_{\rm eq}} \nabla \widetilde{P} = -\varepsilon \dfrac{1}{\rho_{\rm eq} + \varepsilon\rho} \nabla P^{\rm NL} + \varepsilon F_{\rm h} \\ \partial_t \widetilde{w} + \dfrac{1}{\rho_{\rm eq}} \partial_z \widetilde{P} + \dfrac{\widetilde{\rho}}{\rho_{\rm eq}} g = -\varepsilon \dfrac{1}{\rho_{\rm eq} + \varepsilon\rho} \partial_z P^{\rm NL} + \varepsilon F_{\rm v}, \\ \partial_t \widetilde{\rho} + \widetilde{w} \dfrac{\rm d}{{\rm d}z} \rho_{\rm eq} = \varepsilon f, \end{cases}$$

with $f = -\varepsilon U \cdot \nabla_{X,z} \rho$. Proceeding as for Lemma 4, one gets:

$$\|\nabla_{X,z} \widetilde{P}\|_{H^{\nu-1}} \lesssim \|\widetilde{\rho}\|_{H^{\nu-1}} \quad \text{and} \quad \|\nabla_{X,z} P^{\rm NL}\|_{H^{\nu-1}} \leq C\left(\|(U,\rho)\|_{H^\nu}\right),$$

while F and f are also bounded from above in $H^{\nu-1}(\mathcal{S})$ by $C\big(\|(U,\rho)\|_{H^\nu}\big)$. It follows as in the proof of Theorem 1 that:

$$\frac{\rm d}{{\rm d}t}\left(\|\widetilde{U}\|^2_{H^{\nu-1}} + \|\widetilde{\rho}\|^2_{H^{\nu-1}} \right) \leq C\|\widetilde{\rho}\|^2_{H^{\nu-1}} + \varepsilon C\left(\|(U,\rho)\|_{H^\nu}\right)\left(\|\widetilde{U}\|_{H^{\nu-1}} + \|\widetilde{\rho}\|_{H^{\nu-1}} \right),$$

and the estimate provided in the proposition classically follows using a Gronwall-type argument. □

6.3 Proof of Theorem 2

We recall below the statement of Theorem 2 before proving it. This theorem provides a control of the time of existence and on the size of the solution in terms of the parameters ε and μ that appear in the dimensionless version of the equations.

Theorem 2 *Assume that the Brünt–Vaisälä frequency $N > 0$ is independent of z, and let $k_0 \in \mathbb{N}$, $\nu = 2k_0 > d + 2$. Then, for all $\rho^0, U^0 \in H^\nu(\mathcal{S})$, such that $\nabla_{X,z} \cdot U^0 = 0$ and:*

$$\begin{cases} \forall 0 \le k \le k_0 - 1, & \partial_z^{2k} \rho^0 = \partial_z^{2k+1} V^0 = 0 \\ \forall 0 \le k \le k_0, & \partial_z^{2k} w^0 = 0 \end{cases} \quad at\ z = -1, 0,$$

there exists $T > 0$, such that for all $0 < \varepsilon, \mu \le 1$, there is a unique solution $(U, \rho) \in C([0, \frac{T}{\varepsilon/\sqrt{\mu}}]; H^n(\mathcal{S})^{d+2})$ to (20)–(21) with initial condition: (U^0, ρ^0).

Proof A key step in the proof is that w and its vertical derivatives of even order vanish at the boundary. We will use the following lemma.

Lemma 5 *If (U, ρ) is a smooth solution to (20)–(21) on a time interval $[0, T]$, such that for some $k_0 \in \mathbb{N}$, the initial data satisfy:*

$$\forall 0 \le j \le k_0, \quad \partial_z^{2j} w^0 = \partial_z^{2j} \rho^0 = \partial_z^{2j+1} V^0 = 0 \quad at\ z = -1, 0.$$

These cancelations are propagated by the equations on the time interval $[0, T]$. Moreover, if $\partial_z^{2k_0+2} w^{\text{in}}{}_{|z=-1,0} = 0$, then $\partial_z^{2k_0+2} w_{|z=-1,0} = 0$ on $[0, T]$.

