IOP PUBLISHING NONLINEARITY

Nonlinearity **23** (2010) 237–275 doi:10.1088/0951-7715/23/2/003

Well-posedness of the Cauchy problem for models of large amplitude internal waves

Philippe Guyenne[1], David Lannes[2] and Jean-Claude Saut[3]

[1] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716-2553, USA
[2] Ecole Normale Supérieure, DMA et CNRS UMR 8553, 45 Rue d'Ulm, 75005 Paris, France
[3] Laboratoire de Mathématiques, UMR 8628, Université Paris-Sud et CNRS, 91405 Orsay, France

E-mail: guyenne@math.udel.edu, David.Lannes@ens.fr and jean-claude.saut@math.u-psud.fr

Received 14 July 2009
Published 11 January 2010
Online at stacks.iop.org/Non/23/237

Recommended by L Bunimovich

Abstract

We consider in this paper the 'shallow-water/shallow-water' asymptotic model obtained in Choi and Camassa (1999 *J. Fluid Mech.* **396** 1–36), Craig *et al* (2005 *Commun. Pure. Appl. Math.* **58** 1587–641) (one-dimensional interface) and Bona *et al* (2008 *J. Math. Pures Appl.* **89** 538–66) (two-dimensional interface) from the two-layer system with rigid lid, for the description of large amplitude internal waves at the interface of two layers of immiscible fluids of different densities. For one-dimensional interfaces, this system is of hyperbolic type and its local well-posedness does not raise serious difficulties, although other issues (blow-up, loss of hyperbolicity, etc) turn out to be delicate. For two-dimensional interfaces, the system is nonlocal. Nevertheless, we prove that it conserves some properties of 'hyperbolic type' and show that the associated Cauchy problem is locally well posed in suitable Sobolev classes provided some natural restrictions are imposed on the data. These results are illustrated by numerical simulations with emphasis on the formation of shock waves.

Mathematics Subject Classification: 35L45, 76B55

1. Introduction

1.1. General setting

In [5] Bona, Lannes and Saut derived in a systematic way, and for a large class of scaling regimes, asymptotic models for the propagation of internal waves at the interface between two layers of immiscible fluids of different densities, under the rigid lid assumption and with a flat bottom. More precisely, they considered a homogeneous fluid of depth d_1 and density ρ_1 lying over another homogeneous fluid of depth d_2 and density $\rho_2 > \rho_1$. The bottom on which both fluids rest was assumed to be horizontal and featureless while the top of fluid 1 was restricted by the rigid lid assumption, which is to say, the top boundary was viewed as an impenetrable, bounding surface.

The full (Euler) model for this situation was reduced to a system of evolution equations posed spatially on $\mathbb{R}^d$, $d = 1, 2$, that involve two nonlocal operators, a classical Dirichlet to Neumann operator and an interface operator that links the velocity potentials of the two layers.

The different asymptotic models were obtained by expanding the nonlocal operators with respect to suitable small parameters that depend variously on the amplitude, wavelengths and depth ratio of the two layers. Furthermore, the consistency of the full equations with the asymptotic models was rigorously established.

Denoting by a a typical amplitude of the deformation of the interface and λ a typical wavelength, we introduced the dimensionless parameters

$$\gamma := \frac{\rho_1}{\rho_2} \in [0, 1], \qquad \delta := \frac{d_1}{d_2}, \qquad \varepsilon := \frac{a}{d_1}, \qquad \mu := \frac{d_1^2}{\lambda^2}.$$

It also turned out to be useful to introduce two other parameters ε_2 and μ_2 defined as

$$\varepsilon_2 = \frac{a}{d_2} = \varepsilon\delta, \qquad \mu_2 = \frac{d_2^2}{\lambda^2} = \frac{\mu}{\delta^2}.$$

Remark 1. The parameters ε_2 and μ_2 correspond to ε and μ with d_2 rather than d_1 taken as the unit of length in the vertical direction.

The full internal wave equations (or two-layer free interface Euler system) can be written in dimensionless form involving γ, δ, ϵ and μ. An essential step in the analysis is that the internal wave system can be reduced to a system of two evolution equations coupling the free interface ζ to the velocity potential in the upper fluid evaluated at the interface, namely ψ_1. Such a system is a generalization of the Zakharov, Craig–Sulem–Sulem formulation of the classical water waves problem [13, 29] and is written in dimensionless form:

$$\begin{cases} \partial_t \zeta - \dfrac{1}{\mu} G^\mu[\varepsilon\zeta]\psi_1 = 0, \\ \partial_t (\boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]\psi_1 - \gamma\nabla\psi_1) + (1-\gamma)\nabla\zeta \\ \quad + \dfrac{\varepsilon}{2}\nabla(|\boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]\psi_1|^2 - \gamma|\nabla\psi_1|^2) + \varepsilon\nabla\mathcal{N}^{\mu,\delta}(\varepsilon\zeta, \psi_1) = 0, \end{cases} \tag{1}$$

where $\mathcal{N}^{\mu,\delta}$ is defined for all pairs (ζ, ψ) smooth enough by the formula

$$\mathcal{N}^{\mu,\delta}(\zeta, \psi) := \mu \frac{\gamma\left(\dfrac{1}{\mu}G^\mu[\zeta]\psi + \nabla\zeta\cdot\nabla\psi\right)^2 - \left(\dfrac{1}{\mu}G^\mu[\zeta]\psi + \nabla\zeta\cdot\boldsymbol{H}^{\mu,\delta}[\zeta]\psi\right)^2}{2(1+\mu|\nabla\zeta|^2)}.$$

Here $G^\mu[\varepsilon\zeta]$ is a Dirichlet–Neumann operator for the upper fluid, while $\boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]$ is the (nonlocal) interface operator linking the trace on the interface of the velocity potential for the upper fluid to the trace on the interface of the velocity potential for the lower fluid.

More precisely, given the trace ψ_1 of the velocity potential for the upper fluid, let the function Φ_2 be the unique solution (up to a constant) of the boundary-value problem

$$\begin{cases} \mu\Delta\Phi_2 + \partial_z^2\Phi_2 = 0 & -1 - 1/\delta < z < -1 + \varepsilon\zeta, \\ \partial_z\Phi_2\,_{|z=-1-1/\delta} = 0, & \partial_n\Phi_2\,_{|z=-1+\varepsilon\zeta} = \dfrac{1}{(1+\varepsilon^2|\nabla\zeta|^2)^{1/2}} G^\mu[\varepsilon\zeta]\psi_1. \end{cases} \tag{2}$$

Then, the operator $\boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]\cdot$ is defined on ψ_1 by

$$\boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]\psi_1 = \nabla(\Phi_2\,_{|z=-1+\varepsilon\zeta}).$$

Remark 2. In the above formulation, $\partial_n \Phi_{2\,|_{z=-1+\varepsilon\zeta}}$ stands here for the upward conormal derivative associated with the elliptic operator $\mu\Delta\Phi_2 + \partial_z^2\Phi_2$,

$$\sqrt{1+\varepsilon^2|\nabla\zeta|^2}\partial_n\Phi_{2\,|_{z=-1+\varepsilon\zeta}} = -\mu\varepsilon\nabla\zeta\cdot\nabla\Phi_{2\,|_{z=-1+\varepsilon\zeta}} + \partial_z\Phi_{2\,|_{z=-1+\varepsilon\zeta}}.$$

The Neumann boundary condition of (2) at the interface can also be stated as $\partial_n\Phi_{2\,|_{z=-1+\varepsilon\zeta}} = \partial_n\Phi_{1\,|_{z=-1+\varepsilon\zeta}}$.

The asymptotic models derived in [5] couple the elevation ζ of the interface with a variable $\boldsymbol{v}$ defined by

$$\boldsymbol{v} := \boldsymbol{H}^{\mu,\delta}[\varepsilon\zeta]\psi_1 - \gamma\nabla\psi_1, \tag{3}$$

which is the gradient of the second canonical variable in the Hamiltonian formulation of (1) (see for instance [4, 11, 12]).

Remark 3. Linearizing equations (1) around the rest state $\tilde{\zeta} = 0$, $\tilde{\psi}_1 = 0$ one finds the equations

$$\begin{cases} \partial_t\zeta - \dfrac{1}{\mu}G^\mu[0]\psi_1 = 0, \\ \partial_t(\boldsymbol{H}^{\mu,\delta}[0]\psi_1 - \gamma\nabla\psi_1) + (1-\gamma)\nabla\zeta = 0. \end{cases}$$

Since $G^\mu[0]$ and $\boldsymbol{H}^{\mu,\delta}[0]$ can be explicitly computed by Fourier analysis, this leads to the linearized dispersion relation

$$\omega^2 = (1-\gamma)\frac{|\boldsymbol{k}|}{\sqrt{\mu}}\frac{\tanh(\sqrt{\mu}|\boldsymbol{k}|)\tanh\left(\dfrac{\sqrt{\mu}}{\delta}|\boldsymbol{k}|\right)}{\tanh(\sqrt{\mu}|\boldsymbol{k}|)+\gamma\tanh\left(\dfrac{\sqrt{\mu}}{\delta}|\boldsymbol{k}|\right)}; \tag{4}$$

corresponding to plane-wave solutions $e^{i\boldsymbol{k}\cdot X - i\omega t}$.

This paper is devoted to the study of the asymptotic model obtained in the shallow water/shallow water (SW/SW) regime, that is to say we assume that $\mu \ll 1$, $\mu_2 \ll 1$. In this regime, large amplitude waves are allowed with respect to both the upper fluid ($\varepsilon = O(1)$) and the lower fluid ($\varepsilon_2 = O(1)$). For the sake of notational simplicity, we take $\varepsilon = 1$ throughout this paper (and thus $\varepsilon_2 = \delta$). It has been proved in [5] that the internal wave equations (1) are consistent with this asymptotic model in both dimensions 1 and 2.

In the one-dimensional case $d = 1$, the model can be written in the form (see also the more compact formulation (15)):

$$\begin{cases} \partial_t\zeta + \partial_x\left(\dfrac{h_1h_2}{\delta h_1+\gamma h_2}\boldsymbol{v}\right) = 0, \\ \partial_t\boldsymbol{v} + (1-\gamma)\partial_x\zeta + \dfrac{1}{2}\partial_x\left(\dfrac{(\delta h_1)^2-\gamma h_2^2}{(\delta h_1+\gamma h_2)^2}\boldsymbol{v}^2\right) = 0, \end{cases} \tag{5}$$

where $h_1 = 1-\zeta$ and $h_2 = 1+\delta\zeta$. This system has been formally derived in [12] (see system (5.26)) and in a slightly different form in [9] but to our knowledge it has not been analysed as a hyperbolic system.

The two-dimensional ($d = 2$) generalization of (5) has been derived for the first time in [5] and exhibits some new nonlocal terms. More precisely, this two-dimensional generalization of (5) can be written as

$$\begin{cases} \partial_t\zeta + \nabla\cdot(h_1\mathfrak{R}[\zeta]\boldsymbol{v}) = 0, \\ \partial_t\boldsymbol{v} + (1-\gamma)\nabla\zeta + \frac{1}{2}\nabla(|\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2) = 0, \end{cases} \tag{6}$$

where $h_1 = 1-\zeta$, $h_2 = 1+\delta\zeta$, and the operator $\mathfrak{R}[\zeta]$, which contains the nonlocal effects, is defined as below.

Definition 1. *Let $\gamma \in [0,1)$, $\delta > 0$ and $\zeta \in L^\infty(\mathbb{R}^d)$ be such that*

$$1 - |\zeta|_\infty > 0, \qquad 1 - \delta|\zeta|_\infty > 0.$$

The operator $\mathfrak{R}[\zeta]$ is then defined as

$$\mathfrak{R}[\zeta]: \begin{array}{l} L^2(\mathbb{R}^d)^d \to L^2(\mathbb{R}^d)^d \\ \boldsymbol{u} \mapsto \mathfrak{R}[\zeta]\boldsymbol{u} := \dfrac{1}{\gamma+\delta}\left(1 - \Pi\left(\dfrac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\,\cdot\right)\right)^{-1}\Pi(h_2\boldsymbol{u}), \end{array}$$

where $h_2 = 1 + \delta\zeta$, and $\Pi := \nabla\nabla^\top/\Delta$ denotes the projection onto gradient vector fields.

Remark 4. The assumptions on ζ ensure that $|\frac{1-\gamma}{\gamma+\delta}\delta\zeta|_\infty < 1$ and thus allow one to define $(1 - \Pi(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot))^{-1}$ in $Ł^2(\mathbb{R}^d)$ by its Neumann series:

$$\mathfrak{R}[\zeta]\boldsymbol{u} = \frac{1}{\gamma+\delta}\sum_{n=0}^{\infty}\left(\Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\,\cdot\right)\right)^n\Pi(h_2\boldsymbol{u}). \tag{7}$$

Note also that when $d = 1$, one has $\Pi = 1$ and

$$\mathfrak{R}[\zeta]\boldsymbol{u} = \frac{h_2}{\delta h_1 + \gamma h_2}\boldsymbol{u},$$

so that (6) and (5) coincide as expected.

Remark 5. The notations in (6) differ slightly from the notations used in [5] where the equations are written in terms of an operator $\mathfrak{Q}[\frac{\gamma-1}{\gamma+\delta}\delta\zeta]$ (see lemma 3 of this reference) rather than $\mathfrak{R}[\zeta]$. It is straightforward to remark that

$$\mathfrak{R}[\zeta]\boldsymbol{u} = \frac{1}{\gamma+\delta}\mathfrak{Q}\left[\frac{\gamma-1}{\gamma+\delta}\delta\zeta\right](h_2\boldsymbol{u})$$

and that the two formulations coincide.

Remark 6. Since $\nabla\cdot(h_1\mathfrak{R}[\zeta]\boldsymbol{v}) = \nabla\cdot(h_2\boldsymbol{v})$ when $\gamma = 0$ and $\delta = 1$ (see lemma 3 of [5]), (6) degenerates into the two-dimensional nonlinear shallow water equations that do not involve nonlocal terms. These nonlocal terms are therefore specific to *two-dimensional* internal waves. Moreover, Duchêne [15] has recently proved that they are a consequence of the rigid lid assumption.

As kindly pointed out to us by an anonymous referee, the appearance of nonlocal operators resulting from the two-dimensional approximation in shallow water type models was also noticed in [7] in the context of surface waves over a variable bottom.

Remark 7. Performing the approximation one order further in μ would lead to 'Green–Naghdi type' systems, involving nonlinear dispersion (see [9, 15], and [12, 17] in the one-dimensional case). The corresponding well-posedness theory has not been carried out so far: for the Green–Naghdi equations derived in the case of one single fluid, the well-posedness has been established in [20, 23] in the one-dimensional case and [2] in two dimensions.

1.2. Organization of the paper

The paper is organized as follows. In section 2 we present some properties of the nonlocal operator $\mathfrak{R}[\zeta]$ which will be used in what follows in the paper.

Section 3 is devoted to the local well-posedness for the Cauchy problem associated with the SW/SW systems (5) and (6). We first put them in a 'quasilinear system' formulation. While this is straightforward in the one-dimensional case, this is more delicate in the two-dimensional case because of the presence of the nonlocal operator $\mathfrak{R}[\zeta]$. In particular, the 'quasilinear' formulation involves now 0-order nonlocal operators and the equivalence between the two formulations is only established for curl free $\boldsymbol{v}$s.

We next prove the local well-posedness of the transformed systems under suitable conditions on the initial data. These conditions ensure that the system is strictly hyperbolic in dimension 1. Two conditions amount to saying that both fluid layers remain of positive depth. Such a condition has to be imposed in the classical Saint-Venant (shallow water) system for surface waves. The third condition (a smallness condition on $\boldsymbol{v}$ and ζ) has no counterpart for the Saint-Venant system. It can be viewed as a 'trace' of the Kelvin–Helmoltz instabilities inherent to the full system (1) (see [8]). The existence proof for the transformed system is obtained via an energy method implemented on a regularized version of the system. The one-dimensional proof is standard but we nevertheless give it to emphasize the differences with the two-dimensional case. Moreover, we provide precise blow-up conditions and prove that blow-up occurs for 'many' initial data. This is due to the fact that the two characteristic fields can be linearly degenerate only on a 'small' closed subset of the domain of hyperbolicity.

The existence proof for the two-dimensional case follows the strategy of that of the one-dimensional case but is much more delicate. In particular, serious difficulties arise from the nonlocal terms in the construction of a symmetrizer.

To complete the existence proof we need to show that a solution of the transformed system yields a solution of (5) or (6). This is straightforward in dimension 1 while in dimension 2 this amounts to proving that $\boldsymbol{v}(., t)$ remains curl free provided it is so at time $t = 0$, a fact that is established in proposition 6.

Finally, we present some numerical simulations of the SW/SW systems in the last section 4 under periodic boundary conditions. We use a pseudospectral method for the spatial discretization which is well suited for the approximation of the nonlocal operator $\mathfrak{R}[\zeta]$. The numerical results illustrate the various theoretical results related to blow-up in the one-dimensional case and suggest interesting conjectures in dimension 2.

1.3. Notations

- We denote by $C(p_1, \dots, p_n)$ any positive function nondecreasing with respect to the parameters $p_1, \dots p_n$.
- We denote the horizontal variables by x when $d = 1$ and by $X = (x, y)^\top$ when $d = 2$.
- The notation $A \lesssim B$ means that $A \leqslant CB$, for some positive constant C whose exact value has no importance.
- Partial differentiation with respect to x (respectively y) is denoted indifferently by ∂_x or ∂_1 (respectively ∂_y or ∂_2).
- We denote, respectively, by $\boldsymbol{e}^1$ and $\boldsymbol{e}^2$ the unit vectors (of $\mathbb{R}^2$) $(1, 0)^\top$ and $(0, 1)^\top$ and if $\boldsymbol{v} \in \mathbb{R}^d$, we denote by $\boldsymbol{v}_j$ its jth component.
- We use the convention of summation on repeated indices, i.e. we simply write $a_j b_j$ instead of $\sum_{j=1}^2 a_j b_j$.
- We denote by $|\cdot|_p$ $(1 \leqslant p \leqslant \infty)$ the standard norm of the Lebesgue spaces $L^p(\mathbb{R}^d)$ $(d = 1, 2)$.
- If $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2)^\top \in L^2(\mathbb{R}^2)^2$, then we write $|\boldsymbol{v}|_2 = (|\boldsymbol{v}_1|_2^2 + |\boldsymbol{v}_2|_2^2)^{1/2}$.
- If $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2)^\top \in L^\infty(\mathbb{R}^2)^2$, then we write $|\boldsymbol{v}|_\infty = |\boldsymbol{v}_1|_\infty + |\boldsymbol{v}_2|_\infty$.
- For a two-dimensional vector field $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2)^\top$, we define curl $\boldsymbol{v} = \partial_1 \boldsymbol{v}_2 - \partial_2 \boldsymbol{v}_1$.
- We use the Fourier multiplier notation: $f(D)u$ is defined as $\mathcal{F}(f(D)u)(\xi) = f(\xi)\hat{u}(\xi)$, where $\mathcal{F}$ and $\hat{\cdot}$ stand for the Fourier transform.

– As mentioned above, the projection onto gradient fields in $L^2(\mathbb{R}^d)^d$ is written Π and is defined by the formula

$$\Pi = -\frac{\nabla\nabla^\top}{|D|^2}.$$

(Note that $\Pi = \mathrm{Id}$ when $d = 1$.)

– The operator $\Lambda = (1-\Delta)^{1/2}$ is equivalently defined using the Fourier multiplier notation to be $\Lambda = (1+|D|^2)^{1/2}$.

– The standard notation $H^s(\mathbb{R}^d)$, or simply H^s if the underlying domain is clear from the context, is used for the L^2-based Sobolev spaces; their norm is written as $|\cdot|_{H^s}$.

2. Preliminary results on the operator $\mathfrak{R}[\zeta]$

We choose to state here some properties of the operator $\mathfrak{R}[\zeta]$ that will be used in the analysis of the two-dimensional SW/SW equations (6). The study of the systems will be performed in the next section (the reader only interested in a rough overview of the properties of the SW/SW systems can therefore skip this section). The first property deals with the operator norm of the nonlocal operator.

Proposition 1. *Let $\gamma \in [0,1)$, $\delta > 0$ and $t_0 > 1$. Assume also that $\zeta \in L^\infty(\mathbb{R}^2)$ and satisfies*

$$(1 - |\zeta|_\infty) > 0 \qquad \text{and} \qquad (1 - \delta|\zeta|_\infty) > 0.$$

(1) The operator $\mathfrak{R}[\zeta] : L^2(\mathbb{R}^2)^2 \to L^2(\mathbb{R}^2)^2$ is well defined (see definition 1) and (with $h_2 = 1 + \delta\zeta$)

$$\forall \boldsymbol{v} \in L^2(\mathbb{R}^2)^2, \qquad |\mathfrak{R}[\zeta]\boldsymbol{v}|_2 \leqslant \frac{1}{\gamma + \delta - \delta(1-\gamma)|\zeta|_\infty}|h_2\boldsymbol{v}|_2.$$

(2) If moreover $\zeta \in H^s \cap H^{t_0+1}(\mathbb{R}^2)$ $(s \geqslant 0)$ then for all $\boldsymbol{v} \in H^s(\mathbb{R}^2)^2$,

$$|\mathfrak{R}[\zeta]\boldsymbol{v}|_{H^s} \leqslant C\left(\frac{1}{\gamma + \delta - \delta(1-\gamma)|\zeta|_\infty}, \delta(1-\gamma)|\zeta|_{H^{t_0+1}}\right) \times (|h_2\boldsymbol{v}|_{H^s} + \delta(1-\gamma)|\zeta|_{H^s}|\Pi(h_2\boldsymbol{v})|_{H^{t_0}}).$$

Proof. Since $\|\Pi(\frac{\gamma-1}{\gamma+\delta}\delta\zeta\Pi\cdot)\|_{L^2\to L^2} \leqslant \frac{1-\gamma}{\gamma+\delta}\delta|\zeta|_\infty$, the bilinear form $a(\boldsymbol{u}, \boldsymbol{v})$ defined as

$$a(\boldsymbol{u}, \boldsymbol{v}) = \left(\left(1 - \Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\boldsymbol{u}, \boldsymbol{v}\right)$$

is coercive and continuous on $L^2(\mathbb{R}^d)^2$, with coercivity and continuity constants, respectively, given by

$$k(\zeta) = 1 - \frac{1-\gamma}{\gamma+\delta}\delta|\zeta|_\infty \qquad \text{and} \qquad M(\zeta) = 1 + \frac{1-\gamma}{\gamma+\delta}\delta|\zeta|_\infty.$$

It follows therefore from Lax–Milgram's theorem that for all $f \in L^2(\mathbb{R}^d)^2$, there exists a unique solution to the equation

$$\left(1 - \Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\boldsymbol{u} = f,$$

and that $|\boldsymbol{u}|_2 \leqslant k(\zeta)^{-1}|f|_2$. The result is thus proved for the particular case $s = 0$.

If $s \geqslant 0$, one applies Λ^s to the equation, and obtains

$$\left(1 - \Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\Lambda^s\boldsymbol{u} = \Lambda^s f + \delta\frac{1-\gamma}{\gamma+\delta}\Pi[\Lambda^s, \zeta]\boldsymbol{u}.$$

From the estimate established above for $s = 0$, we deduce that

$$|\boldsymbol{u}|_{H^s} \leqslant \frac{1}{k(\zeta)}\left(\left|f|_{H^s} + \delta\frac{1-\gamma}{\gamma+\delta}|[\Lambda^s, \zeta]\boldsymbol{u}\right|_2\right). \tag{8}$$

If $0 \leqslant s \leqslant t_0 + 1$, we use the Calderón–Coifman–Meyer commutator estimate

$$|[\Lambda^s, g]h|_2 \lesssim |g|_{H^{t_0+1}}|h|_{H^{s-1}}$$

to get

$$|\boldsymbol{u}|_{H^s} \lesssim \frac{1}{k(\zeta)}\left(|f|_{H^s} + \delta\frac{1-\gamma}{\gamma+\delta}|\zeta|_{H^{t_0+1}}|\boldsymbol{u}|_{H^{s-1}}\right);$$

we thus obtain by continuous induction that

$$\forall 0 \leqslant s \leqslant t_0 + 1, \qquad \frac{1}{\gamma+\delta}|\boldsymbol{u}|_{H^s} \leqslant C\left(\frac{1}{\gamma+\delta-\delta(1-\gamma)|\zeta|_\infty}, \delta(1-\gamma)|\zeta|_{H^{t_0+1}}\right)|f|_{H^s}.$$

When $s > t_0 + 1$, we use (8) and the following Kato–Ponce commutator estimate

$$|[\Lambda^s, g]h|_2 \lesssim |g|_{H^{t_0+1}}|h|_{H^{s-1}} + |g|_{H^s}|h|_{H^{t_0}}$$

to obtain

$$|\boldsymbol{u}|_{H^s} \lesssim \frac{1}{k(\zeta)}\left(|f|_{H^s} + \delta\frac{1-\gamma}{\gamma+\delta}|\zeta|_{H^{t_0+1}}|\boldsymbol{u}|_{H^{s-1}} + \delta\frac{1-\gamma}{\gamma+\delta}|\zeta|_{H^s}|\boldsymbol{u}|_{H^{t_0}}\right).$$

Using the estimate for $s = t_0$ previously established and using a continuous induction, we deduce that

$$\frac{1}{\gamma+\delta}|\boldsymbol{u}|_{H^s} \leqslant C\left(\frac{1}{\gamma+\delta-\delta(1-\gamma)|\zeta|_\infty}, \delta(1-\gamma)|\zeta|_{H^{t_0+1}}\right)\left(|f|_{H^s} + \delta(1-\gamma)|\zeta|_{H^s}|f|_{H^{t_0}}\right).$$

The result is then a direct consequence of the definition of $\mathfrak{R}[\zeta]$. □

In the following proposition, we show how the divergence and partial differentiation operators act on the operator $\mathfrak{R}[\zeta]$. Let us introduce first the following notation:

$$\mathfrak{S}[\zeta]\boldsymbol{v} = \boldsymbol{v} + (1-\gamma)\mathfrak{R}[\zeta]\boldsymbol{v} \tag{9}$$

(so that $\mathfrak{S}[\zeta]\boldsymbol{v}$ degenerates into $\mathfrak{S}[\zeta]\boldsymbol{v} = \dfrac{1+\delta}{\delta h_1 + \gamma h_2}\boldsymbol{v}$ when $d = 1$).

Proposition 2. *Let $\gamma \in [0, 1)$, $\delta > 0$ and $t_0 > 1$. Assume also that $\zeta \in H^s(\mathbb{R}^2)$, with $s \geqslant t_0 + 1$, and satisfies*

$$\inf_{\mathbb{R}}(1 - |\zeta|_\infty) > 0 \qquad \text{and} \qquad \inf_{\mathbb{R}}(1 - \delta|\zeta|_\infty) > 0.$$

Then, for all $\boldsymbol{v} \in L^2(\mathbb{R}^2)^2$, one has

$$\nabla\cdot\mathfrak{R}[\zeta]\boldsymbol{v} = \delta\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{\delta h_1 + \gamma h_2}\cdot\nabla\zeta + \frac{h_2}{\delta h_1 + \gamma h_2}\nabla\cdot\boldsymbol{v}$$

and, for $j = 1, 2$,

$$\partial_j(\mathfrak{R}[\zeta]\boldsymbol{v}) = \delta\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\partial_j\zeta\right) + \mathfrak{R}[\zeta]\partial_j\boldsymbol{v}.$$

Proof. For the first identity, first note that from lemma 3 of [5] and remark 4 one has

$$\nabla\cdot((\delta h_1+\gamma h_2)\mathfrak{R}[\zeta]\boldsymbol{v}) = \nabla\cdot(h_2\boldsymbol{v});$$

since the assumptions on ζ imply $\delta h_1+\gamma h_2>0$, it follows that

$$\nabla\cdot\mathfrak{R}[\zeta]\boldsymbol{v} = \frac{1}{\delta h_1+\gamma h_2}\nabla\cdot(h_2\boldsymbol{v}) + \delta(1-\gamma)\frac{\nabla\zeta}{\delta h_1+\gamma h_2}\cdot\mathfrak{R}[\zeta]\boldsymbol{v},$$

from which the result follows easily.

For the second identity, note that by definition of $\mathfrak{R}[\zeta]\boldsymbol{v}$, one has

$$(\gamma+\delta)\left(1+\Pi\left(\frac{\gamma-1}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\mathfrak{R}[\zeta]\boldsymbol{v} = \Pi(h_2\boldsymbol{v});$$

differentiating this identity, one gets

$$(\gamma+\delta)\left(1+\Pi\left(\frac{\gamma-1}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\partial_j(\mathfrak{R}[\zeta]\boldsymbol{v}) = \partial_j\Pi(h_2\boldsymbol{v}) + (1-\gamma)\Pi(\delta\partial_j\zeta\mathfrak{R}[\zeta]\boldsymbol{v}),$$

and thus

$$\partial_j(\mathfrak{R}[\zeta]\boldsymbol{v}) = \mathfrak{R}[\zeta]\left(\frac{\partial_j(h_2\boldsymbol{v})}{h_2}\right) + \delta(1-\gamma)\mathfrak{R}[\zeta]\left(\frac{\partial_j\zeta\mathfrak{R}[\zeta]\boldsymbol{v}}{h_2}\right),$$

from which the result follows. □

As said in remark 4, one has $\mathfrak{R}[\zeta](\frac{\boldsymbol{v}}{h_2}) = \frac{1}{\delta h_1+\gamma h_2}\boldsymbol{v}$ in the one-dimensional case $d=1$; when $d=2$, this identity is of course false but the following proposition establishes that when $\boldsymbol{v}$ is a gradient vector field (i.e. when $\Pi\boldsymbol{v}=\boldsymbol{v}$) then this identity is true up to a more regular term.

Proposition 3. *Let* $\gamma\in[0,1)$, $\delta>0$ *and* $t_0>1$. *Let also* $\zeta\in H^{t_0+1}(\mathbb{R}^2)$ *be such that*

$$\inf_{\mathbb{R}}(1-|\zeta|_\infty)>0 \qquad \text{and} \qquad \inf_{\mathbb{R}}(1-\delta|\zeta|_\infty)>0.$$

Then, for all $\boldsymbol{v}\in L^2(\mathbb{R}^2)^2$,

$$\left|\mathfrak{R}[\zeta]\left(\frac{\boldsymbol{v}}{h_2}\right) - \frac{1}{\delta h_1+\gamma h_2}\Pi\boldsymbol{v}\right|_2 \leqslant C\left(\frac{1}{\gamma+\delta-\delta(1-\gamma)|\zeta|_\infty}, \delta(1-\gamma)|\zeta|_{H^{t_0+1}}\right)|\Pi\boldsymbol{v}|_{H^{-1}}.$$

Proof. Note first that one can write

$$\mathfrak{R}[\zeta]\left(\frac{\boldsymbol{v}}{h_2}\right) - \frac{1}{\delta h_1+\gamma h_2}\Pi\boldsymbol{v} = \left(1-\Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)^{-1}\boldsymbol{w},$$

with

$$\boldsymbol{w} = \frac{1}{\gamma+\delta}\Pi\boldsymbol{v} - \left(1-\Pi\left(\frac{1-\gamma}{\gamma+\delta}\delta\zeta\Pi\cdot\right)\right)\left(\frac{1}{\delta h_1+\gamma h_2}\Pi\boldsymbol{v}\right).$$

With the same notation as in the proof of proposition 1, we thus have

$$\left|\mathfrak{R}[\zeta]\left(\frac{\boldsymbol{v}}{h_2}\right) - \frac{1}{\delta h_1+\gamma h_2}\Pi\boldsymbol{v}\right|_2 \lesssim \frac{1}{k(\zeta)}|\boldsymbol{w}|_2,$$

and we are thus led to control $|\boldsymbol{w}|_2$. In order to do so, let f_1 and f_2 be defined as

$$f_1 = \frac{1-\gamma}{\gamma+\delta}\delta\zeta, \qquad f_2 = \frac{1}{\delta h_1+\gamma h_2}.$$

Simple computations show that

$$\boldsymbol{w} = f_1[\Pi, f_2]\Pi\boldsymbol{v} + [\Pi, f_1]\Pi(f_2\Pi\boldsymbol{v}).$$

Using the commutator estimate (which can be deduced from the general commutator estimates for pseudo-differential operators given in theorem 6 of [22])

$$\forall r, \qquad -t_0 < r \leqslant t_0 + 1, \qquad |[\Pi, g]h|_{H^r} \lesssim |g|_{H^{t_0+1}} |h|_{H^{r-1}} \tag{10}$$

with $r = 0$ and the product estimate (valid for $t_0 > 1 = d/2$, see for instance [16]),

$$|fg|_{H^{-1}} \lesssim |f|_{H^{t_0}} |g|_{H^{-1}} \tag{11}$$

we deduce that

$$|\boldsymbol{w}|_2 \lesssim |f_1|_{H^{t_0+1}} |f_2|_{H^{t_0+1}} |\Pi \boldsymbol{v}|_{H^{-1}},$$

and the result follows easily. □

3. Local well-posedness of the SW/SW equations

3.1. The 'quasilinear system' formulation

3.1.1. The one-dimensional case $d = 1$. When $d = 1$, it is a tedious but simple computation to check that the SW/SW equations (5) can be put in the 'quasilinear' form:

$$\partial_t U + A(U)\partial_x U = 0, \qquad U = (\zeta, \boldsymbol{v})^\top, \tag{12}$$

with

$$A(U) = \begin{pmatrix} a(U) & b(\zeta) \\ c(U) & a(U) \end{pmatrix}$$

and

$$a(U) = f'(\zeta)\boldsymbol{v}, \qquad b(\zeta) = f(\zeta), \qquad c(U) = (1-\gamma) + \tfrac{1}{2} f''(\zeta)\boldsymbol{v}^2, \tag{13}$$

and where the function $f(\zeta)$ is given by

$$f(\zeta) = \frac{h_1 h_2}{\delta h_1 + \gamma h_2}. \tag{14}$$

The following proposition is thus straightforward:

Proposition 4 (The case $d = 1$). *Let $T > 0$, $t_0 > 1/2$ and $s \geqslant t_0 + 1$. Then $U = (\zeta, \boldsymbol{v})^\top \in C([0, T]; H^s(\mathbb{R})^2)$ solves (5) if and only if it solves (12).*

Remark 8. Equation (5) can thus be recast in the conservative form

$$\begin{cases} \partial_t \zeta + \partial_x (f(\zeta)\boldsymbol{v}) = 0, \\ \partial_t \boldsymbol{v} + (1-\gamma)\partial_x \zeta + \frac{1}{2}\partial_x (f'(\zeta)\boldsymbol{v}^2) = 0. \end{cases} \tag{15}$$

Remark 9. The system studied in [26] (see also [10]) appears to be a particular case of (12) in the situation where $\gamma \sim 1$ (two fluids of almost identical densities). In order to derive the system of [26] from (12), one just has to change the velocity unknown $\boldsymbol{v}$ into $\boldsymbol{u} = \frac{\delta h_1}{\delta h_1 + \gamma h_2}\boldsymbol{v}$ ($\boldsymbol{u}$ is then an approximation of the mean horizontal velocity in the lower fluid); setting $\gamma = 1$ in the resulting equations (except in the term involving gravity—the term $(1-\gamma)$ in $c(U)$) yields equations (2.18) and (2.19) of [26].

3.1.2. The two-dimensional case $d = 2$. When $d = 2$, it is trickier to put (6) in a quasilinear form because of the presence of the nonlocal term $\mathfrak{R}[\zeta]\boldsymbol{v}$. The main result of this section is to prove that one can write (6) in the equivalent form

$$\partial_t U + A^j[U]\partial_j U = 0, \qquad U = (\zeta, \boldsymbol{v})^\top, \tag{16}$$

where

$$A^j[U] = \begin{pmatrix} a^j(U) & \boldsymbol{b}^j(U)^\top \\ \boldsymbol{c}^j[U] & D^j[U] \end{pmatrix} \qquad (j = 1, 2),$$

and

$$a^j(U) = (\boldsymbol{v} - \gamma\mathfrak{R}[\zeta]\boldsymbol{v})_j - \gamma(\mathfrak{S}[\zeta]\boldsymbol{v})_j \frac{h_2}{\delta h_1 + \gamma h_2}, \tag{17}$$

$$\boldsymbol{b}^j(U) = \frac{h_1 h_2}{\delta h_1 + \gamma h_2}\boldsymbol{e}^j, \tag{18}$$

$$\boldsymbol{c}^j[U]\bullet = \boldsymbol{e}^j - \gamma\left[\boldsymbol{e}^j + \delta(\mathfrak{S}[\zeta]\boldsymbol{v})_j\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\bullet\right)\right], \tag{19}$$

$$D^j[U]\bullet = (\boldsymbol{v} - \gamma\mathfrak{R}[\zeta]\boldsymbol{v})_j \mathrm{Id}_{2\times 2} - \gamma(\mathfrak{S}[\zeta]\boldsymbol{v})_j\mathfrak{R}[\zeta]\bullet, \tag{20}$$

the operator $\mathfrak{S}[\zeta]$ being as defined in (9).

Proposition 5 (The case $d = 2$). *Let $T > 0$, $t_0 > 1$ and $s \geqslant t_0 + 1$. Let also $U = (\zeta, \boldsymbol{v})^\top \in C([0, T]; H^s(\mathbb{R}^2)^3)$ be such that for all $t \in [0, T]$,*

$$(1 - |\zeta(t, \cdot)|_\infty) > 0 \qquad (1 - \delta|\zeta(t, \cdot)|_\infty) > 0 \qquad \text{and} \qquad \operatorname{curl} \boldsymbol{v}(t, \cdot) = 0.$$

Then, U solves (6) if and only U solves (16).

Remark 10. System (16) is not *stricto sensu* a quasilinear system since $\boldsymbol{c}^j[U]$ (respectively $D^j[U]$) is not an $\mathbb{R}^2$-vector-valued (respectively 2×2-matrix-valued) function but a linear operator defined over the space of $\mathbb{R}^2$-vector-valued (respectively 2×2-matrix-valued) functions. However, these operators are of order zero and, as shown below, (16) can be handled roughly as a quasilinear system.