Proof of the lemma Let us prove first the following identities for $0 \le j \le k_0$:

$$(\partial_z^{2j} w)_{|z=0,-1} = 0, \qquad (\partial_z^{2j} \rho)_{|z=0,-1} = 0 \qquad (\partial_z^{2j+1} P) = 0, \qquad \partial_z^{2j+1} V = 0,$$

at $z = -1, 0$. Proceeding by induction, we assume that these identities hold for $0 \le j \le k \le k_0 - 1$ and want to prove that they hold for $j = k + 1$. We recall that:

$$-(\partial_z^2 + \mu \Delta) P = \partial_z \rho + \varepsilon \mu \nabla_{X,z} \cdot (U \cdot \nabla_{X,z} U),$$

so that, after differentiating $2k + 1$ times with respect to z:

$$-\partial_z^{2k+2}(\partial_z P + \rho) = \varepsilon \mu \partial_z^{2k+2}(U \cdot \nabla_{X,z} w) + \varepsilon \mu \partial_z^{2k+1} \nabla \cdot (U \cdot \nabla_{X,z} V). \quad (34)$$

Applying ∂_z^{2k+2} to the equation on w, one also gets:

$$\mu \partial_t \partial_z^{2k+2} w + \varepsilon \mu \partial_z^{2k+2}(U \cdot \nabla_{X,z} w) = -\partial_z^{2k+2}(\partial_z P + \rho),$$

which, together with (34), yields:

$$\begin{aligned} \mu \partial_t \partial_z^{2k+2} w &= \varepsilon \mu \partial_z^{2k+1} \nabla \cdot (U \cdot \nabla_{X,z} V) \\ &= \varepsilon \mu \partial_z^{2k+1} \left[V \cdot \nabla(\nabla \cdot V) + w \partial_z (\nabla \cdot V) + \sum_{l=1}^{d} \nabla V_l \cdot \partial_l V + \nabla w \cdot \partial_z V \right]. \end{aligned}$$

Recalling that $\nabla \cdot V = -\partial_z w$, taking the trace of the above identity at $z = -1, 0$ and using the induction assumption, one deduces:

$$\mu(\partial_t + \varepsilon V \cdot \nabla + \varepsilon(2k+1)(\partial_z w))\partial_z^{2k+2} w = 0 \quad \text{at} \quad z = -1, 0.$$

Therefore, if $\partial_z^{2k+2} w = 0$ at $t = 0$ on $z = -1, 0$, it remains equal to 0 for all times, which proves the first relation of the induction assumption for $j = k+1$. We can now apply ∂_z^{2k+2} to the equation on ρ and take the trace at the boundaries to obtain:

$$(\partial_t + \varepsilon V \cdot \nabla)\partial_z^{2k+2} \rho = 0$$

where we used the fact that $\partial_z^{2k+2}(N^2 w) = 0$ at $z = -1, 0$ for N^2 chosen as in the statement of the lemma and the fact that $\partial_z^{2j} w = 0$ at $z = -1, 0$ for $0 \leq j \leq k+1$. This shows that the condition $\partial_z^{2k+2} \rho = 0$ at $z = -1, 0$ propagates from the initial data and establishes the second relation of the induction assumption for $j = k+1$. Taking the trace of (34) at $z = -1, 0$ and using the results just proved yields the third relation. For the fourth one, we apply ∂_z^{2k+3} to the equation on V and take the trace at the boundaries to obtain:

$$(\partial_t + \varepsilon V \cdot \nabla + \varepsilon(2k+3)(\partial_z w))\partial_z^{2k+3} V = 0 \quad \text{at } z = -1, 0,$$

showing that the desired relation is propagated from the initial condition.

To conclude the induction, we need to show that the assumption is true for $j = 0$. The condition on w is obvious, and the fact that ρ, $\partial_z P$ and $\partial_z V$ also vanish at the boundaries can be deduced proceeding as above. The last point of the lemma is established by running the first step of the induction proof for $k = k_0 + 1$.

Applying $\Lambda^{\nu-2k}\partial_z^{2k}$ $(0 \leq k \leq k_0)$ to the equations of (20), we get:

$$\begin{cases} (\partial_t \tilde{V}_k + \varepsilon U \cdot \nabla_{X,z} \tilde{V}_k) = -\nabla \tilde{P}_k + \varepsilon F_{\rm h}, \\ \mu(\partial_t \tilde{w}_k + \varepsilon U \cdot \nabla_{X,z} \tilde{w}_k) = -(\partial_z \tilde{P}_k + \tilde{\rho}_k) + \varepsilon\mu F_{\rm v}, \\ \partial_t \tilde{\rho}_k + \varepsilon U \cdot \nabla_{X,z} \tilde{\rho}_k - N^2 \tilde{w}_k = \varepsilon f, \\ \nabla_{X,z} \cdot \tilde{U}_k = 0, \end{cases}$$

where we denoted $\tilde{U}_k = \Lambda^{\nu-k}\partial_z^{2k} U$ and $\tilde{\rho}_k = \Lambda^{\nu-k}\partial_z^{2k}\rho$, and where:

$$F_{\rm h} = -[\Lambda^{\nu-2k}\partial_z^{2k}, U] \cdot \nabla_{X,z} V, \qquad F_{\rm v} = -[\Lambda^{\nu-2k}\partial_z^{2k}, U] \cdot \nabla_{X,z} w,$$

$$f = -[\Lambda^{\nu-2k}\partial_z^{2k}, U] \cdot \nabla_{X,z}\rho.$$

By the lemma, we also have the boundary conditions:

$$\tilde{w}_{k\ |z=-1} = \tilde{w}_{k\ |z=0} = 0.$$

Multiplying the first equation by $\tilde{U}_k$ and the second one by $\tilde{\rho}_k$, and integrating by parts, we, therefore, get (recall that $\nabla_{X,z} \cdot \tilde{U}_k = 0$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{E}_k \leq \varepsilon \|F_{\mathrm{h}}\|_2 \left\|\tilde{V}_h\right\|_2 + \varepsilon\sqrt{\mu}\,\|F_{\mathrm{v}}\|_2 \sqrt{\mu}\,\|\tilde{w}_k\|_2 + \varepsilon \|f\|_2 \|\tilde{\rho}_k\|_2$$
$$\leq \varepsilon \left(\|F_{\mathrm{h}}\|_2 + \sqrt{\mu}\,\|F_{\mathrm{v}}\|_2 + \|f\|_2\right) \tilde{E}_k^{1/2},$$

where

$$\tilde{E}_k = \frac{1}{2}\left\|\tilde{V}_k\right\|_2^2 + \frac{1}{2}\mu\,\|\tilde{w}_k\|_2^2 + \frac{1}{2}\,\|\tilde{\rho}_k\|_2^2 .$$

With $\nu \geq 2t_0 + 2$, the commutator estimates of Appendix A imply that:

$$\|F_{\mathrm{h}}\|_2 + \sqrt{\mu}\|F_{\mathrm{v}}\|_2 + \|f\|_2 \lesssim \frac{1}{\sqrt{\mu}}\left(\|V\|_{H^\nu} + \sqrt{\mu}\,\|w\|_{H^\nu} + \|\rho\|_{H^\nu}\right)^2;$$

we note that the singular factor $\mu^{-1/2}$ comes from the components $-[\Lambda^\nu, w]\partial_z V$ of F_{h} and $-[\Lambda^\nu, w]\partial_z \rho$ of f in the case $k = 0$.