Proof. One can use proposition 2 to express the quantities involved in (6) in the following form:

Lemma 1. *Let $t_0 > 1$ and $U = (\zeta, \boldsymbol{v})^\top \in H^{t_0+1}(\mathbb{R}^2)^3$ be such that*

$$\inf_{\mathbb{R}^2}(1 - |\zeta|_\infty) > 0 \qquad \text{and} \qquad \inf_{\mathbb{R}^2}(1 - \delta|\zeta|_\infty) > 0.$$

Then, one has

$$\nabla \cdot (h_1\mathfrak{R}[\zeta]\boldsymbol{v}) = (a^j(U)\partial_j\zeta + \boldsymbol{b}^j(U) \cdot \partial_j\boldsymbol{v}),$$

where $a^j(U)$ and $\boldsymbol{b}^j(U)$ are given by (17) and (18).

If moreover $\operatorname{curl} \boldsymbol{v} = 0$ *then*

$$(1 - \gamma)\nabla\zeta + \tfrac{1}{2}\nabla(|\boldsymbol{v} - \gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2) = \boldsymbol{c}^j[U]\partial_j\zeta + D^j[U]\partial_j\boldsymbol{v},$$

where the operators $\boldsymbol{c}^j[U]$ and $D^j[U]$ are given by (19) and (20).

Proof. To prove the first identity, just compute

$$\nabla\cdot(h_1\mathfrak{R}[\zeta]\boldsymbol{v}) = -\nabla\zeta\cdot\mathfrak{R}[\zeta]\boldsymbol{v} + h_1\nabla\cdot\mathfrak{R}[\zeta]\boldsymbol{v},$$

and use the first part of proposition 2 to check that

$$\nabla\cdot(h_1\mathfrak{R}[\zeta]\boldsymbol{v}) = \frac{\delta h_1\boldsymbol{v} - \gamma(1+\delta)\mathfrak{R}[\zeta]\boldsymbol{v}}{\delta h_1 + \gamma h_2}\cdot\nabla\zeta + \frac{h_1 h_2}{\delta h_1+\gamma h_2}\nabla\cdot\boldsymbol{v},$$

which implies directly the result.

In order to prove the second identity, recall first that for all $\boldsymbol{u}\in H^1(\mathbb{R}^2)^2$, one has

$$\tfrac{1}{2}\nabla|\boldsymbol{u}|^2 = (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + (\mathrm{curl}\ \boldsymbol{u})\boldsymbol{u}^\perp$$

with $\boldsymbol{u}^\perp = (\boldsymbol{u}_2, -\boldsymbol{u}_1)^\top$ and we recall that $\mathrm{curl}\ \boldsymbol{u} = \partial_1\boldsymbol{u}_2 - \partial_2\boldsymbol{u}_1$. Since $\mathrm{curl}\ \boldsymbol{v} = 0$ (by assumption) and $\mathrm{curl}\ \mathfrak{R}[\zeta]\boldsymbol{v} = 0$ (by definition of $\mathfrak{R}[\zeta]$), we have

$$\begin{aligned}\tfrac{1}{2}\nabla(|\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2) &= ((\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})\cdot\nabla)(\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})\\ &\quad-\gamma(\mathfrak{R}[\zeta]\boldsymbol{v}\cdot\nabla)\mathfrak{R}[\zeta]\boldsymbol{v},\end{aligned}$$

or equivalently

$$\begin{aligned}\tfrac{1}{2}\nabla(|\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2) &= (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} - \gamma(\mathfrak{R}[\zeta]\boldsymbol{v}\cdot\nabla)\boldsymbol{v} - \gamma(\boldsymbol{v}\cdot\nabla)\mathfrak{R}[\zeta]\boldsymbol{v}\\ &\quad-\gamma(1-\gamma)(\mathfrak{R}[\zeta]\boldsymbol{v}\cdot\nabla)\mathfrak{R}[\zeta]\boldsymbol{v}.\end{aligned}\qquad(21)$$

We can now use the second identity of proposition 2 to check that

$$(\boldsymbol{v}\cdot\nabla)\mathfrak{R}[\zeta]\boldsymbol{v} = \boldsymbol{v}_j\left(\delta\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\partial_j\zeta\right) + \mathfrak{R}[\zeta]\partial_j\boldsymbol{v}\right)$$

and

$$(\mathfrak{R}[\zeta]\boldsymbol{v}\cdot\nabla)\mathfrak{R}[\zeta]\boldsymbol{v} = (\mathfrak{R}[\zeta]\boldsymbol{v})_j\left(\delta\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\partial_j\zeta\right) + \mathfrak{R}[\zeta]\partial_j\boldsymbol{v}\right).$$

It follows therefore from (21) that

$$\begin{aligned}\frac{1}{2}\nabla(|\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2) &= -\gamma\delta(\mathfrak{S}[\zeta]\boldsymbol{v})_j\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\partial_j\zeta\right)\\ &\quad+\big((\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})\cdot\nabla\big)\boldsymbol{v} - \gamma(\mathfrak{S}[\zeta]\boldsymbol{v})_j\mathfrak{R}[\zeta]\partial_j\boldsymbol{v}.\end{aligned}$$ □

The proposition is then a direct consequence of lemma 1. □

The next proposition is crucial in order to prove that a solution of (16) which is initially curl free remains curl free and thus yields a solution of (6).

Proposition 6. *Let $T>0$, $t_0>1$ and $s\geqslant t_0+1$. Let also $U=(\zeta,\boldsymbol{v})^\top\in C([0,T];H^s(\mathbb{R}^2)^3)$ be a solution of (16) such that* $\mathrm{curl}\ \boldsymbol{v}(0,.)=0$. *Then* $\mathrm{curl}\ \boldsymbol{v}(t,.)=0$ *for all $t\in[0,T]$.*

Proof. We consider the equation for $\boldsymbol{v}$ in (16). According to the computations in the proof of lemma 1, and using that $\mathrm{curl}\ \mathfrak{R}[\zeta]\boldsymbol{v}=0$, it is written

$$\partial_t\boldsymbol{v} + \nabla F - (\mathrm{curl}\ \boldsymbol{v})(\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})^\perp = 0,\qquad(22)$$

where

$$F = (1-\gamma\zeta) + \tfrac{1}{2}(|\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v}|^2 - \gamma|\mathfrak{R}[\zeta]\boldsymbol{v}|^2).$$

Let $\omega = \mathrm{curl}\ \boldsymbol{v}$. One deduces from (22) that ω satisfies

$$\partial_t\omega + \mathrm{curl}\ [\omega(\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})^\perp] = 0.\qquad(23)$$

We multiply (23) by ω and integrate over $\mathbb{R}^2$ using the fact that $\nabla(\boldsymbol{v}-\gamma\mathfrak{R}[\zeta]\boldsymbol{v})\in L^\infty([0,T]\times\mathbb{R}^2)$ and proposition 1 to get after integration by parts

$$\frac{\mathrm{d}}{\mathrm{d}t}|\omega(t,.)|_2^2 \leqslant C\int_0^t|\omega(s,.)|_2^2\,\mathrm{d}s,$$

and the result follows by Gronwall's lemma. □

3.2. Solving the equations

3.2.1. The one-dimensional case $d = 1$. A simple computation shows that (12) is strictly hyperbolic provided that

$$\begin{cases} \inf_{\mathbb{R}}(1-\zeta) > 0, \\ \inf_{\mathbb{R}}(1+\delta\zeta) > 0, \\ \inf_{\mathbb{R}}\left[1-\gamma\left(1+\delta\frac{(1+\delta)^2}{(\delta+\gamma-\delta(1-\gamma)\zeta)^3}\boldsymbol{v}^2\right)\right] > 0. \end{cases} \tag{24}$$

The following theorem follows directly from standard results on hyperbolic systems. For the sake of completeness and because we will use it as a guideline for the more complex two-dimensional case, we nevertheless give a sketch of proof. Improved blow-up conditions are given in corollary 1.

Theorem 1. *Let $\delta > 0$ and $\gamma \in [0, 1)$. Let also $t_0 > 1/2$, $s \geqslant t_0 + 1$ and $U^0 = (\zeta^0, \boldsymbol{v}^0)^\top \in H^s(\mathbb{R})^2$ be such that (24) is satisfied. Then*

- *There exists $T_{\max} > 0$ and a unique maximal solution $U = (\zeta, \boldsymbol{v})^\top \in C([0, T_{\max}); H^s(\mathbb{R})^2)$ to (5) satisfying (24) on $[0, T_{\max})$ and with initial condition U^0.*
- *This solution satisfies the conservation of energy on $[0, T_{\max})$:*
$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}}[(1-\gamma)\zeta^2 + \boldsymbol{v}^2 f(\zeta)]\,\mathrm{d}x = 0,$$
with $f(\zeta)$ as in (14).
- *If $T_{\max} < \infty$ then $\lim_{t\to T_{\max}} |U(t)|_{W^{1,\infty}} = \infty$ or one of the three conditions of (24) ceases to be true as $t \to T_{\max}$.*

Proof. Thanks to proposition 4, it is equivalent to show existence/uniqueness of a solution to (12) with initial data U^0. Throughout this proof, we denote by $\mathfrak{c}_j(U)$ ($j = 1, 2, 3$) constants depending on the following quantities:

$$\mathfrak{c}_1(U) = C\left(\left|\frac{1}{1-\zeta}\right|_\infty, \left|\frac{1}{1+\delta\zeta}\right|_\infty, \left|\frac{1}{1-\gamma\left(1+\delta\frac{(1+\delta)^2}{(\delta+\gamma-\delta(1-\gamma)\zeta)^3}\boldsymbol{v}^2\right)}\right|_\infty\right)$$

and

$$\mathfrak{c}_2(U) = C(|U|_{W^{1,\infty}}, \mathfrak{c}_1(U)) \qquad \text{and} \qquad \mathfrak{c}_3(U) = C(|U|_{H^{t_0+1}}, \mathfrak{c}_1(U)).$$

Step 1. Construction of a regularized system of equations. Let χ be a smooth, even, compactly supported function defined over $\mathbb{R}$ and with values in $[0, \infty)$, and equal to 1 in a neighbourhood of the origin. For all $\iota > 0$, we define the operator χ_ι as

$$\chi_\iota = \chi(\iota D);$$

the operator χ_ι is thus a smoothing operator mapping continuously H^s into H^r for all $s, r \in \mathbb{R}$. The regularization of (12) is then defined as

$$\partial_t U^\iota + \chi_\iota(A(U^\iota)\chi_\iota(\partial_x U^\iota)) = 0. \tag{25}$$

Since U^0 satisfies (24), the mapping $U \mapsto \chi_\iota(A(U)\chi_\iota(\partial_x U))$ is locally Lipschitz in a neighbourhood of U^0 in H^s, for all $s \geqslant t_0 > 1/2$. Existence/uniqueness of a maximal solution $U^\iota \in C([0,T^\iota); H^s)$ (with $T^\iota > 0$) to (25) with initial condition U^0 and satisfying (24) is thus a direct consequence of the Cauchy–Lipschitz theorem for ODEs in Banach spaces.

Step 2. Choice of a symmetrizer. Let us introduce

$$S(U) = \begin{pmatrix} b(\zeta)^{-1} & 0 \\ 0 & c(U)^{-1} \end{pmatrix},$$

with $b(\zeta)$ and $c(U)$ given by (13). The matrix $S(U)$ is a symmetrizer in the sense that:

(1) The matrix $S(U)$ is symmetric and $|S(U)\cdot|_2$ is uniformly equivalent to $|\cdot|_2$: for all $V \in L^2(\mathbb{R})^{1+1}$,

$$|V|_2^2 \leqslant \mathfrak{c}_2(U)\,(S(U)V, V) \qquad \text{and} \qquad (S(U)V, V) \leqslant \mathfrak{c}_2(U)|V|_2^2. \tag{26}$$

(2) The matrix $S(U)A(U)$ is symmetric.

Step 3. Energy estimate. Let $\tilde{U} = \Lambda^s U^\iota$ and multiply (25) on the left-hand side by Λ^s. This yields the following equation on $\tilde{U}$:

$$\partial_t \tilde{U} + \chi_\iota(A(U^\iota)\chi_\iota(\partial_x \tilde{U})) = \chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)). \tag{27}$$

Since U^ι satisfies (24), $S(U^\iota(t,\cdot))$ is well defined on $[0,T^\iota)$, and we can thus multiply (27) on the left-hand side by $S(U^\iota)$ to obtain

$$S(U^\iota)\partial_t \tilde{U} + \chi_\iota(S(U^\iota)A(U^\iota)\chi_\iota(\partial_x \tilde{U})) = S(U^\iota)\chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)) + [\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}). \tag{28}$$

Multiplying (28) on the right-hand side by $\tilde{U}$ and integrating with respect to the space variable, one obtains

$$\tfrac{1}{2}\partial_t(S(U^\iota)\tilde{U}, \tilde{U}) + (S(U^\iota)A(U^\iota)\partial_x\chi_\iota\tilde{U}, \chi_\iota\tilde{U}) = \tfrac{1}{2}(\partial_t S(U^\iota)\tilde{U}, \tilde{U}) + (\chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)), S(U^\iota)\tilde{U}) + ([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}), \tilde{U}), \tag{29}$$

where we used the fact that $S(U^\iota)$ and χ_ι are self-adjoint and that ∂_x and χ_ι commute. We now turn to control the different terms involved in (29):

– *Control of* $(S(U^\iota)A(U^\iota)\partial_x\chi_\iota\tilde{U}, \chi_\iota\tilde{U})$. As said above, $S(U^\iota)A(U^\iota)$ is a symmetric matrix and it is thus standard to obtain the estimate

$$|(S(U^\iota)A(U^\iota)\partial_x\chi_\iota\tilde{U}, \chi_\iota\tilde{U})| \leqslant \tfrac{1}{2}|\partial_x(S(U^\iota)A(U^\iota))|_{L^2\to L^2}|\chi_\iota\tilde{U}|_2^2 \leqslant \mathfrak{c}_2(U^\iota)|\tilde{U}|_2^2. \tag{30}$$

– *Control of* $\frac{1}{2}(\partial_t S(U^\iota)\tilde{U}, \tilde{U})$. By a simple Cauchy–Schwarz inequality, we get

$$|\tfrac{1}{2}(\partial_t S(U^\iota)\tilde{U}, \tilde{U})| \leqslant C(\mathfrak{c}_2(U^\iota), |\partial_t U^\iota|_\infty)|\tilde{U}|_2^2.$$

Since for all $f \in L^\infty(\mathbb{R})$, one has $|\chi_\iota f|_\infty \lesssim |f|_\infty$, we can use equation (25) to control $|\partial_t U|_\infty$ and obtain

$$|\tfrac{1}{2}(\partial_t S(U^\iota)\tilde{U}, \tilde{U})| \leqslant \mathfrak{c}_2(U^\iota)|\tilde{U}|_2^2. \tag{31}$$

– *Control of* $(\chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)), S(U^\iota)\tilde{U})$. It is straightforward to get

$$|(\chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)), S(U^\iota)\tilde{U})| \leqslant \mathfrak{c}_2(U^\iota)|([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota))|_2|\tilde{U}|_2;$$

The commutator term is classically controlled thanks to Kato–Ponce and Moser estimates (see e.g. [28, chapter 16]) and one gets

$$|(\chi_\iota([A(U^\iota), \Lambda^s]\chi_\iota(\partial_x U^\iota)), S(U^\iota)\tilde{U})| \leqslant \mathfrak{c}_2(U^\iota)|\tilde{U}|_2^2. \tag{32}$$

– *Control of* $([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}), \tilde{U})$. Cauchy–Schwarz inequality yields

$$|([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}), \tilde{U})| \leqslant |([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U})|_2 |\tilde{U}|_2.$$

Thanks to the commutator estimate $|[\chi_\iota, f]u|_2 \lesssim |\nabla f|_{H^{t_0+1}} |u|_{H^{-1}}$ (the uniformness with respect to ι is a consequence of the general commutator estimate given in theorem 6 of [22]), we get

$$|([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}), \tilde{U})| \leqslant C(|U^\iota|_{H^{t_0+1}}, \mathfrak{c}_2(U^\iota))\,|A(U^\iota)\chi_\iota(\partial_x \tilde{U})|_{H^{-1}} |\tilde{U}|_2.$$

We can thus use the elementary product estimate (recall that $t_0 > \frac{1}{2}$) $|fg|_{H^{-1}} \lesssim |f|_{H^{t_0+1}} |g|_{H^{-1}}$ to get

$$|([\chi_\iota, S(U^\iota)]A(U^\iota)\chi_\iota(\partial_x \tilde{U}), \tilde{U})| \leqslant \mathfrak{c}_3(U^\iota)\,|\tilde{U}|_2^2. \tag{33}$$

We can now deduce from (29) and (30)–(33) that

$$\partial_t (S(U^\iota)\tilde{U}, \tilde{U}) \leqslant \mathfrak{c}_3(U^\iota)\,|\tilde{U}|_2^2. \tag{34}$$

By (26) and Gronwall's lemma, one deduces classically that for all $s \geqslant t_0 + 1$, there exists $T > 0$ independent of ι such that $T^\iota > T$ and

$$|U^\iota|_{L^\infty([0,T];H^s)} \leqslant M,$$

for some $M > 0$ independent of ι.

Step 4. Convergence of U^ι to a solution U to (12) as $\iota \to 0$. It is completely standard (see for instance chapter 16 of [28]) to prove that the sequence $(U^\iota)_\iota$ converges to some function $U \in C([0,T]; H^s(\mathbb{R}))$ $(s \geqslant t_0 + 1)$ which solves (12) and that such a solution is unique.

Step 5. Blow-up condition. Proceeding as in step 3, but working on the exact system and not on the regularized one, one can improve (34) into

$$\partial_t (S(U^\iota)\tilde{U}, \tilde{U}) \leqslant \mathfrak{c}_2(U^\iota)\,|\tilde{U}|_2^2$$

since (33) is not necessary if one works with the non-regularized system. The blow-up condition of the theorem then follows classically.

Step 6. Lastly, the conservation of energy is obtained by multiplying the first equation in (15) by $(1-\gamma)\zeta$, the second by $v f(\zeta)$ and then integrating by parts over $\mathbb{R}$. □

Remark 11. Setting

$$H(\zeta, v) = \frac{1}{2}\int_{\mathbb{R}} [(1-\gamma)\zeta^2 + v^2 f(\zeta)]\,\mathrm{d}x,$$

one can write (15) in Hamiltonian form (this corresponds to (5.24) in [12])

$$\partial_t U + \mathfrak{J}\nabla H(U) = 0, \tag{35}$$

where $\mathfrak{J}$ is the skew-adjoint operator $\mathfrak{J} = \partial_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

We now make more precise the blow-up condition of theorem 1.

Corollary 1. Under the assumptions of theorem 1, if the maximal existence time $T_{\max}$ is finite and if $\gamma > 0$ then:

- $U = (\zeta, v)^\top$ remains uniformly bounded on $[0, T_{\max}) \times \mathbb{R}$.
- $\lim_{t \to T_{\max}} |\partial_x U(t,\cdot)|_\infty = \infty$.

It is possible that the height of one of the fluids vanishes as $t \to T_{\max}$. In that case, additional information can be given on the blow-up of $\partial_x U(t,\cdot)$:

- If $\lim_{t\to T_{\max}} \inf_{\mathbb{R}}(1-\zeta(t,\cdot)) = 0$ then $\lim_{t\to T_{\max}} \sup_{\mathbb{R}} \partial_x \boldsymbol{v}(t,\cdot) = \infty$.
- If $\lim_{t\to T_{\max}} \inf_{\mathbb{R}}(1+\delta\zeta(t\cdot)) = 0$ then $\lim_{t\to T_{\max}} \inf_{\mathbb{R}} \partial_x \boldsymbol{v}(t,\cdot) = -\infty$.

Remark 12. The fact that a vanishing height for the upper (respectively lower) fluid induces an 'upward' (respectively 'downward') shock on the velocity can be observed on our numerical simulations (see figures 4 and 5, respectively).

Remark 13. It is of course possible to have a shock on the velocity without vanishing of the depth for the upper or lower fluid. This scenario can also be observed in our numerical computations (see figure 2). It follows also from step 2 of the proof of corollary 1 that in such a configuration, the integral $\int_0^{T_{\max}} \partial_x \boldsymbol{v}(s, Y_s^{-1}(\cdot))\,\mathrm{d}s$ converges (see the proof for a definition of Y_s).

Proof. The corollary follows from theorem 1 and the results proved in the three steps below.

Step 1. ζ remains uniformly bounded for all $t \in [0, T_{\max})$. First note that one can define a mapping $Y \in C^1([0, T_{\max}) \times \mathbb{R})$ by solving the following transport equation:

$$\partial_t Y + a(U(t))\partial_x Y = 0, \qquad Y_{|t=0} = x.$$

Setting $\zeta(t,x) = \tilde{\zeta}(t, Y(t,x))$ one can recast the first equation of (12) in the form

$$\forall (t,x) \in [0, T_{\max}) \times \mathbb{R}, \qquad \partial_t \tilde{\zeta}(t, Y(t,x)) + b(\tilde{\zeta}(t, Y(t,x)))\partial_x \boldsymbol{v}(t,x) = 0,$$

or equivalently,

$$\forall (t,y) \in [0, T_{\max}) \times \mathbb{R}, \qquad \partial_t \tilde{\zeta}(t,y) + b(\tilde{\zeta}(t,y))\partial_x \boldsymbol{v}(t, Y_t^{-1}(y)) = 0,$$

where $Y(t, Y_t^{-1}(y)) = y$. Since $b(\tilde{\zeta})$ does not vanish on $[0, T_{\max})$, we can use its explicit expression provided by (13) to restate this equation in the form

$$\partial_t(\ln(\tilde{h}_2/\tilde{h}_1^{\gamma}))(t,y) + \partial_x \boldsymbol{v}(t, Y_t^{-1}(y)) = 0,$$

where $\tilde{h}_1 = 1 - \tilde{\zeta}$ and $\tilde{h}_2 = 1 + \delta\tilde{\zeta}$. We thus deduce the relation

$$\tilde{h}_2(t,y) = \tilde{h}_1^{\gamma}(t,y)\frac{h_2(0,y)}{h_1(0,y)^{\gamma}} \exp\left(-\int_0^t \partial_x \boldsymbol{v}(s, Y_s^{-1}(y))\,\mathrm{d}s\right). \tag{36}$$

Since $\gamma > 0$, it follows from this expression that for $(t,x) \in [0, T_{\max}) \times \mathbb{R}$, one must have $\tilde{h}_1(t,x) > 0$ and $\tilde{h}_2(t,x) > 0$. This implies some bounds on $\tilde{\zeta}$ (and thus on ζ), namely,

$$\forall (t,x) \in [0, T_{\max}) \times \mathbb{R}, \qquad \zeta(t,x) \in \left(-\frac{1}{\delta}, 1\right),$$

and the claim is proved.

Step 2. We show here that

(1) If $h_1(t,\cdot)$ vanishes as $t \to T_{\max}$ then $\lim_{t\to T_{\max}} \sup_{\mathbb{R}} \partial_x \boldsymbol{v}(t,\cdot) = \infty$.
(2) If $h_2(t,\cdot)$ vanishes as $t \to T_{\max}$ then $\lim_{t\to T_{\max}} \inf_{\mathbb{R}} \partial_x \boldsymbol{v}(t,\cdot) = -\infty$.

Since quite obviously $\tilde{h}_1$ and $\tilde{h}_2$ cannot vanish together, it also follows from (36) and the fact that ζ remains bounded that

$$\lim_{t\to T} \inf_{\mathbb{R}} \tilde{h}_2(t,\cdot) = 0 \iff \lim_{t\to T} \sup_{\mathbb{R}} \int_0^t \partial_x \boldsymbol{v}(s, Y_s^{-1}(\cdot))\,\mathrm{d}s = \infty$$

$$\lim_{t\to T} \inf_{\mathbb{R}} \tilde{h}_1(t,\cdot) = 0 \iff \lim_{t\to T} \inf_{\mathbb{R}} \int_0^t \partial_x \boldsymbol{v}(s, Y_s^{-1}(\cdot))\,\mathrm{d}s = -\infty.$$

The first and second assertions of the claim then follow easily.

Step 3. We prove here that $v(t,\cdot) > 0$ remains uniformly bounded for all $t \in [0, T_{\max})$. This results directly from the fact that, by definition of $T_{\max}$, $c(U(t,\cdot)) > 0$ on $[0, T_{\max})$. It follows therefore from the explicit expression of $c(U)$ that

$$v^2 < \frac{1-\gamma}{\gamma\delta(1+\delta)^2}(\gamma + \delta - \delta(1-\gamma)\zeta)^3,$$

and the result stems from the uniform boundedness of ζ proved in step 1. □

We now show that the solutions in theorem 1 blow up for 'many' initial data (see figure 2 for a numerical illustration of this generic behaviour).

Proposition 7. *Under the assumptions of theorem 1 and with the same notations, one always has* $T_{\max} < \infty$ *if* $U^0 \neq (0,0)$ *is suitably compactly supported.*

Proof. The main point of the proof is the following lemma:

Lemma 2. *The hyperbolicity domain* $\mathcal{H}$ *for (12) is the domain delimited in the* (ζ, v^2) *plane by the axis* $\zeta = 0$*, the vertical lines* $\zeta = \frac{-1}{\delta}$, $\zeta = 1$ *and the graph of the decreasing function* $g(\zeta) = -2\frac{1-\gamma}{f''(\zeta)}$.

System (12) is genuinely nonlinear on a subdomain $\mathcal{H}_{\mathrm{GNL}}$ *of* $\mathcal{H}$ *such that* $\mathcal{H}\backslash\mathcal{H}_{\mathrm{GNL}}$ *contains at most four* C^1 *curves in the* (ζ, v^2) *plane. Moreover, there is at most one value of* ζ *such that the two characteristic fields are simultaneously linearly degenerate.*

Before proving this lemma, let us show how it allows us to conclude the proof of proposition 7. The proof is an immediate adaptation of classical blow-up results for systems having two genuinely nonlinear fields (see [14, theorem 7.8.2], [1, theorem 3.1], [21, 24, 25], sections 3.2 and 3.3). By [25] (theorem 3.2), one can, from any function $\tilde{U}_0 = (\tilde{\zeta}_0, \tilde{v}_0) \in \mathcal{H}_{\mathrm{GNL}}$, construct a family of simple wave solutions with compactly supported initial data which develop singularities in finite time. □

Proof of lemma 2. The hyperbolicity domain $\mathcal{H}$ (see (24)) can be recast as

$$\begin{cases} (1-\zeta) > 0, \\ (1+\delta\zeta) > 0, \\ (1-\gamma) + \frac{1}{2}f''(\zeta)v^2 > 0, \end{cases} \tag{37}$$

and the first part of the lemma follows easily. In order to determine $\mathcal{H}_{\mathrm{GNL}}$ we now compute the spectral quantities associated with (12).

The eigenvalues $\lambda_+(U)$ and $\lambda_-(U)$ of $A(U)$ are given by

$$\begin{cases} \lambda_+(U) = a(U) + \sqrt{b(\zeta)c(U)} = f'(\zeta)v + f(\zeta)^{\frac{1}{2}}F(\zeta,v)^{\frac{1}{2}}, \\ \lambda_-(U) = a(U) - \sqrt{b(\zeta)c(U)} = f'(\zeta)v - f(\zeta)^{\frac{1}{2}}F(\zeta,v)^{\frac{1}{2}}, \end{cases} \tag{38}$$

where a, b and c are given by (13) and $F(\zeta, v) = (1-\gamma)+\frac{1}{2}f''(\zeta)v^2$, provided the hyperbolicity conditions (37) are satisfied. Corresponding systems of right and left eigenvectors are

$$\begin{cases} R_+(U) = (\sqrt{b(\zeta)}, \sqrt{c(U)})^\top = (f(\zeta)^{\frac{1}{2}}, F(\zeta,v)^{\frac{1}{2}})^\top, \\ R_-(U) = (\sqrt{b(\zeta)}, -\sqrt{c(U)})^\top = (f(\zeta)^{\frac{1}{2}}, -F(\zeta,v)^{\frac{1}{2}})^\top. \end{cases} \tag{39}$$

$$\begin{cases} L_+(U) = (\sqrt{c(U)}, \sqrt{b(\zeta)})^\top = (F(\zeta,v)^{\frac{1}{2}}, f(\zeta)^{\frac{1}{2}})^\top, \\ L_-(U) = (\sqrt{c(U)}, -\sqrt{b(\zeta)})^\top = (F(\zeta,v)^{\frac{1}{2}}, -f(\zeta)^{\frac{1}{2}})^\top. \end{cases} \tag{40}$$

(Note that, on the other hand, the Riemann invariants do not seem to have an obvious explicit form.)

The two characteristic fields are genuinely nonlinear if $(\nabla\lambda_\pm(U), R_\pm(U)) \neq 0$. Let us consider first the field associated with $\lambda_+(U)$. A simple computation shows that (we skip the arguments of the functions):

$$4F^{\frac{1}{2}}(\nabla\lambda_+(U), R_+(U)) = G(\zeta, v), \qquad \text{with} \qquad G(\zeta,v) = 6f'F + ff'''v^2 + 6f^{\frac{1}{2}}F^{\frac{1}{2}}f''v.$$

Since $F(\zeta, v) = (1-\gamma) + \frac{1}{2}f''(\zeta)v^2$, if follows that $G(\zeta, v) = 0$ if and only if

$$6f'(1-\gamma) + 3f'f''v^2 + ff'''v^2 + 6f^{\frac{1}{2}}f''v((1-\gamma) + \tfrac{1}{2}f''v^2)^{\frac{1}{2}} = 0.$$

Squaring and setting $X = v^2$, this amounts to saying that

$$[\alpha_+^2 - 18ff''^3]X^2 + 12(1-\gamma)[f'\alpha^+ - 3ff''^2]X + 36f'^2(1-\gamma)^2 = 0, \quad (41)$$

with $\alpha_+ = 3f'f'' + ff'''$. The reduced discriminant of (41) is written, after some calculations

$$\Delta_+ = 108(1-\gamma)^2 f^2 f''^2[3f''^2 - 2f'f'''] > 0.$$

There are therefore possibly two C^1 curves in the (ζ, v^2) plane on which the first characteristic field is not genuinely nonlinear; these curves correspond to the solutions of (41). For the characteristic field associated with $\lambda_-(U)$ one finds instead of (41)

$$[\alpha_-^2 - 18ff''^3]X^2 + 12(1-\gamma)[f'\alpha_- - 3ff''^2]X + 36f'^2(1-\gamma)^2 = 0, \quad (42)$$

with $\alpha_- = 3f'f'' - ff'''$. The corresponding discriminant is not necessarily positive, but in any case, there are at most two C^1 curves on which the second characteristic field is degenerate; they correspond to the solutions of (42).

Note finally that both characteristic fields are simultaneously linearly degenerate if and only if $f'(\zeta)$ vanishes since (41)–(42) $= 72(1-\gamma)v^2 ff'f'''$. When $-\frac{1}{\delta} < \zeta < 1$, this can happen for only one value ζ_0 of ζ, $-\frac{1}{\delta} < \zeta_0 < 0$ if $\delta^2 - \gamma < 0$, and $0 < \zeta_0 < 1$ if $\delta^2 - \gamma > 0$. The condition $\delta^2 = \gamma$ is the critical depth ratio which corresponds to the well-known singularity for small-amplitude long internal waves in two-layer flows (see e.g. [3]). □

Remark 14. Using the conservative form (5) in the formulation (15), one derives the Rankine–Hugoniot condition for a piecewise C^1 solution

$$\begin{cases} (\eta_+ - \eta_-)n_t + [f(\zeta_+) - f(\zeta_-)]n_x = 0, \\ (v_+ - v_-)n_t + (1-\gamma)(\zeta_+ - \zeta_-)n_x + \frac{1}{2}[(f'(\zeta_+)v_+^2) - (f'(\zeta_-)v_-^2)]n_x = 0, \end{cases} \quad (43)$$

along the surfaces of discontinuity. A mathematical study of weak solutions to (5) has not been performed yet, but Bouchut and Zeitlin [6] recently managed to handle numerically shock waves for a similar type of system.

Remark 15. We plot in figure 1 the domain of hyperbolicity $\mathcal{H}$ and the curves on which the two characteristic fields are degenerate (in the (ζ, v^2) plane). For the computations, we took $\gamma = 0.1$ and $\delta = 5$. The degeneracy curves for the second characteristic field do not exist for the smallest values of ζ. This is because there exists a critical value of ζ below which the discriminant Δ_- of (42) is negative and (42) does not have real solutions.

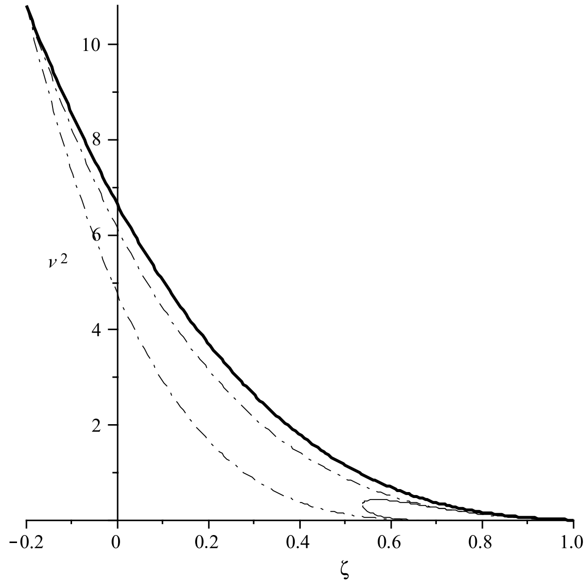

Figure 1. The curve g delimiting the hyperbolicity domain (bold solid line), the degeneracy curves for the first and second characteristic fields (dashed-dotted and solid lines).

3.2.2. The two-dimensional case $d = 2$. We show here that the two-dimensional shallow water/shallow water equations (6) are locally well-posed under the following conditions that generalize the hyperbolicity conditions (24) of the one-dimensional case,

$$\begin{cases} 1 - |\zeta|_\infty > 0, \\ 1 - \delta|\zeta|_\infty > 0, \\ 1 - \gamma - \gamma\delta \dfrac{|\mathfrak{S}[\zeta]\boldsymbol{v}|_\infty^2}{\gamma + \delta - \delta(1-\gamma)|\zeta|_\infty} > 0, \end{cases} \tag{44}$$

with $\mathfrak{S}[\zeta]\boldsymbol{v}$ as in (9).

Theorem 2. *Let $\delta > 0$ and $\gamma \in [0, 1)$. Let also $t_0 > 1$, $s \geqslant t_0 + 1$ and $U^0 = (\zeta^0, \boldsymbol{v}^0)^\top \in H^s(\mathbb{R}^2)^3$ be such that (44) is satisfied and* curl $\boldsymbol{v}^0 = 0$. *Then, there exists $T_{\max} > 0$ and a unique maximal solution $U = (\zeta, \boldsymbol{v})^\top \in C([0, T_{\max}); H^s(\mathbb{R}^2)^3)$ to (6) with initial condition U^0. Moreover, if $T_{\max} < \infty$ then at least one of the following conditions holds:*

(i) $\lim_{t \to T_{\max}} |U(t)|_{H^{t_0+1}} = \infty.$

(ii) One of the three conditions of (44) is enforced as $t \to T_{\max}$.

Proof. For the sake of clarity, we try to follow as much as possible the structure of the proof of theorem 1. Throughout this proof, we denote by $\mathfrak{c}(U)$ any constant of the form

$$\mathfrak{c}(U) = C\left(\frac{1}{1 - |\zeta|_{H^{t_0}}}, \frac{1}{1 - \delta|\zeta|_{H^{t_0}}}, \frac{1}{1 - \gamma - \gamma\delta \dfrac{|\mathfrak{S}[\zeta]\boldsymbol{v}|_\infty^2}{\delta + \gamma - \delta(1-\gamma)|\zeta|_\infty}}, |U|_{H^{t_0+1}}\right).$$

Step 1. Regularized equations. As in the one-dimensional case, we construct a regularized system of equations. We still denote by χ_ι the two-dimensional generalization of the smoothing operator χ_ι introduced in the first step of the proof of theorem 1,

$$\chi_\iota = \chi(\iota|D|);$$

the regularization of (16) is then

$$\partial_t U^\iota + \chi_\iota(A^j[U^\iota]\chi_\iota(\partial_j U^\iota)) = 0; \tag{45}$$

existence/uniqueness of a maximal solution $U^\iota = (\zeta^\iota, \boldsymbol{v}^\iota)^\top \in C([0, T^\iota); H^s)$ $(s \geqslant t_0, T^\iota > 0)$ with initial condition U^0 to (45) satisfying (44) is thus obtained as in the proof of theorem 1 thanks to the following lemma (recall that a^j, $\boldsymbol{b}^j$, $\boldsymbol{c}^j$ and D^j are defined in (17)–(20)):

Lemma 3. *Let $s \geqslant t_0 > 1$ and $U^0 = (\zeta^0, \boldsymbol{v}^0)^\top \in H^s(\mathbb{R}^2)^3$ be such that (44) is satisfied. Then, there exists a neighbourhood $\mathcal{U}$ of U^0 in $H^s(\mathbb{R}^2)^3$ such that the mappings*

$$U \in \mathcal{U} \mapsto a^j(U) \in H^s(\mathbb{R}^2);$$

$$U \in \mathcal{U} \mapsto \boldsymbol{b}^j(U) \in H^s(\mathbb{R}^2)^2,$$

$$U \in \mathcal{U} \mapsto \boldsymbol{c}^j[U]\bullet \in \mathcal{L}(H^s(\mathbb{R}^2); H^s(\mathbb{R}^2)^2),$$

$$U \in \mathcal{U} \mapsto D^j[U]\bullet \in \mathcal{L}(H^s(\mathbb{R}^2)^2; H^s(\mathbb{R}^2)^2),$$

are well defined and smooth.