Summing over $0 \leq k \leq k_0$ and using the above estimates, it follows that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{k_0}\tilde{E}_k \lesssim \frac{\varepsilon}{\sqrt{\mu}}\left(\|V\|_{H^\nu} + \sqrt{\mu}\,\|w\|_{H^\nu} + \|\rho\|_{H^\nu}\right)^2 \left(\sum_{k=0}^{k_0}\tilde{E}_k\right)^{1/2}.$$

It is, therefore, possible to conclude by a nonlinear Gronwall-type argument if we are able to show that:

$$\left(\|V\|_{H^\nu} + \sqrt{\mu}\,\|w\|_{H^\nu} + \|\rho\|_{H^\nu}\right)^2 \lesssim \sum_{k=0}^{k_0}\tilde{E}_k,$$

which is a direct consequence of the following lemma.

Lemma 6 ***i.*** *There is a constant C, such that for all smooth functions V, w, and ρ such that:*

$$\partial_z^{2k+1}V_{|z=-1,0} = \partial_z^{2k}w_{|z=-1,0} = \partial_z^{2k}\rho_{|z=-1,0} + 0 \quad \mathit{for} \quad k = 0, \ldots, k_0 - 1,$$

one has, with $u = V$, w or ρ:

$$\frac{1}{C}\|u\|_{H^\nu} \leq \sum_{k=0}^{k_0}\|\Lambda^{\nu-2k}\partial_z^{2k}u\|_2 \leq C\|u\|_{H^\nu}.$$

Proof of the lemma We need to show that it is possible to control $\Lambda^{\nu-2k-1}\partial_z^{2k+1}V$ $(1 \leq k \leq k_0)$ in $L^2(\mathcal{S})$ by $\sum_{k=0}^{k_0}\|\Lambda^{\nu-2k}\partial_z^{2k}V\|_2$. This is a consequence of Lemma 5 which, under the assumption made on the vanishing of the vertical derivatives of V of odd order, allows one to write:

$$\begin{aligned}\|\Lambda^{\nu-2k-1}\partial_z^{2k+1}V\|_2^2 &= \int_{\mathcal{S}} \Lambda^{\nu-2k-1}\partial_z^{2k+1}V \cdot \Lambda^{\nu-2k-1}\partial_z^{2k+1}V,\\ &= -\int_{\mathcal{S}} \Lambda^{\nu-2k}\partial_z^{2k}V \cdot \Lambda^{\nu-2k-2}\partial_z^{2k+2}V,\\ &\leq \left\|\Lambda^{\nu-2k}\partial_z^{2k}V\right\|_2 \left\|\Lambda^{\nu-2k-2}\partial_z^{2k+2}V\right\|_2.\end{aligned}$$

The result for w and ρ is proved similarly.

The proof of the theorem is, therefore, complete. □

Acknowledgements D. L. would like to thank J.-F. Bony and N. Popoff for fruitful discussions about this work.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

A Product and Commutator Estimates

We use here the notations:

$$A = B + \langle C\rangle_{s>s_0}$$

to mean that $A = B$ if $s \leq s_0$, and $A = B + C$ otherwise.

From the following classical product estimate for functions on $\mathbb{R}^d$:

$$\forall t_0 > d/2, \quad \forall s \geq 0, \qquad |fg|_{H^s} \lesssim |f|_{H^s}|g|_{H^{t_0}} + \big\langle |f|_{H^{t_0}}|g|_{H^s}\big\rangle_{s>t_0},$$

we deduce from $L^\infty \times L^2$ product estimates on the strip $\mathcal{S}$ and the continuous embedding $H^{s+1/2,1} \subset L^\infty H^s$ ($s \in \mathbb{R}$) that functions on the strip $\mathcal{S}$ satisfy the following product estimate, for all $t_0 > d/2$, and $s \geq 0$:

$$\|FG\|_{H^{s,0}} \lesssim \|F\|_{H^{s,0}} \|G\|_{H^{t_0+1/2,1}} + \big\langle \|F\|_{H^{t_0+1/2,1}} \|G\|_{H^{s,0}}\big\rangle_{s>t_0}, \tag{35}$$