Proof. The assertions concerning $a^j(\cdot)$ and $\boldsymbol{b}^j(\cdot)$ are obvious from the explicit expressions of these mappings provided by (17) and (18), respectively. For $\boldsymbol{c}^j[U]$ and $D^j[U]$, the result follows from the regularity estimates on $\mathfrak{R}[\zeta]$ given in proposition 1. □

Moreover, proceeding as for proposition 6, one has (since curl $\boldsymbol{v}^0 = 0$)

$$\forall t \in [0, T^\iota), \qquad \text{curl } \boldsymbol{v}^\iota = 0. \tag{46}$$

Step 2. Choice of a symmetrizer. Let us look for $S[U]$ in the form

$$S[U] = \begin{pmatrix} s_1(U) & 0 \\ 0 & S_2[U] \end{pmatrix}, \tag{47}$$

with $s_1(\cdot) : H^s(\mathbb{R}^2)^3 \mapsto H^s(\mathbb{R}^2)$ and $S_2[U]$ a linear operator mapping $L^2(\mathbb{R}^2)^2$ into itself. Defining $C[U]$ as

$$\forall \tilde{\boldsymbol{v}} = (\tilde{\boldsymbol{v}}_1, \tilde{\boldsymbol{v}}_2)^\top \in L^2(\mathbb{R}^2)^2, \qquad C[U]\tilde{\boldsymbol{v}} = \boldsymbol{c}_1[U]\tilde{\boldsymbol{v}}_1 + \boldsymbol{c}_2[U]\tilde{\boldsymbol{v}}_2,$$

a straightforward generalization of the one-dimensional case consists of taking $s_1(U) = b(U)^{-1}$ and $S_2[U] = C[U]^{-1}$; unfortunately, such a choice is not correct because the operator $C[U]$ is not self-adjoint. It turns out, however, that $C[U]$ is self-adjoint (up to a smoothing term) on the restriction of $L^2(\mathbb{R}^2)^2$ to gradient vector fields, as shown in the following lemma. We first need to define the operator $C_1[U]$ as

$$C_1[U] = (1-\gamma)\text{Id} + \frac{1}{2}\delta\gamma \begin{pmatrix} c_1[U] + c_1[U]^* & 0 \\ 0 & c_1[U] + c_1[U]^* \end{pmatrix}, \tag{48}$$

with $c_1[U] : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ given by

$$c_1[U] = \frac{1}{\delta h_1 + \gamma h_2}(2\mathfrak{S}_1\mathfrak{S}_2\Pi(e^2\cdot)_1 + \mathfrak{S}_1^2\Pi(e^1\cdot)_1 + \mathfrak{S}_2^2\Pi(e^2\cdot)_2), \tag{49}$$

and where $\mathfrak{S}_j = (\mathfrak{S}[\zeta]\boldsymbol{v})_j$.

Lemma 4. *Let* $t_0 > 1$ *and* $U = (\zeta, \boldsymbol{v})^\top \in H^{t_0+1}(\mathbb{R}^2)^3$ *be such that (44) is satisfied. Define also* $C_1[U]$ *as in (48) and let* $C_2[U] = C[U] - C_1[U]$*. For all* $\tilde{\zeta} \in L^2(\mathbb{R}^2)$*, one has*

$$|C_2[U]\nabla\tilde{\zeta}|_2 \leqslant \mathfrak{c}(U)|\tilde{\zeta}|_2.$$

Proof. It follows from the definition of $C[U]$ and the explicit expression of $\boldsymbol{c}_1[U]$ and $\boldsymbol{c}_2[U]$ given by (19) that one can decompose $C[U]$ into

$$C[U] = \tilde{C}_1[U] + \tilde{C}_{2,1}[U] + \tilde{C}_{2,2}[U]$$

where for all $\boldsymbol{w} \in L^2(\mathbb{R}^2)^2$, using the summation convention on repeated indices,

$$\tilde{C}_1[U]\boldsymbol{w} = (1-\gamma)\boldsymbol{w} + \delta\gamma\frac{\mathfrak{S}_j\mathfrak{S}_k}{\delta h_1 + \gamma h_2}\Pi(\boldsymbol{e}^k\boldsymbol{w}_j),$$

$$\tilde{C}_{2,1}[U]\boldsymbol{w} = \delta\gamma\mathfrak{S}_j\left(\mathfrak{R}[\zeta]\left(\frac{\mathfrak{S}[\zeta]\boldsymbol{v}}{h_2}\boldsymbol{w}_j\right) - \frac{1}{\delta h_1 + \gamma h_2}\Pi\left((\mathfrak{S}[\zeta]\boldsymbol{v})\boldsymbol{w}_j\right)\right),$$

$$\tilde{C}_{2,2}[U]\boldsymbol{w} = \delta\gamma\frac{\mathfrak{S}_j}{\delta h_1 + \gamma h_2}[\Pi, \mathfrak{S}_k](\boldsymbol{e}^k\boldsymbol{w}_j).$$

It is a direct consequence of proposition 3 and of the product estimate (11) that

$$|\tilde{C}_{2,1}[U]\boldsymbol{w}|_2 \lesssim \mathfrak{c}(U)|\mathfrak{S}_j|_\infty|\Pi((\mathfrak{S}[\zeta]\boldsymbol{v})\boldsymbol{w}_j)|_{H^{-1}}$$
$$\lesssim \mathfrak{c}(U)|\mathfrak{S}_j|_\infty|\mathfrak{S}[\zeta]\boldsymbol{v}|_{H^{t_0}}|\boldsymbol{w}_j|_{H^{-1}},$$

which in turn yields (using proposition 1 to control $|\mathfrak{R}[\zeta]\boldsymbol{v}|_{H^{t_0}}$ and thus $|\mathfrak{S}[\zeta]\boldsymbol{v}|_{H^{t_0}}$),

$$|\tilde{C}_{2,1}[U]\nabla\tilde{\zeta}|_2 \leqslant \mathfrak{c}(U)|\tilde{\zeta}|_2. \tag{50}$$

With the help of the commutator estimate (10), we can also note that

$$|\tilde{C}_{2,2}[U]\nabla\tilde{\zeta}|_2 \lesssim \delta\gamma\left|\frac{\mathfrak{S}_j}{\delta h_1 + \gamma h_2}\right|_\infty|\mathfrak{S}[\zeta]\boldsymbol{v}|_{H^{t_0+1}}|\tilde{\zeta}|_2;$$

using the last part of proposition 1 to control $|\mathfrak{R}[\zeta]\boldsymbol{v}|_{H^{t_0+1}}$ and thus $|\mathfrak{S}[\zeta]\boldsymbol{v}|_{H^{t_0+1}}$, we get therefore

$$|\tilde{C}_{2,2}[U]\nabla\tilde{\zeta}|_2 \leqslant \mathfrak{c}(U)|\tilde{\zeta}|_2. \tag{51}$$

Noting also that $\Pi(\boldsymbol{e}^2\partial_1\tilde{\zeta})_1 = \Pi(\boldsymbol{e}^1\partial_1\tilde{\zeta})_2 = \Pi(\boldsymbol{e}^1\partial_2\tilde{\zeta})_1$ and $\Pi(\boldsymbol{e}^2\partial_2\tilde{\zeta})_1 = \Pi(\boldsymbol{e}^2\partial_1\tilde{\zeta})_2$, we can rewrite $\tilde{C}_1[U]\nabla\tilde{\zeta}$ as

$$\tilde{C}_1[U]\nabla\tilde{\zeta} = (1-\gamma)\nabla\tilde{\zeta} + \delta\gamma\begin{pmatrix} c_1[U] & 0 \\ 0 & c_1[U]\end{pmatrix}\nabla\tilde{\zeta}$$
$$= C_1[U] + \tilde{C}_{2,3}[U]\nabla\tilde{\zeta},$$

with $C_1[U]$ and $c_1[U]$ as in (48) and (49) while $\tilde{C}_{2,3}[U]$ is given by

$$\tilde{C}_{2,3}[U] = \frac{1}{2}\begin{pmatrix} c_1[U] - c_1[U]^* & 0 \\ 0 & c_1[U] - c_1[U]^*\end{pmatrix}.$$

With the commutator estimate (10), we get

$$|\tilde{C}_{2,3}[U]\nabla\tilde{\zeta}|_2 \leqslant \mathfrak{c}(U)|\tilde{\zeta}|_2. \tag{52}$$

Setting $C_2[U] = \tilde{C}_{2,1}[U] + \tilde{C}_{2,2}[U] + \tilde{C}_{2,3}[U]$, the result follows from (50), (51) and (52). □

In view of this lemma, we now choose the coefficients $s_1[U]$ and $S_2[U]$ of the symmetrizer $S[U]$ given by (47) as follows:

$$s_1(U) = b(U)^{-1}, \tag{53}$$

$$S_2[U] = C_1[U]^{-1}; \tag{54}$$

the invertibility of $C_1[U]$ is ensured by the following lemma:

Lemma 5. *Let $t_0 > 1$ and $U = (\zeta, \boldsymbol{v})^\top \in H^{t_0}(\mathbb{R}^2)^3$ be such that (44) is satisfied.*

(1) The operator $C_1[U]$ is invertible in $\mathcal{L}(L^2(\mathbb{R}^2)^2; L^2(\mathbb{R}^2))$ and

$$\|C_1[U]^{-1}\|_{L^2\to L^2} \leqslant \mathfrak{c}(U).$$

(2) The following coercivity property holds, for all $\tilde{\boldsymbol{v}} \in L^2(\mathbb{R}^2)^2$,

$$|\tilde{\boldsymbol{v}}|_2^2 \leqslant \mathfrak{c}(U)(C_1[U]^{-1}\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{v}}).$$

Proof. It follows from (48) that $C_1[U]$ can be seen as a perturbation of $(1-\gamma)\mathrm{Id}$ in $\mathcal{L}(L^2(\mathbb{R}^2)^2; L^2(\mathbb{R}^2)^2)$; the norm of this perturbation being bounded from above by $\delta\gamma\|c_1[U]\|_{L^2\to L^2}$, it follows that $C_1[U]$ is invertible if

$$\delta\gamma\|c_1[U]\|_{L^2\to L^2} < (1-\gamma). \tag{55}$$

Recalling that by convention, $|\mathfrak{S}[\zeta]\boldsymbol{v}|_\infty = |(\mathfrak{S}[\zeta]\boldsymbol{v})_1|_\infty + |(\mathfrak{S}[\zeta]\boldsymbol{v})_2|_\infty$, we get from (49) that

$$\|c_1[U]\|_{L^2\to L^2} \leqslant \frac{|\mathfrak{S}[\zeta]\boldsymbol{v}|_\infty^2}{\gamma+\delta-\delta(1-\gamma)|\zeta|_\infty}.$$

Condition (55) is thus satisfied under the assumptions of the lemma, and the invertibility of $C_1[U]$ follows. The bound on $\|C_1[U]^{-1}\|_{L^2\to L^2}$ can be deduced easily from a coercivity property on $C_1[U]$ that we establish now. Just note that

$$(C_1[U]\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{v}}) \geqslant (1-\gamma)|\tilde{\boldsymbol{v}}|_2^2 - \delta\gamma\|c_1[U]\|_{L^2\to L^2}|\tilde{\boldsymbol{v}}|_2^2$$

and use the above bound on $\|c_1[U]\|_{L^2\to L^2}$ to get

$$\mathfrak{c}(U)(C_1[U]\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{v}}) \geqslant |\tilde{\boldsymbol{v}}|_2^2.$$

The bound on $\|C_1[U]^{-1}\|_{L^2\to L^2}$ is then obtained by replacing $\tilde{\boldsymbol{v}}$ by $C_1[U]^{-1}\tilde{\boldsymbol{v}}$ in this inequality.

We now turn to prove the coercivity property on $C_1[U]^{-1}$ stated in the lemma. Since $C_1[U]$ and $C_1[U]^{-1}$ are both self-adjoint, bounded, positive operators, they admit unique square roots, denoted, respectively, $C_1[U]^{-1/2}$ and $C_1[U]^{1/2}$, which are also positive. We can therefore write

$$\begin{aligned}|\tilde{\boldsymbol{v}}|_2^2 &= |C_1[U]^{1/2}C_1[U]^{-1/2}|_2^2\\ &\leqslant \|C_1[U]^{1/2}\|_{L^2\to L^2}^2|C_1[U]^{-1/2}|_2^2\\ &= \|C_1[U]^{1/2}\|_{L^2\to L^2}^2(C_1[U]^{-1}\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{v}}).\end{aligned}$$

Since $\|C_1[U]^{1/2}\|_{L^2\to L^2} = \|C_1[U]\|_{L^2\to L^2}^{1/2}$, we can deduce from the above computations that

$$|\tilde{\boldsymbol{v}}|_2^2 \leqslant \mathfrak{c}(U)(C_1[U]^{-1}\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{v}}),$$

which is exactly the result claimed in the lemma. □

It follows directly from this lemma that $S[U]$ satisfies the following generalization of (26): for all $V \in L^2(\mathbb{R}^2)^{1+2}$,

$$|V|_2^2 \leqslant \mathfrak{c}(U)\,(S[U]V, V) \qquad \text{and} \qquad (S[U]V, V) \leqslant \mathfrak{c}(U)|V|_2^2. \tag{56}$$

The operator $S[U]$ would therefore be a symmetrizer in the sense given in step 2 of the proof of theorem 1 if $S[U]A^j[U]$ $(j = 1, 2)$ were symmetric, which is unfortunately not the case. However, $\Pi S[U]A^j[U]\Pi$, where Π denotes as before the projection onto gradient vector fields, is symmetric at leading order. This crucial property will be exploited in step 3 below and is a consequence of lemma 4 and of the following lemma:

Lemma 6. *Let* $t_0 > 1$ *and* $U = (\zeta, \boldsymbol{v})^\top \in H^{t_0+1}(\mathbb{R}^2)^3$ *satisfying (44). One can decompose the operators* $D^j[U]$ $(j = 1, 2)$ *given by (20) into*

$$D^j[U] = d_1^j(U)\mathrm{Id} + D_2^j[U]$$

with

$$d_1^j(U) = \left(\boldsymbol{v} - \gamma\mathfrak{R}[\zeta]\boldsymbol{v} - \gamma\frac{h_2}{\delta h_1 + \gamma h_2}\mathfrak{S}[\zeta]\boldsymbol{v}\right)_j.$$

Then, for all $\tilde{\boldsymbol{v}} \in L^2(\mathbb{R}^2)^2$ *such that* $\Pi\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}$, *one has*

$$|D_2^j[U]\partial_j\tilde{\boldsymbol{v}}|_2 \leqslant \mathfrak{c}(U)|\tilde{\boldsymbol{v}}|_2.$$

Proof. It follows from the explicit expression (20) of $D^j[U]$ and the assumption that $\Pi\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}$ that one can write

$$D_2^j[U]\partial_j\tilde{\boldsymbol{v}} = -\gamma(\mathfrak{S}[\zeta]\boldsymbol{v})_j\left(\mathfrak{R}[\zeta]\left(\frac{h_2\partial_j\tilde{\boldsymbol{v}}}{h_2}\right) - \frac{1}{\delta h_1 + \gamma h_2}\Pi(h_2\partial_j\tilde{\boldsymbol{v}})\right)$$
$$-(\mathfrak{S}[\zeta]\boldsymbol{v})_j\frac{1}{\delta h_1 + \gamma h_2}[\Pi, h_2]\partial_j\tilde{\boldsymbol{v}}.$$

Using propositions 1 and 3 and (11) to control the first component of the right-hand side of the above identity, and proposition 1 and (10) to control the second one, we get the result. □

Step 3. Energy estimate. Proceeding as in the proof of theorem 1, one can check that $\tilde{U} = \Lambda^s U^\iota$ solves

$$\partial_t\tilde{U} + \chi_\iota(A^j[U^\iota]\chi_\iota(\partial_j\tilde{U})) = \chi_\iota([A^j[U^\iota], \Lambda^s]\chi_\iota(\partial_j U^\iota)), \tag{57}$$

and one obtains the following two-dimensional generalization of (29):

$$\tfrac{1}{2}\partial_t(S[U^\iota]\tilde{U}, \tilde{U}) + (S[U^\iota]A^j[U^\iota]\partial_j\chi_\iota\tilde{U}, \chi_\iota\tilde{U}) = \tfrac{1}{2}([\partial_t, S[U^\iota]]\tilde{U}, \tilde{U})$$
$$+ (\chi_\iota([A^j[U^\iota], \Lambda^s]\chi_\iota(\partial_j U^\iota)), S[U^\iota]\tilde{U}) + ([\chi_\iota, S[U^\iota]]A^j[U^\iota]\chi_\iota(\partial_j\tilde{U}), \tilde{U}). \tag{58}$$

As in the proof of theorem 1, we now intend to control all the components of (58).

– *Control of* $(S[U^\iota]A^j[U^\iota]\partial_j\chi_\iota\tilde{U}, \chi_\iota\tilde{U})$. Using the explicit expression of $A^j[U^\iota]$, we get

$$(S[U^\iota]A^j[U^\iota]\partial_j\chi_\iota\tilde{U}, \chi_\iota\tilde{U}) = (s_1(U^\iota)a^j(U^\iota)\partial_j\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\zeta}) + (s_1(U^\iota)b(U^\iota)\nabla\cdot\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\zeta})$$
$$+ (S_2[U^\iota]C[U^\iota]\nabla\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\boldsymbol{v}}) + (S_2[U^\iota]D^j[U^\iota]\partial_j\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}}), \tag{59}$$

and we thus have to bound from above the different components of the right-hand side of (59).

- *Estimate on* $(s_1(U^\iota)a^j(U^\iota)\partial_j\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\zeta})$. With a simple integration by parts, and using the explicit formulas of $s_1(U^\iota)$ and $a^j(U^\iota)$ provided by (53) and (17), one obtains

$$|(s_1(U^\iota)a^j(U^\iota)\partial_j\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\zeta})| \leqslant \mathfrak{c}(U^\iota)|\tilde{\zeta}|_2^2. \tag{60}$$

- *Estimate on* $I := (s_1(U^\iota)b(U^\iota)\nabla\cdot\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\zeta}) + (S_2[U^\iota]C[U^\iota]\nabla\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\boldsymbol{v}})$. Replacing $s_1(U^\iota)$ and $S_2[U^\iota]$ by their expressions given by (53) and (54), we get immediately

$$I = -(\nabla\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\boldsymbol{v}}) + (C_1[U^\iota]^{-1}C[U^\iota]\nabla\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\boldsymbol{v}})$$
$$= (C_1[U^\iota]^{-1}C_2[U^\iota]\nabla\chi_\iota\tilde{\zeta}, \chi_\iota\tilde{\boldsymbol{v}}),$$

where we used the decomposition $C[U^\iota] = C_1[U^\iota] + C_2[U^\iota]$ of lemma 4. Using the bounds on $\|C_1[U^\iota]^{-1}\|_{L^2\to L^2}$ and $|C_2[U^\iota]\nabla\chi_\iota\tilde{\zeta}|$ provided by lemmas 5 and 4, respectively, we deduce that

$$|I| \leqslant \mathfrak{c}(U^\iota)|\tilde{\zeta}|_2|\tilde{\boldsymbol{v}}|_2. \tag{61}$$

- *Estimate on* $J := (S_2[U^\iota]D^j[U^\iota]\partial_j\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}})$. With $d_1^j[U^\iota]$ and $D_2^j[U^\iota]$ as in lemma 6, we can write

$$J = (S_2[U^\iota]d_1^j(U^\iota)\partial_j\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}}) + (S_2[U^\iota]D_2^j[U^\iota]\partial_j\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}})$$
$$:= J_1 + J_2.$$

Let us decompose J_1 into

$$2J_1 = -(S_2[U^\iota]\partial_j(d_1^j(U^\iota))\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}}) - ([\partial_j, S_2[U^\iota]]d_1^j(U^\iota)\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}})$$
$$+ ([S_2[U^\iota], d_1^j(U^\iota)]\partial_j\chi_\iota\tilde{\boldsymbol{v}}, \chi_\iota\tilde{\boldsymbol{v}}).$$

Recalling that $S_2[U^\iota] = C_1[U^\iota]^{-1}$, the first term on the right-hand side is easily controlled thanks to lemma 5 and the explicit expression of $d_1^j(U^\iota)$. We can thus deduce that

$$|J_1| \leqslant \mathfrak{c}(U^\iota)|\tilde{\boldsymbol{v}}|_2^2$$

from the following lemma:

Lemma 7. *Let* $t_0 > 1$. *For all* $U \in H^{t_0+1}$ *satisfying (44), one has*

$$\|[\partial_j, C_1[U]^{-1}]\|_{L^2\to L^2} \leqslant \mathfrak{c}(U)$$

and

$$\|[C_1[U]^{-1}, d_1^j(U)]\partial_j\|_{L^2\to L^2} \leqslant \mathfrak{c}(U).$$

Proof. For the first part of the lemma, note that

$$[\partial_j, C_1[U]^{-1}] = C_1[U]^{-1}[\partial_j, C_1[U]]C_1[U]^{-1}.$$

It is thus a consequence of lemma 5 that

$$\|[\partial_j, C_1[U]^{-1}]\|_{L^2\to L^2} \leqslant \mathfrak{c}(U)\|[\partial_j, C_1[U]]\|_{L^2\to L^2};$$

the operator norm of the commutator estimate on the right-hand side is bounded from above by $\mathfrak{c}(U)$ as a consequence of the general commutator estimate given in theorem 6 of [22], so that the result follows. Since the operator norm of a bounded operator is equal to the operator norm of its adjoint, one has

$$\|[C_1[U]^{-1}, d_1^j(U)]\partial_j\|_{L^2\to L^2} = \|\partial_j[C_1[U]^{-1}, d_1^j(U)]\|_{L^2\to L^2}.$$

Noting that

$$[C_1[U]^{-1}, d_1^j(U)] = C_1[U]^{-1}[C_1[U], d_1^j(U)]C_1[U]^{-1},$$

and recalling that $\|Q_1[U]^{-1}\|_{L^2\to L^2} \leqslant \mathfrak{c}(U)$, we deduce that

$$\|[C_1[U]^{-1}, d_1^j(U)]\partial_j\|_{L^2\to L^2} \leqslant \|\partial_j C_1[U]^{-1}[C_1[U], d_1^j(U)]C_1[U]^{-1}\|_{L^2\to L^2}$$
$$\leqslant \mathfrak{c}(U)(1 + \|[\partial_j, C_1[U]^{-1}]\|_{L^2\to L^2})\|[C_1[U], d_1^j(U)]\|_{L^2\to H^1}.$$

Using the first point of the lemma, this gives

$$\|[C_1[U]^{-1}, d_1^j(U)]\partial_j\|_{L^2\to L^2} \leqslant \mathfrak{c}(U)\|[C_1[U], d_1^j(U)]\|_{L^2\to H^1}.$$

Since (10), with $r = 1$, can be invoked to control the operator norm of the commutator on the right-hand side by $\mathfrak{c}(U)$, the result follows. □

In order to control J_2, first note that $\Pi(\chi_\iota \tilde{\boldsymbol{v}}) = \chi_\iota \tilde{\boldsymbol{v}}$ (this follows from the identity curl $\boldsymbol{v}^\iota = 0$ stated in (46)). It is thus a direct consequence of lemmas 6 and 5 that $|J_2|$ has the same upper bound as $|J_1|$, so that

$$|J| \leqslant \mathfrak{c}(U^\iota)|\tilde{\boldsymbol{v}}|_2^2. \tag{62}$$

It is now easy to deduce from (59), (60), (61) and (62) that the following two-dimensional generalization of (30) holds

$$|(S[U^\iota]A^j[U^\iota]\partial_j \chi_\iota \tilde{U}, \chi_\iota \tilde{U})| \leqslant \mathfrak{c}(U^\iota)|\tilde{U}|_2^2. \tag{63}$$

– *Control of* $\frac{1}{2}([\partial_t, S[U^\iota]]\tilde{U}, \tilde{U})$. Without particular additional difficulty with respect to the one-dimensional case, we get the following two-dimensional generalization of (31),

$$|\tfrac{1}{2}([\partial_t, S[U^\iota]]\tilde{U}, \tilde{U})| \leqslant \mathfrak{c}(U^\iota)|\tilde{U}|_2^2. \tag{64}$$

– *Control of* $(\chi_\iota([A^j[U^\iota], \Lambda^s]\chi_\iota(\partial_j U^\iota)), S[U^\iota]\tilde{U})$. This term can be handled with standard commutator estimates as in the one-dimensional case (and using also lemma 5); one thus obtains the following two-dimensional generalization of (32),

$$|(\chi_\iota([A^j[U^\iota], \Lambda^s]\chi_\iota(\partial_j U^\iota)), S[U^\iota]\tilde{U})| \leqslant \mathfrak{c}(U^\iota)|\tilde{U}|_2^2. \tag{65}$$

– *Control of* $([\chi_\iota, S[U^\iota]]A^j[U^\iota]\chi_\iota(\partial_j \tilde{U}), \tilde{U})$. By Cauchy–Schwarz inequality, one gets easily

$$|([\chi_\iota, S[U^\iota]]A^j[U^\iota]\chi_\iota(\partial_j \tilde{U}), \tilde{U})| \leqslant \|[\chi_\iota, S[U^\iota]]A^j[U^\iota]\partial_j\|_{L^2\to L^2}|\tilde{U}|_2^2.$$

We can now note that

$$\begin{aligned}\|[\chi_\iota, S[U^\iota]]A^j[U^\iota]\partial_j\|_{L^2\to L^2} &= \|\partial_j A^j[U^\iota]^*[\chi_\iota, S[U^\iota]]\|_{L^2\to L^2}\\ &\leqslant \mathfrak{c}(U^\iota)\|[\chi_\iota, S[U^\iota]]\|_{L^2\to H^1},\end{aligned}$$

where we used the fact that $\|A^j[U^\iota]\|_{L^2\to L^2}$ and $\|[\partial_j, A^j[U^\iota]]\|_{L^2\to L^2}$ are bounded from above by $\mathfrak{c}(U)$ to derive the last inequality. To see that $\|[\chi_\iota, S[U^\iota]]\|_{L^2\to H^1}$ is also bounded from above by the same quantity, we proceed as in the proof of lemma 7. We thus generalize (33) as

$$|([\chi_\iota, S[U^\iota]]A^j[U^\iota]\chi_\iota(\partial_j \tilde{U}), \tilde{U})| \leqslant \mathfrak{c}(U^\iota)|\tilde{U}|_2^2. \tag{66}$$

Thanks to (63)–(66), we can proceed exactly as in the one-dimensional case to conclude that there exists $T > 0$ independent of ι such that $T^\iota > T$ and

$$|U^\iota|_{L^\infty([0,T];H^s)} \leqslant M,$$

for some $M > 0$ independent of ι.

Step 4. Convergence of U^ι to a solution U of (6) as $\iota \to 0$. Exactly as in the one-dimensional case, one can check that there exists $T > 0$ such that U^ι converges to $U \in C([0, T]; H^s(\mathbb{R}^2))^3$ ($s \geqslant t_0+1$) which solves (16) and that such a solution is unique. The fact that this solution is also the unique solution to (6) is not automatic as in the one-dimensional case since proposition 5 requires $\boldsymbol{v}$ to remain curl free. This property is ensured by proposition 6 since we assumed that curl $\boldsymbol{v}^0 = 0$.

Step 5. Blow-up condition. The blow-up condition is provided by a completely standard continuation argument. □

Remark 16. We do not know whether or not system (6) possesses a conserved energy or a Hamiltonian structure. It would be of interest to use the approach proposed in [12] to derive a Hamiltonian version of (6). The existence of two-dimensional blow-up solutions, although highly expected, is also unknown.

4. Numerical computations

In this section, we describe the numerical methods for spatial and temporal discretizations of the SW/SW systems (5) and (6). Unless stated otherwise, for expository convenience, we describe these methods in the general two-dimensional case. We then present numerical simulations that support the well-posedness results established in the previous sections.

4.1. Numerical methods

For spatial discretization, we assume periodic boundary conditions in both x- and y-directions, and use a pseudospectral method [17]. In particular, this is a natural choice for the computation of the nonlocal operator $\mathfrak{R}[\zeta]$ since each term in its Neumann series (7) is a concatenation of Fourier multipliers (through the operator Π) with ζ. We will restrict our numerical simulations to solutions that decay fast at infinity so that, by specifying a sufficiently large domain, the effects of the lateral boundary conditions can be neglected.

The pseudospectral method consists of approximating ζ and $\boldsymbol{v}$ by truncated Fourier series

$$\begin{pmatrix} \zeta \\ \boldsymbol{v} \end{pmatrix} = \sum_{\boldsymbol{k}} \begin{pmatrix} \hat{\zeta} \\ \hat{\boldsymbol{v}} \end{pmatrix} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot X},$$

where $\boldsymbol{k} = (k_1, k_2)^\top = (\frac{2\pi n_1}{L_1}, \frac{2\pi n_2}{L_2})^\top$,

$$n_1 = \left\{-\frac{N_1}{2} + 1, \ldots, \frac{N_1}{2}\right\}, \qquad n_2 = \left\{-\frac{N_2}{2} + 1, \ldots, \frac{N_2}{2}\right\},$$

N_1 and N_2 are the numbers of modes (or equivalently the numbers of grid points) and L_1 and L_2 are the lengths of the domain, in the x- and y-directions, respectively. Applications of Fourier multipliers are performed in Fourier space, while nonlinear products are calculated in physical space at a discrete set of equally spaced points $X = (x, y)^\top = (\frac{j_1 L_1}{N_1}, \frac{j_1 L_2}{N_2})^\top$ where

$$j_1 = \{0, \ldots, N_1 - 1\}, \qquad j_2 = \{0, \ldots, N_2 - 1\}.$$

For example, if we wish to apply the operator $\Pi = \frac{\nabla\nabla^\top}{\Delta}$ to a function $\boldsymbol{u}$ in physical space, we transform $\boldsymbol{u}$ to Fourier space, multiply the operator $\frac{\boldsymbol{k}\boldsymbol{k}^\top}{|\boldsymbol{k}|^2}$ to the Fourier coefficients of $\boldsymbol{u}$, and then transform back to physical space. These operations can be performed efficiently by using the fast Fourier transform.

The nonlocal operator $\mathfrak{R}[\zeta]$ that appears in the two-dimensional system (6) can be represented by its Neumann series (7) and, in numerical simulations, this series is also approximated by a finite number of terms,

$$\mathfrak{R}[\zeta]\boldsymbol{v} = \frac{1}{\gamma + \delta} \sum_{n=0}^{N_t} \left(\Pi \left(\frac{1-\gamma}{\gamma+\delta} \varepsilon\delta\zeta\,\Pi\, \cdot \right)\right)^n \Pi(h_2 \boldsymbol{v}),$$

where N_t is chosen according to the level of accuracy desired. As will be shown below, we have found $N_t = 4$ to be a good compromise between accuracy and speed.

No special attention is paid to deal with aliasing errors for two practical reasons. First, aliasing occurs when evaluating nonlinear terms. A common and effective way to remove these errors would be to extend the spectra of ζ and $\boldsymbol{v}$ and set the values of the extra modes to zero (zero-padding technique). The higher the order of nonlinearity, the more the spectra have to be extended. However, in the two-dimensional case, the Neumann series of $\mathfrak{R}[\zeta]$ involves nonlinearities of increasing order whose treatment by zero padding would require substantial memory storage, which would severely affect the computational time. Second, aliasing errors are most significant when a small number of modes is used (i.e. when the solution is poorly resolved). Therefore, to reduce their effects, care is taken to specify a sufficiently large number of modes in all of our simulations. By doing so, we have always observed that very good global accuracy (such as mass conservation) is achieved and that the solution evolves well towards the predicted instability.

For time integration, we employ a fourth-order Runge–Kutta scheme with constant time step, which represents a good compromise between accuracy, speed and stability. It can be applied to systems (5) and (6) indifferently in Fourier or physical space. For the convenience of the reader, below we describe this scheme when it is applied to the one-dimensional model (5) in physical space:

$$\begin{aligned}
\dot{U}_1 &= F(U^n), \\
\dot{U}_2 &= F\left(U^n + \frac{\Delta t}{2}\dot{U}_1\right), \\
\dot{U}_3 &= F\left(U^n + \frac{\Delta t}{2}\dot{U}_2\right), \\
\dot{U}_4 &= F\left(U^n + \Delta t\,\dot{U}_3\right), \\
U^{n+1} &= U^n + \frac{\Delta t}{6}(\dot{U}_1 + 2\dot{U}_2 + 2\dot{U}_3 + \dot{U}_4),
\end{aligned}$$

for the solution $U^{n+1} = (\zeta^{n+1}, \boldsymbol{v}^{n+1})^\top$ at time $t_{n+1} = t_n + \Delta t$, where Δt denotes the time step and

$$F(U) = \begin{pmatrix} -\mathcal{F}^{-1}\left[\mathrm{i}k_1\mathcal{F}\left(\dfrac{h_1 h_2}{\delta h_1 + \gamma h_2}\boldsymbol{v}\right)\right] \\ -(1-\gamma)\mathcal{F}^{-1}[\mathrm{i}k_1\mathcal{F}(\zeta)] - \dfrac{1}{2}\mathcal{F}^{-1}\left[\mathrm{i}k_1\mathcal{F}\left(\dfrac{(\delta h_1)^2 - \gamma h_2^2}{(\delta h_1 + \gamma h_2)^2}\boldsymbol{v}^2\right)\right] \end{pmatrix}.$$

The procedure is similar in the two-dimensional case.

4.2. Numerical results

4.2.1. The one-dimensional case $d = 1$. In this section, we solve the one-dimensional system (5) using the numerical methods described above, and we check the information provided by corollary 1 on finite time singularity formation by choosing suitable initial conditions and parameter values. We focus on the following situations:

– *Generic shock generation*. We specify a 'hump' (i.e. a localized wave profile) as initial condition for both ζ and $\boldsymbol{v}$, of the form

$$\zeta^0(x) = 2\frac{c}{\eta}\sqrt{\frac{\alpha}{\beta}}\frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}},$$

$$\boldsymbol{v}^0(x) = 2\frac{\sqrt{\alpha\beta}}{\eta}\frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}},$$

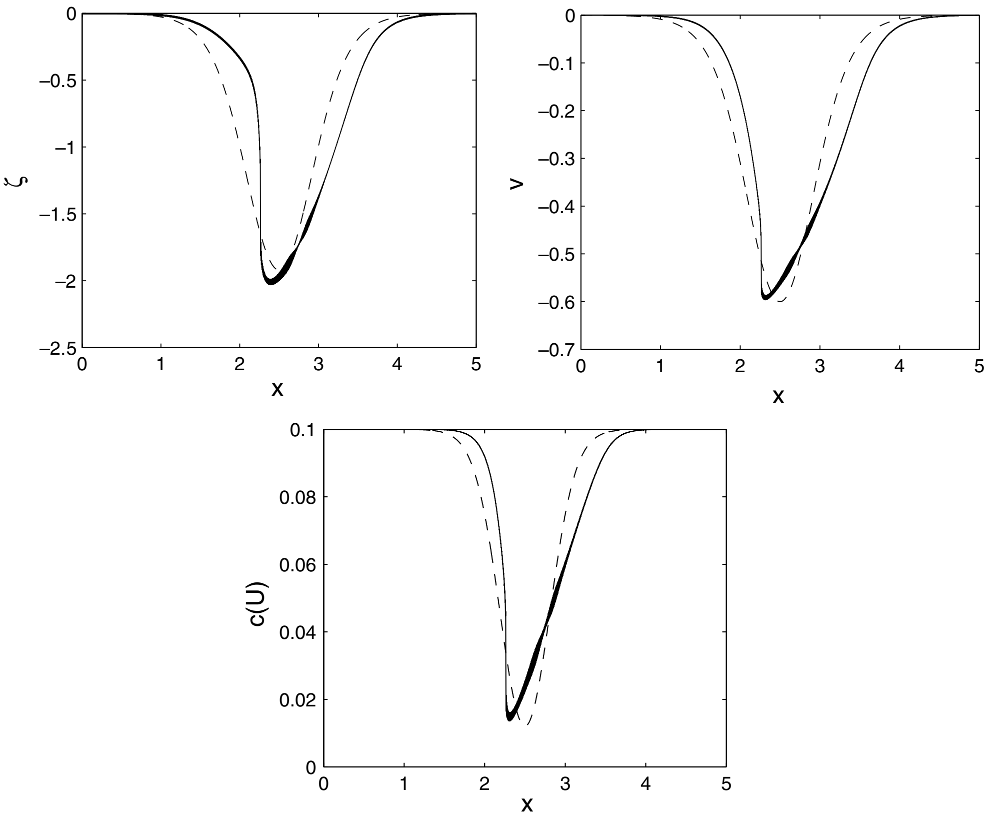

Figure 2. Profiles of ζ, v and $c(U)$ at $t = 0$ (- - - -) and $t = 0.78$ (——) for $\delta = 1/3$, $\gamma = 0.9$ and $c = 0.32$.

where

$$\alpha = \frac{\delta}{\delta + \gamma}, \qquad \beta = 1 - \gamma, \qquad \eta = \frac{\delta^2 - \gamma}{(\delta + \gamma)^2}.$$

We use the following values for the numerical parameters: $\delta = 1/3$, $\gamma = 0.9$, $c = 0.32$, $L_1 = 5$, $N_1 = 2048$ and $\Delta t = 10^{-5}$.

Figure 2 shows that both ζ and v tend to form a shock (i.e. $|\partial_x \zeta|_\infty \to \infty$ and $|\partial_x v|_\infty \to \infty$ while (ζ, v) remain bounded) as predicted. In the present parameter regime, the bottom and rigid lid are located at $z = -3$ and $z = 1$, respectively. We can see that the bottom is never reached by the interface as the shock develops, and $c(U)$ as defined by (13) remains positive. Small oscillations around the crest and on the right side of the profiles at $t = 0.78$ are the sign of the imminent computation breakdown due to the shock occurrence. These oscillations are manifestations of the Gibbs phenomenon, indicating the inability of the pseudospectral scheme to accurately resolve the full development of the shock.