It is possible to deduce another product estimate from the product estimate on $\mathbb{R}^d$; for instance:

$$\|FG\|_{H^{s,0}} \lesssim \|F\|_{H^{s,0}} \|G\|_{H^{t_0+1/2,1}} + \big\langle \|F\|_{H^{t_0,0}} \|G\|_{H^{s+1/2,1}}\big\rangle_{s>t_0}; \tag{36}$$

in particular, if $r_0 = t_0$ if $s \leq t_0$ and $r_0 = s$ otherwise, then:

$$\|FG\|_{H^{s,0}} \lesssim \|F\|_{H^{r_0,0}} \|G\|_{H^{r_0+1/2,1}}. \tag{37}$$

We also need in this paper product estimates in $H^{s,k}$. Remarking that:

$$\|FG\|_{H^{s,k}} \lesssim \sum_{k'+k''=k} \left\|(\partial_z^{k'} F)(\partial_z^{k''} G)\right\|_{H^{s-k,0}},$$

we easily deduce from (36) that:

$$\forall t_0 > d/2, \quad \forall s \geq 2t_0 + 1, \quad \forall 0 \leq k \leq s, \qquad \|FG\|_{H^{s,k}} \lesssim \|F\|_{H^{s,k}} \|G\|_{H^{s,k}}. \tag{38}$$

For commutators with horizontal derivatives, we first recall the following estimate that combines the Kato–Ponce estimates (for large s) and Coifmann–Meyer estimates (for small s), for functions defined over $\mathbb{R}^d$:

$$\forall t_0 > d/2, \quad \forall s \geq 0, \qquad \left|[\Lambda^s, f]g\right|_2 \lesssim |f|_{H^{t_0+1}} |g|_{H^{s-1}} + \left\langle |f|_{H^s} |g|_{H^{t_0}} \right\rangle_{s>t_0+1}$$

(see, for instance, Theorems 3 and 6 in [34]). Using the continuous embedding $H^{s+1/2,1} \subset L^\infty H^s$ ($s \in \mathbb{R}$), one readily deduces the following estimate for functions defined in the strip $\mathcal{S}$, and for all $t_0 > d/2$, and $s \geq 0$:

$$\left\|[\Lambda^s, F]G\right\|_2 \lesssim \|F\|_{H^{t_0+3/2,1}} \|G\|_{H^{s-1,0}} + \left\langle \|F\|_{H^{s,0}} \|G\|_{H^{t_0+1/2,1}} \right\rangle_{s>t_0+1} \tag{39}$$

or also

$$\left\|[\Lambda^s, F]G\right\|_2 \lesssim \|F\|_{H^{t_0+3/2,1}} \|G\|_{H^{s-1,0}} + \left\langle \|F\|_{H^{s+1/2,1}} \|G\|_{H^{t_0,0}} \right\rangle_{s>t_0+1} \tag{40}$$

(see, for instance, §B.2.2. in [35]).

Let us finally consider commutators with horizontal *and* vertical derivatives, namely, with $\Lambda^{s-k}\partial_z^k$ ($k = 1, 2$), we first remark that:

$$[\Lambda^{s-k}\partial_z^k, F]G = \Lambda^{s-k}[\partial_z^k, F]G + [\Lambda^{s-k}, F]\partial_z^k G.$$

Using the product estimates (35)–(36) for the first term of the right-hand side, and the commutator estimates (39)–(40) for the second one, we get:

$$\forall n \geq 2t_0 + 2, \quad \forall 0 \leq k \leq n, \qquad \left\|[\Lambda^{n-k}\partial_z^k, F]G\right\|_2 \lesssim \|F\|_{H^{n-1}} \|G\|_{H^{n-1}}. \tag{41}$$

References

1. Alinhac, S., Gérard, P.: Opérateurs pseudo-différentiels et théorème de Nash-Moser. Savoirs Actuels. InterEditions, Paris (1991)
2. Barros, R., Choi, W.: Inhibiting shear instability induced by large amplitude internal solitary waves in two-layer flows with a free surface. Stud. Appl. Math. **122**, 325–346 (2009)
3. Barros, R., Choi, W.: On regularizing the strongly nonlinear model for two-dimensional internal waves. Phys. D **264**, 27–34 (2013)

4. Benjamin, T.B.: Internal waves of finite amplitude and permanent form. J. Fluid Mech. **25**, 241–270 (1966)
5. Benjamin, T.B.: Internal waves of permanent form in fluids of great depth. J. Fluid Mech. **29**, 559–592 (1967)
6. Bona, J.L., Lannes, D., Saut, J.-C.: Asymptotic models for internal waves. J. Math. Pures Appl. **89**, 538–566 (2008)
7. Bresch, D., Klein, R.: Spectral analysis of the compressible Euler equations for atmospheric mescales, in preparation
8. Bresch, D., Renardy, M.: Well-posedness of two-layer shallow-water flow between two horizontal rigid plates. Nonlinearity **24**, 1081–1088 (2011)
9. Brown, D.J., Christie, D.R.: Fully nonlinear internal waves in continuously stratified Boussineq fluids. Phys. Fluids **10**, 2569–2586 (1998)
10. Castro, A., Lannes, D.: Well-posedness and shallow-water stability for a new Hamiltonian formulation of the water waves equations with vorticity. Indiana Univ. Math. J. **64**, 1169–1270 (2015)
11. Chelton, D.B., DeSzoeke, R.A., Schlax, M.G., El Naggar, K., Siwertz, N.: Geographical variability of the first baroclinic Rossby radius of deformation. J. Phys. Oceanogr. **28**, 433–460 (1998)
12. Choi, W., Camassa, R.: Long internal waves of finite amplitude. Phys. Rev. Lett. **77**, 1759–1762 (1996)
13. Craig, W., Guyenne, P., Kalisch, H.: Hamiltonian long-wave expansions for free surfaces and interfaces. Commun. Pure. Appl. Math. **58**, 1587–1641 (2005)
14. Danchin, R.: On the well-posedness of the incompressible density-dependent Euler equations in the L^p framework. J. Differ. Equ. **248**, 2130–2170 (2010)
15. Davis, R.E., Acrivos, A.: Solitary waves in deep water. J. Fluid Mech. **29**, 593–601 (1967)
16. Dubreil-Jacotin, M.L.: Sur les théorèmes d'existence relatifs aux ondes périodiques dans les liquides hétérogènes. J. Math. Pures Appl. **16**, 43- (1937)
17. Duchêne, V., Israwi, S., Talhouk, R.: A new class of two-layer Green–Naghdi systems with improved frequency dispersion. Stud. Appl. Math. **137** (2016)
18. Ebin, G.: Ill-posedness of the Rayleigh–Taylor and Kelvin–Helmotz problems for incompressible fluids. Commun. Partial Differ. Equ. **13**, 1265–1295 (1988)
19. Feliks, Y., Ghil, M., Simonnet, E.: Low-frequency variability in the midlatitude atmosphere induced by an oceanic thermal front. J. Atmos. Sci. **61**, 961–981 (2004)
20. Flierl, G.R.: Models of vertical structure and the calibration of two-layer models. Dyn. Atmos. Oceans **2**, 341–381 (1978)
21. Fulton, C.T., Pruess, S.A.: Eigenvalue and eigenfunction asymptotics for regular Sturm–Liouville problems. J. Math. Anal. Appl. **188**, 297–340 (1994)
22. Gill, A.E.: Atmosphere-Ocean Dynamics, International Geophysics Series, vol. 30. Academic Press, London (1982)
23. Griffiths, S.D., Grimshaw, R.H.: Internal tide generation at the continental shelf modeled using a modal decomposition: two-dimensional results. J. Phys. Oceanogr. **37**, 428–451 (2007)
24. Guyenne, P., Lannes, D., Saut, J.-C.: Well-posedness of the Cauchy problem for models of large amplitude internal waves. Nonlinearity **23**, 237 (2010)
25. Helfrich, K.R., Melville, W.K.: Long nonlinear internal waves. Annu. Rev. Fluid Mech. **38**, 395–425 (2006)
26. Iguchi, T., Tanaka, N., Tani, A.: On the two-phase free boundary problem for two- dimensional water waves. Math. Ann. **309**, 199–223 (1997)
27. Itoh, S.: Cauchy problem for the Euler equations of a nonhomogeneous ideal incompressible fluid. II. J. Korean Math. Soc. **32**, 41–50 (1995)
28. Itoh, S., Tani, A.: Solvability of nonstationary problems for nonhomogeneous incompressible fluids and the convergence with vanishing viscosity. Tokyo J. Math. **22**, 17–42 (1999)
29. James, G.: Internal travelling waves in the limit of a discontinuously stratified fluid. Arch. Ration. Mech. Anal. **160**, 41–90 (2001)
30. Kamotski, V., Lebeau, G.: On 2D Rayleigh–Taylor instabilities. Asymptot. Anal. **42**, 1–27 (2005)
31. Klein, R., Achatz, U., Bresch, D., et al.: Regime of validity of soundproof atmospheric flow models. J. Atmos. Sci. **67**, 3226–3237 (2010)
32. Kubota, T., Ko, D.R.S., Dobbs, L.D.: Weakly nonlinear, long internal gravity waves in stratified fluids of finite depth. J. Hydronautics **12**, 157–165 (1978)
33. Lannes, D.: Dispersive effects for nonlinear geometrical optics with rectification. Asympt. Anal. **18**, 111–146 (1998)