As a typical illustration, figure 3 depicts the time evolution of the relative errors in mass and energy conservation in this case. The total mass M and energy E [12] of the system are defined, respectively, by

$$M = \int_{\mathbb{R}} \zeta \, \mathrm{d}x,$$

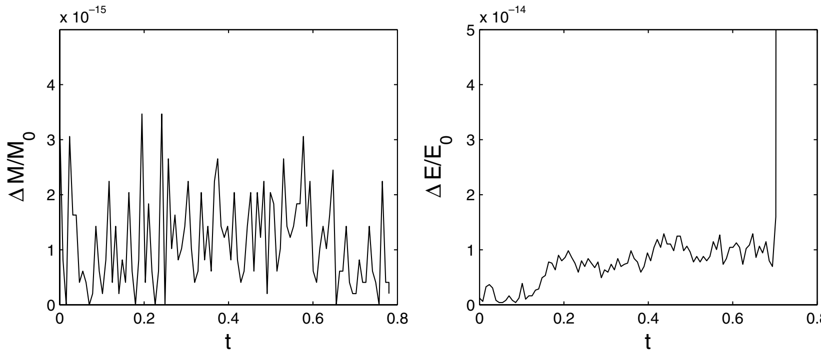

Figure 3. Relative errors on mass and energy conservation versus time for $\delta = 1/3$, $\gamma = 0.9$ and $c = 0.32$.

and

$$E = \frac{1}{2}\int_{\mathbb{R}} \left[(1-\gamma)\zeta^2 + \frac{h_1 h_2}{\delta h_1 + \gamma h_2} v^2\right] \mathrm{d}x,$$

M_0 and E_0 being the values at the initial time $t = 0$. As indicated in figure 3, both mass and energy are very well conserved, with relative errors of order $O(10^{-15})$–$O(10^{-14})$ (near machine precision). This confirms the overall accuracy and effectiveness of the numerical methods used to solve the models. Note, however, that, from $t \simeq 0.7$, there is a rapid deterioration of energy conservation, which is likely due to rapid accumulation of numerical errors and the occurrence of instabilities as the solution approaches the shock.

– *Vanishing depth for the upper fluid.* Corollary 1 predicts that if the depth of the upper fluid vanishes, then a positive shock appears simultaneously on the velocity profile $\boldsymbol{v}$ (by positive shock, we mean that $\sup_{\mathbb{R}} \partial_x \boldsymbol{v} \to +\infty$). Here the initial conditions are given by the superposition of two humps,

$$\zeta^0(x) = \frac{7}{2} c_g \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}} + \frac{7}{2} c_g \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} - 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}, \tag{67}$$

$$\boldsymbol{v}^0(x) = \frac{7}{2} \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}} - \frac{7}{2} \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} - 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}, \tag{68}$$

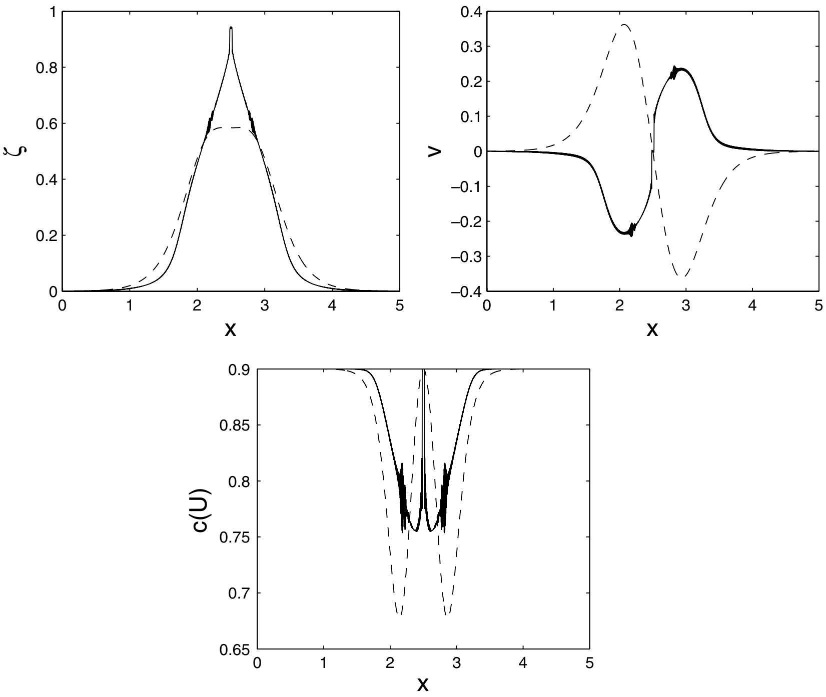

Figure 4. Profiles of ζ, v and $c(U)$ at $t = 0$ (- - - -) and $t = 0.58$ (——) for $\delta = 1$, $\gamma = 0.1$ and $c = 1$.

where

$$c_g = \frac{1}{\sqrt{(1-\gamma)(\delta+\gamma)}}.$$

The numerical parameters are set to be $\delta = 1$, $\gamma = 0.1$, $c = 1$, $L_1 = 5$, $N_1 = 2048$ and $\Delta t = 10^{-5}$.

Figure 4 shows that, as ζ increases and gets closer to the rigid lid (located at $z = 1$), the profile of v tends to form a shock, approaching the vertical asymptote with a positive slope. This occurs at two points located symmetrically near $x = L_1/2$. Here again, we can see that $c(U)$ remains positive and high-wavenumber instabilities start to develop in the profiles at $t = 0.58$. The computation breaks down shortly after this time.

– *Vanishing depth for the lower fluid*. Corollary 1 predicts that if the depth of the lower fluid vanishes, then a negative shock appears simultaneously on the velocity profile v (that is, $\inf_{\mathbb{R}} \partial_x v \to -\infty$). Here again the initial conditions are given by the superposition of two humps but with opposite signs,

$$\zeta^0(x) = -\frac{7}{2}c_g\frac{\sqrt{\alpha\beta}}{\eta}\frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(x-\frac{L_1}{2}+0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}} - \frac{7}{2}c_g\frac{\sqrt{\alpha\beta}}{\eta}\frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(x-\frac{L_1}{2}-0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}},$$

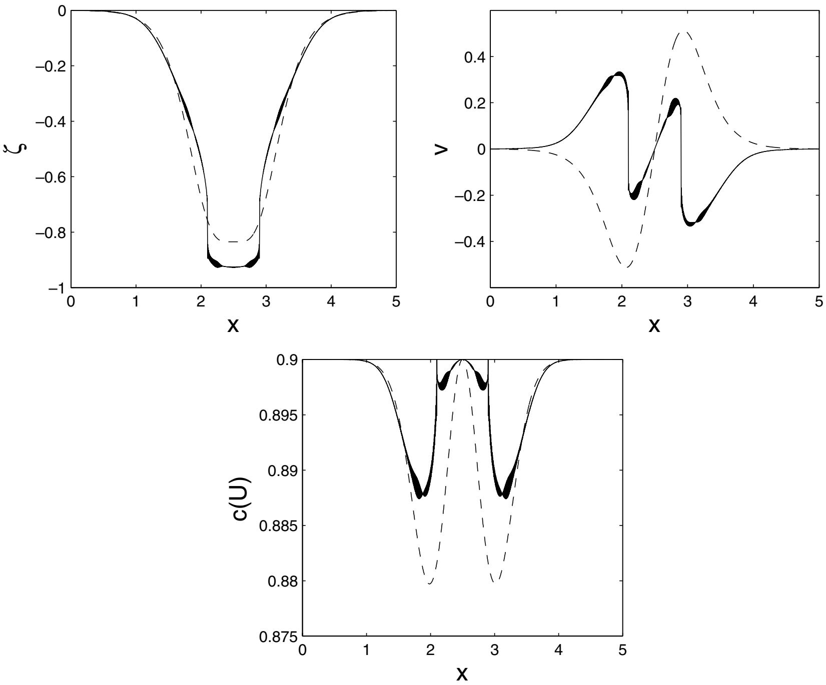

Figure 5. Profiles of ζ, v and $c(U)$ at $t = 0$ (- - - -) and $t = 0.68$ (——) for $\delta = 1$, $\gamma = 0.1$ and $c = 1.04$.

$$v^0(x) = -\frac{7}{2}\frac{\sqrt{\alpha\beta}}{\eta}\frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}$$

$$+\frac{7}{2}\frac{\sqrt{\alpha\beta}}{\eta}\frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} - 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}.$$

The numerical parameters are set to be $\delta = 1$, $\gamma = 0.1$, $c = 1.04$, $L_1 = 5$, $N_1 = 2048$ and $\Delta t = 10^{-5}$.

As predicted, we observe the situation opposite to that described previously (figure 5). The interface reaches down to the bottom, while the profile of v develops a shock at two points located symmetrically about $x = L_1/2$. At these two points, v approaches the vertical asymptote with a negative slope. A noticeable difference with the previous situation is that the wave crest for ζ is broader and smoother, which may be explained by the fact that here the interface is of depression (i.e. of negative amplitude) and is penetrating the much denser lower fluid ($\gamma = 0.1$).

– *Vanishing of the coefficient* $c(U)$. In all the computations presented above, the coefficient $c(U)$ remained nonnegative. It is, however, possible to take initial conditions that lead to vanishing of $c(U)$. Here, we numerically check this behaviour. The initial conditions are

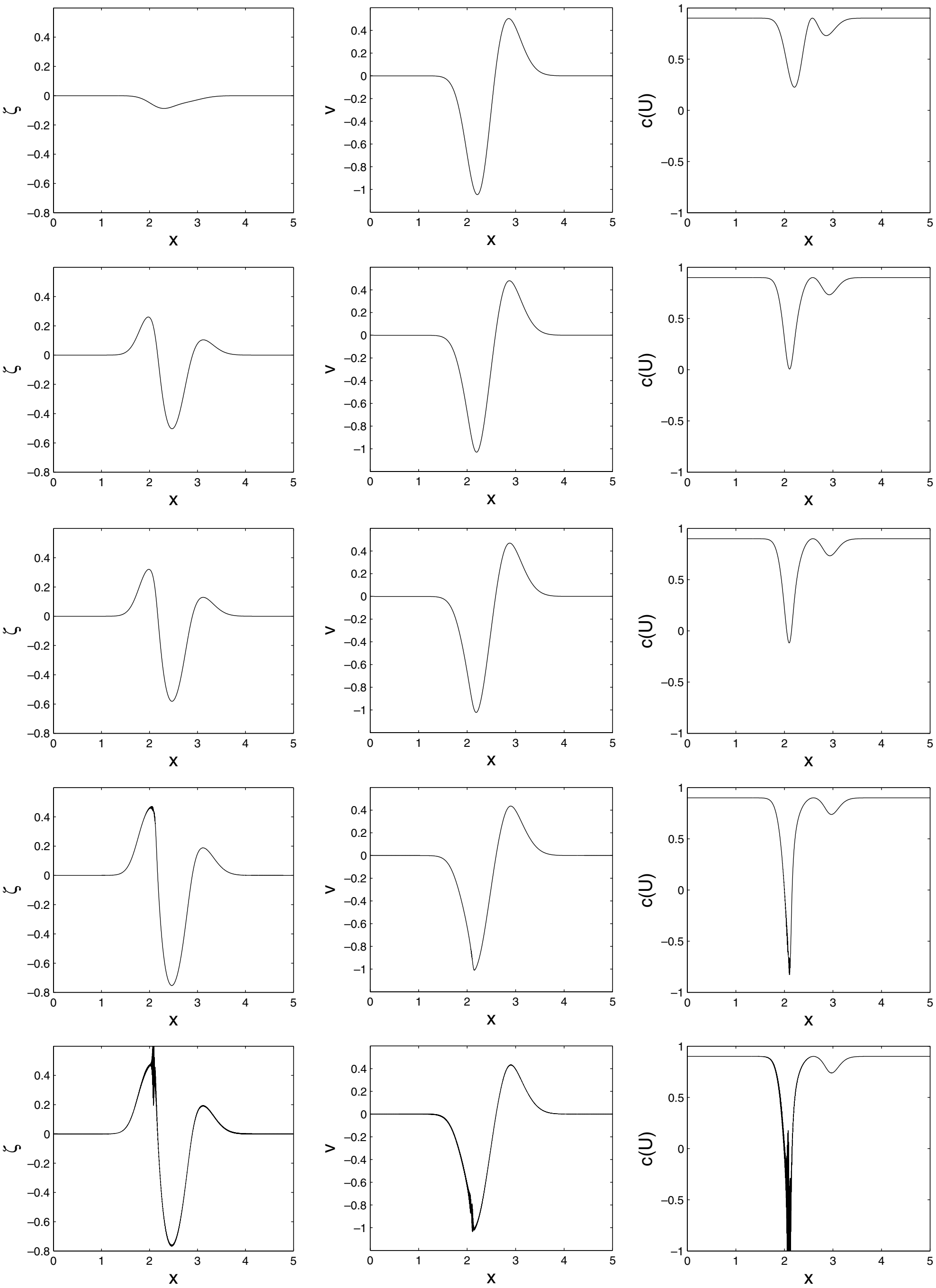

Figure 6. Sequence of profiles of ζ (left), v (middle) and $c(U)$ (right) at $t = 0$, 0.050, 0.061, 0.088, 0.090 (from top to bottom) for $\delta = 1/3$, $\gamma = 0.1$ and $c = -0.075$.

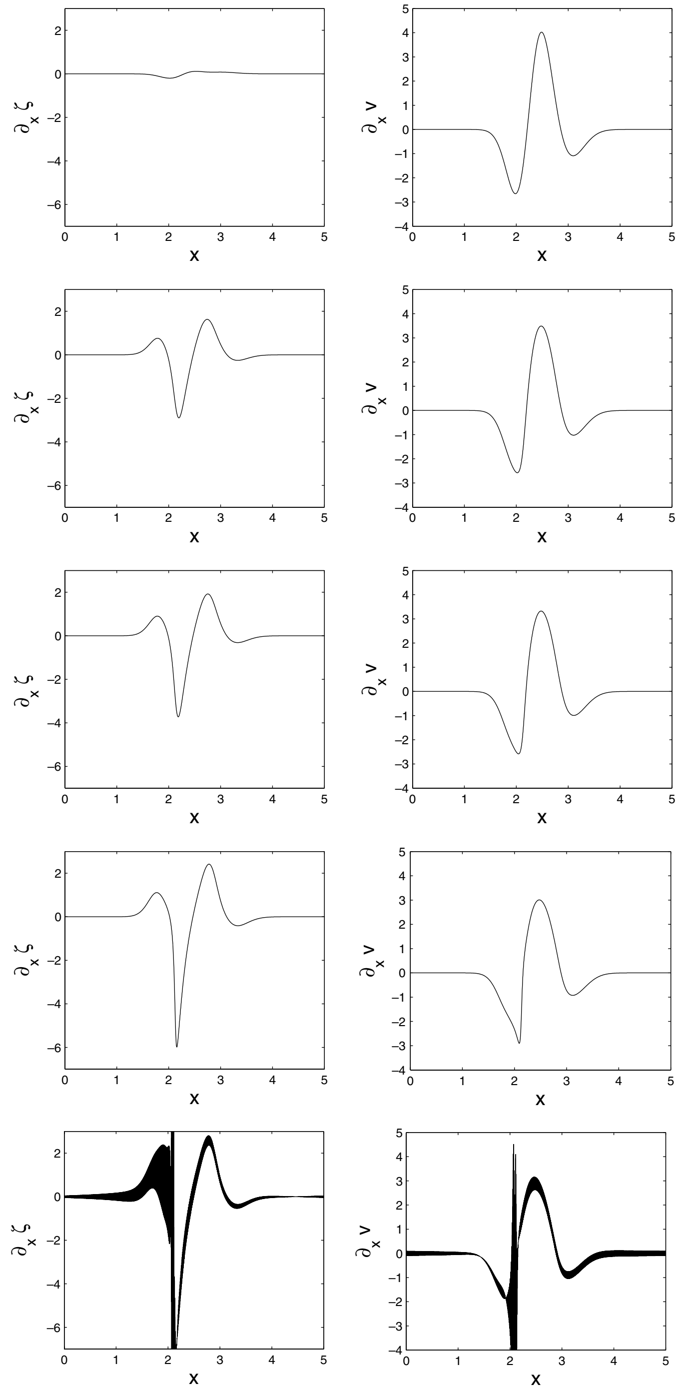

Figure 7. Sequence of profiles of $\partial_x \zeta$ and $\partial_x v$ at $t = 0$, 0.050, 0.061, 0.081, 0.088 (from top to bottom) for $\delta = 1/3$, $\gamma = 0.1$ and $c = -0.075$.

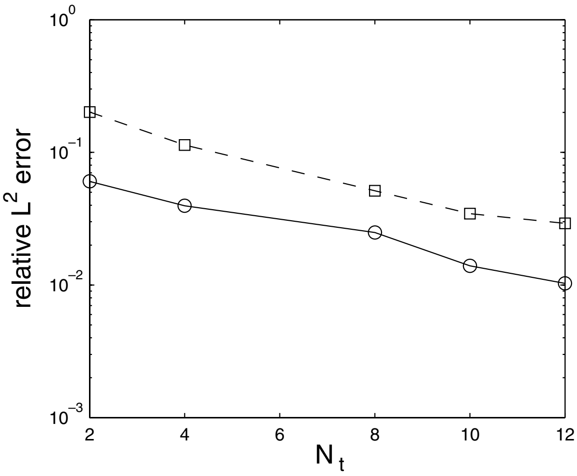

Figure 8. Relative L^2 errors in ζ (——) and v_1 (- - - -) between one- and two-dimensional profiles at $t = 0.58$, as functions of N_t ($\delta = 1$, $\gamma = 0.1$, $c = 1$).

given by

$$\zeta^0(x) = c\mathrm{e}^{-8(x-0.45L_1)^2} + \frac{c}{2}\mathrm{e}^{-5(x-0.55L_1)^2},$$

$$v^0(x) = 10c_g c\mathrm{e}^{-8(x-0.45L_1)^2} - 5c_g c\mathrm{e}^{-5(x-0.55L_1)^2}.$$

The numerical parameters are set to be $\delta = 1/3$, $\gamma = 0.1$, $c = -0.075$, $L_1 = 5$, $N_1 = 2048$ and $\Delta t = 10^{-5}$.

Figure 6 plots the profiles of ζ, v and $c(U)$ for various values of t ($t = 0$, 0.050, 0.061, 0.088, 0.090). Initially, $c(U)$ is strictly positive. Consistent with corollary 1, we see that instabilities occur shortly after $c(U)$ changes sign, which quickly leads to the computation breakdown. Here these instabilities start to appear at $t \simeq 0.088$, around the tallest crest in the profile of ζ, which corresponds to the trough in the profile of v. The fact that they do not occur as soon as $c(U) = 0$ has to do with numerical diffusion (i.e. the discretization of the problem). We have checked that, by refining the spatial resolution, instabilities occur sooner, in closer agreement with the theoretical prediction. Note also that neither ζ nor v has yet formed a shock when the solution blows up. This observation is supported by figure 7 which shows that both $\partial_x \zeta$ and $\partial_x v$ remain bounded.