34. Lannes, D.: Sharp estimates for pseudo-differential operators with symbols of limited smoothness and commutators. J. Funct. Anal. **232**, 495–539 (2006)
35. Lannes, D.: The Water Waves Problem: Mathematical Analysis and Asymptotics, Mathematical Surveys and Monographs, vol. 188. AMS, Providence (2013)
36. Lannes, D.: A stability criterion for two-fluid interfaces and applications. Arch. Ration. Mech. Anal. **208**, 481–567 (2013)
37. Lannes, D.: Modeling shallow water waves. Nonlinearity **33**(5), R1 (2020)
38. Lannes, D., Ming, M.: The Kelvin-Helmholtz instabilities in two-fluids shallow water models. In: Hamiltonian Partial Differential Equations and Applications, Fields Institute Communications, vol. 75. Springer, New York (2015)
39. Long, R.R.: Some aspects of the flow of stratified fluids.I. A theoretical investigation. Tellus **5**, 42- (1953)
40. Long, R.R.: On the Boussinesq approximation and its role in the theory of internal waves. Tellus **17**, 46 (1965)
41. Martin, S., Simmons, W., Wunsch, C.: The excitation of resonant triads by single internal waves. J. Fluid Mech. **53**, 17–44 (1972)
42. Maslowe, S., Redekopp, L.: Long nonlinear waves in stratified shear flows. J. Fluid Mech. **101**, 321–348 (1980)
43. Ono, H.: Algebraic solitary waves in stratified fluids. J. Phys. Soc. Jpn. **39**, 1082–1091 (1975)
44. Saut, J.-C.: Asymptotic Models for Surface and Internal Waves. IMPA, London (2013)
45. Taylor, M.E.: Partial differential equations. III, Applied Mathematical Sciences, vol. 117. Springer, New York (1997). Nonlinear equations, Corrected reprint of the 1996 original
46. Yih, C.S.: Gravity waves in a stratified fluid. J. Fluid Mech. **8**, 48- (1960)
47. Yih, C.S.: Exact solutions for steady two-dimensional flow of a stratified fluid. J. Fluid Mech. **9**, 161- (1960)