4.2.2. The two-dimensional case $d = 2$. We now solve numerically the two-dimensional system (6). A key parameter that controls the accuracy of our two-dimensional computations is N_t, the number of terms in the Neumann series of $\mathfrak{R}[\zeta]$. Therefore, we first want to test the convergence of numerical results with respect to this parameter. This will help us decide on a suitable choice for N_t to be used in subsequent simulations. For this purpose, we consider the same one-dimensional problem used to study the 'vanishing depth for the upper fluid' in the previous section, and compare two-dimensional computations with one-dimensional results for the same parameter values. The one-dimensional model does not involve any nonlocal operator and thus constitutes a good reference for this comparison. In addition, we feel this test is suitable for testing the convergence of the Neumann series defining $\mathfrak{R}[\zeta]$ with respect to N_t because it is a highly nonlinear situation where the solution reaches a large amplitude and develops a sharp crest.

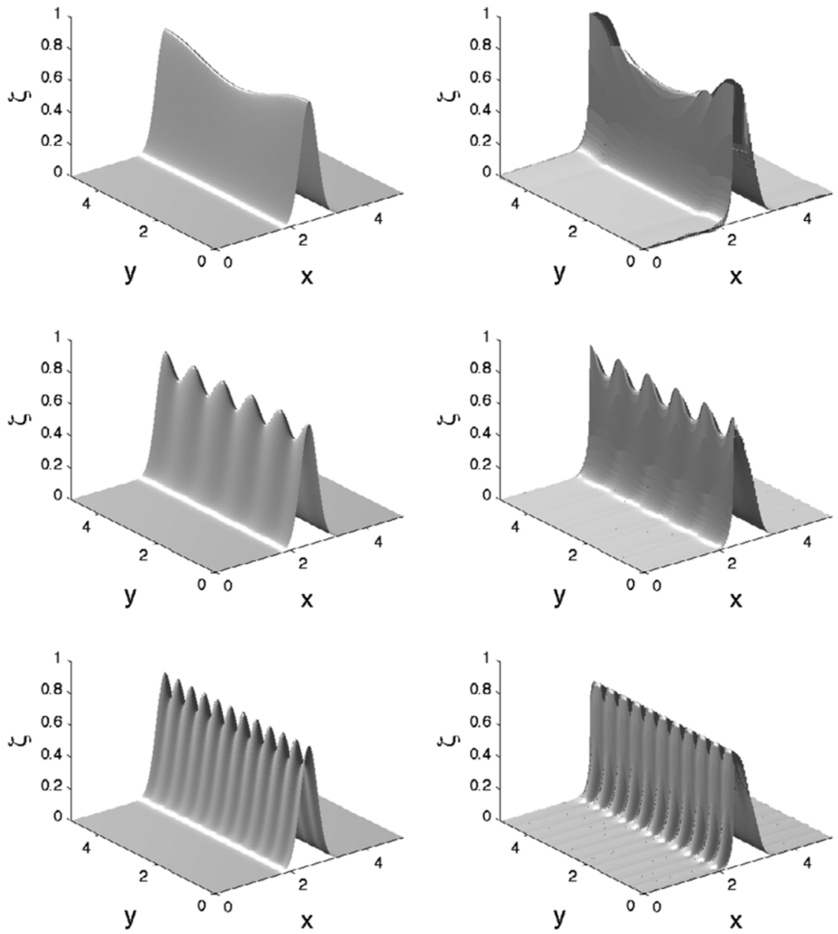

Figure 9. Two-dimensional profiles of ζ at $t = 0$ (left) and $t = 0.75$ (right). The parameters are $\delta = 1$, $\gamma = 0.9$, $c = 0.3$ and $\sigma = 0.1$. The top, middle and bottom panels correspond to $n_2 = 1$, $n_2 = 5$ and $n_2 = 11$, respectively.

Two-dimensional profiles are obtained by repeating the initial conditions (67) and (68) in the y-direction, with $\boldsymbol{v}_1^0 = \boldsymbol{v}^0$ from (68) and $\boldsymbol{v}_2^0 = 0$. The following parameter values are specified: $\delta = 1$, $\gamma = 0.1$, $c = 1$, $L_1 = 5$, $N_1 = 512$, $\Delta t = 10^{-4}$, together with $L_2 = 0.1$ and $N_2 = 8$ in the two-dimensional case. Since the solution is invariant in the transverse direction, we do not need to specify a large value of L_2 and, consequently, a small number of grid points ($N_2 = 8$) is sufficient to ensure a fine resolution in y.

Figure 8 shows the relative L^2 (squared norm) errors on ζ and $\boldsymbol{v}_1$ between one- and two-dimensional results, evaluated at $t = 0.58$, for $N_t = 2, 4, 8, 10, 12$. The so-obtained errors are not ideal in the sense that they include truncation errors from the spatial and temporal discretizations (of both one- and two-dimensional computations), from the approximation of the Neumann series of $\mathfrak{R}[\zeta]$, as well as their accumulation in time but, nevertheless, they do give an idea of how well the numerical solution is resolved as a function of N_t. Our results confirm that these errors are small and decrease as N_t increases, meaning that a better approximation is achieved as more terms are added in the Neumann series of $\mathfrak{R}[\zeta]$. However, keep also in mind that the larger N_t, the more computation this requires. Therefore, in order to achieve reasonable computational times in our two-dimensional simulations, we have used a relatively small number of terms, $N_t = 4$, as a compromise between accuracy and speed.

Since we do not have the equivalent of corollary 1 to describe precisely the formation of singularities in the two-dimensional case, it is of particular interest to investigate numerically the behaviour of the solution in situations that are natural generalizations of those considered in the one-dimensional case. We thus consider the following situations:

- *Stability with respect to two-dimensional perturbations*. It is also of interest to examine the stability of one-dimensional solutions with respect to two-dimensional perturbations, as this gives an idea of the differences between the one- and two-dimensional situations. For this purpose, we impose the following initial conditions:

$$\boldsymbol{v}_1^0 = c\,\mathrm{e}^{-10(x-L_1/2)^2}[1+\sigma\cos(k_2 y)], \qquad \boldsymbol{v}_2^0 = 0, \qquad \zeta_1^0 = c_g \boldsymbol{v}_1^0,$$

  where σ and $k_2 = \frac{2\pi n_2}{L_2}$ denote the amplitude and wavenumber of the transverse perturbation, respectively. For $\sigma = 0$, the one-dimensional solution typically forms a shock.

  Setting $\delta = 1$, $\gamma = 0.9$, $c = 0.3$, $\sigma = 0.1$, $L_1 = L_2 = 5$, $N_1 = 256$, $N_2 = 128$ and $\Delta t = 10^{-3}$, figure 9 shows the two-dimensional profiles of ζ at $t = 0$ and $t = 0.75$ (when the shock occurs), for three different perturbation wavenumbers $n_2 = 1, 5$ and 11. These two-dimensional profiles are compared with the one-dimensional solution in figure 10, for $t = 0.134$, 0.434, 0.584 and 0.750. As expected, in the three cases, the transverse perturbations grow in time; the larger the perturbation wavenumber, the faster the growth as suggested by figure 10. However, overall, the longitudinal dynamics prevail and the solution always evolves towards a shock in the x-direction.
- *Vanishing depth for the upper fluid* (we do not present here the numerical computations corresponding to the vanishing depth in the lower fluid because they do not bring any significant new information). To observe this singularity in the two-dimensional setting, we specify initial conditions similar to (67) and (68):

$$\begin{aligned}\zeta^0(x,y) = \frac{7}{2}c_g\frac{\sqrt{\alpha\beta}}{\eta}&\left[\frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(x-\frac{L_1}{2}+0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}} + \frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(x-\frac{L_1}{2}-0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}}\right]\\ &\times\left[\frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(y-\frac{L_2}{2}+0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}} + \frac{\left(\frac{c^2}{\alpha\beta}-1\right)}{\cosh\left[4\left(y-\frac{L_2}{2}-0.35\right)\right]+\frac{c}{\sqrt{\alpha\beta}}}\right],\end{aligned} \tag{69}$$

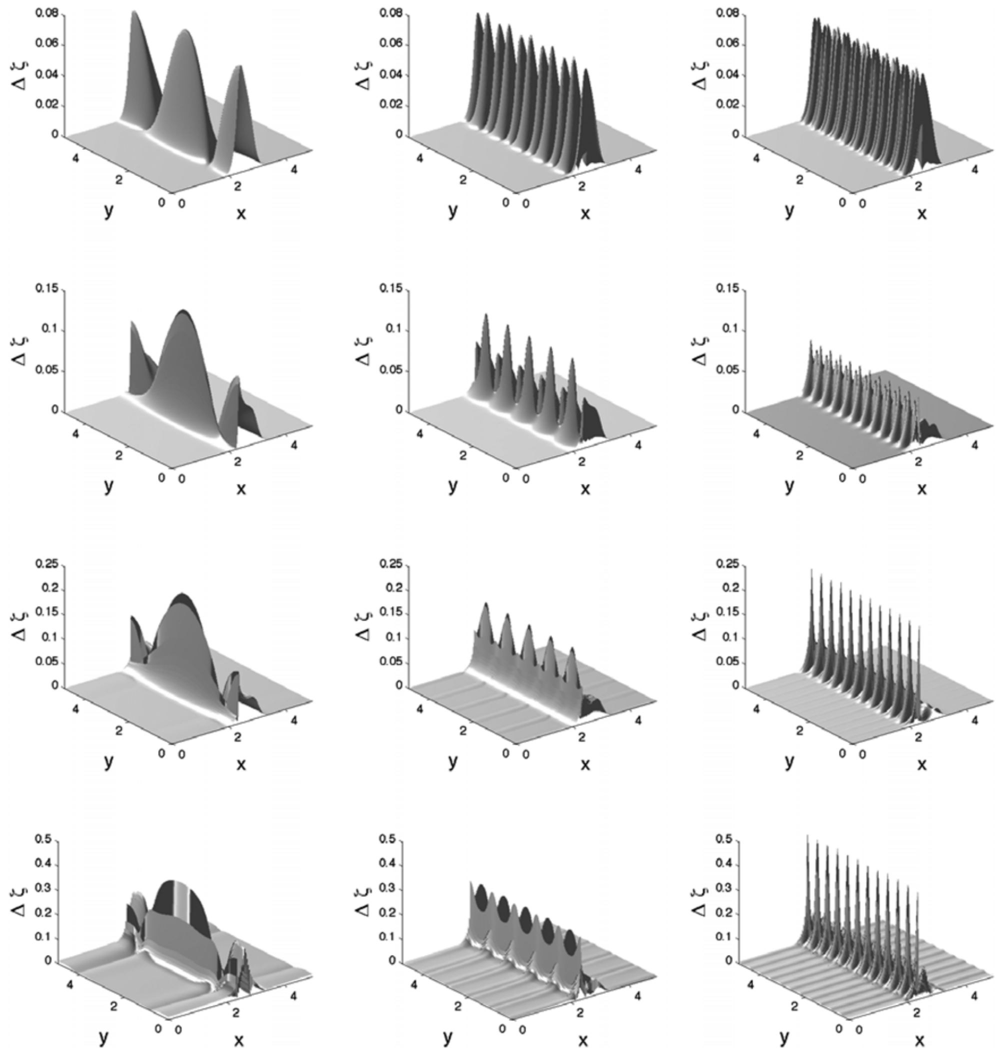

Figure 10. Absolute difference between one- and two-dimensional profiles of ζ at $t = 0.134$, 0.434, 0.584 and 0.750 (from top to bottom). The parameters are $\delta = 1$, $\gamma = 0.9$, $c = 0.3$ and $\sigma = 0.1$. The left, middle and right columns correspond to $n_2 = 1$, $n_2 = 5$ and $n_2 = 11$, respectively.

$$v_1^0(x, y) = \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}} - \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(x - \dfrac{L_1}{2} - 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}, \tag{70}$$

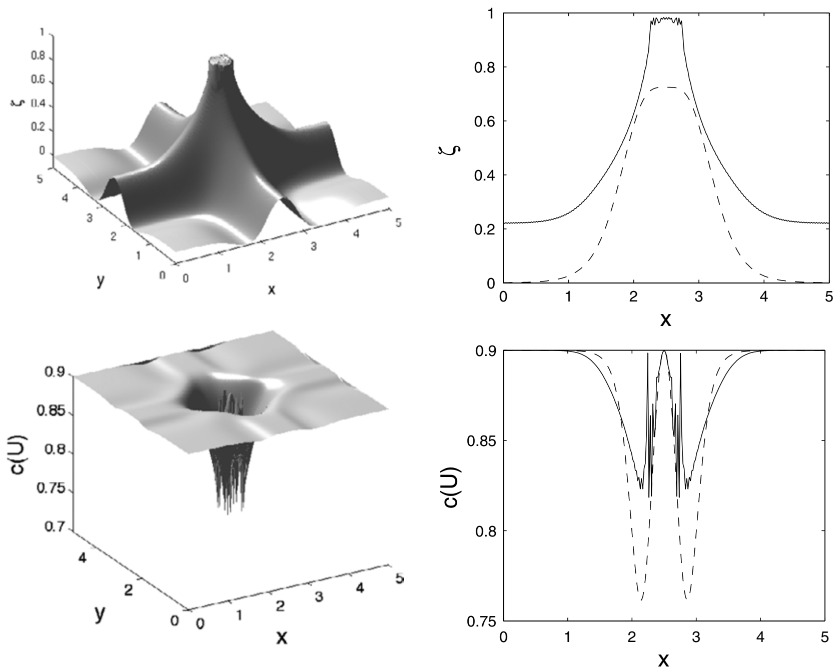

Figure 11. Left: profiles of ζ and $c(U)$ at $t = 0.45$. Right: cross-sections at $y = 2.5$ for $t = 0$ (- - - -) and $t = 0.45$ (——). The parameters are $\delta = 1$, $\gamma = 0.1$ and $c = 1.1$.

$$v_2^0(x, y) = \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(y - \dfrac{L_2}{2} + 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}} - \frac{\sqrt{\alpha\beta}}{\eta} \frac{\left(\dfrac{c^2}{\alpha\beta} - 1\right)}{\cosh\left[4\left(y - \dfrac{L_2}{2} - 0.35\right)\right] + \dfrac{c}{\sqrt{\alpha\beta}}}, \tag{71}$$

with $\delta = 1$, $\gamma = 0.1$, $c = 1.1$, $L_1 = L_2 = 5$, $N_1 = N_2 = 256$ and $\Delta t = 10^{-3}$. This corresponds to two superimposed waves whose crests are perpendicular to each other (along the x- and y-axes), and which intersect in the center of the domain. Such a superposition promotes wave focusing and will quickly lead to vanishing of the upper fluid depth.

Figure 11 plots the profile of ζ at $t = 0.45$ when it has almost reached the rigid lid ($z = 1$) in the center of the domain; $c(U)$, as given by the left-hand side of the third condition in (44), remains positive throughout the computation. Here again, instabilities that have started developing in the profiles at $t = 0.45$ will rapidly grow and eventually spoil the solution.

We point out that, since we are dealing with a generalization of the one-dimensional case, it is also observed that v_1 (respectively v_2) tends to form a positive shock in the x-direction (respectively in the y-direction), as discussed previously.

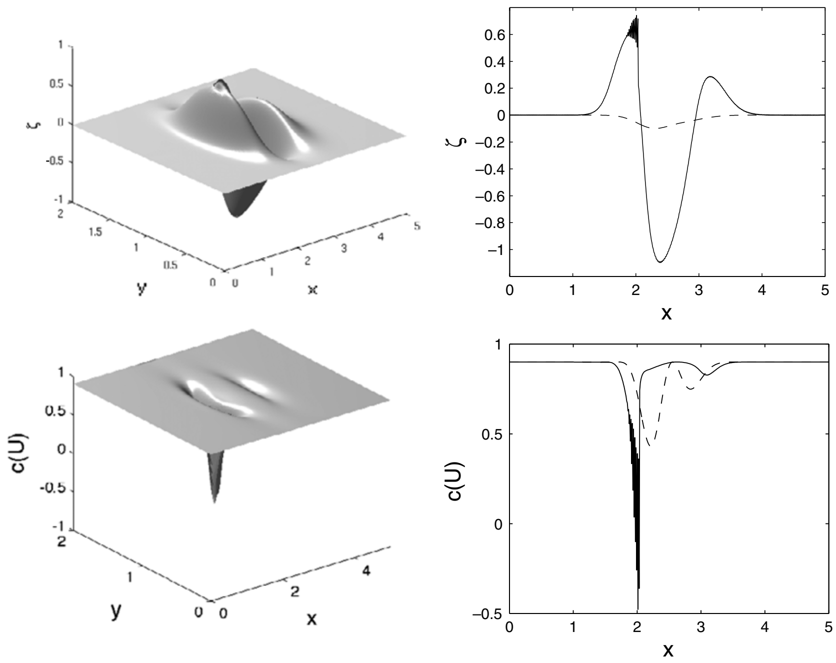

Figure 12. Left: profiles of ζ and $c(U)$ at $t = 0.16$. Right: cross-sections at $y = 1$ for $t = 0$ (- - - -) and $t = 0.16$ (——). The parameters are $\delta = 1/3$, $\gamma = 0.1$ and $c = -0.085$.

By changing the signs in the initial conditions (69)–(71), we observe a similar behaviour but, this time, with ζ sinking down to the bottom (vanishing depth for the lower fluid). For convenience, the results are not shown here.

– *Vanishing of the coefficient* $c(U)$. With initial conditions

$$\zeta^0(x, y) = c[\mathrm{e}^{-8(x-0.45L_1)^2} + \tfrac{1}{2}\mathrm{e}^{-5(x-0.55L_1)^2}]\mathrm{e}^{-8(y-L_2/2)^2},$$

$$v_1^0(x, y) = 10\, c_g\, c[\mathrm{e}^{-8(x-0.45L_1)^2} - \tfrac{1}{2}\mathrm{e}^{-5(x-0.55L_1)^2}]\mathrm{e}^{-8(y-L_2/2)^2},$$

$$v_2^0(x, y) = 0,$$

and parameter values $\delta = 1/3$, $\gamma = 0.1$, $c = -0.085$, $L_1 = 5$, $L_2 = 2$, $N_1 = 512$, $N_2 = 256$ and $\Delta t = 10^{-4}$, the solution is depicted in figure 12. As in the one-dimensional case, it can be checked that instabilities appear shortly after $c(U)$ takes negative values, in accordance with (44).

We would like to emphasize again that the goal of the present numerical simulations is more to illustrate the various scenarios identified analytically than to provide a careful examination of results. Although the numerical methods used here are globally very accurate and allow for an efficient evaluation of the nonlocal operator $\mathfrak{R}[\zeta]$ in the two-dimensional case, they cannot accurately resolve the full development of shock waves. It would be of interest to use shock-capturing numerical schemes such as in [6] to further compare with the analytical predictions, in particular for the one-dimensional system (5).

Acknowledgments

PG acknowledges support from the National Science Foundation through grants Nos DMS-0625931 and DMS-0920850. DL acknowledges support from the ANR-08-BLAN-0301-01 project and J-CS from the project ANR-07-BLAN-0250 of the Agence Nationale de la Recherche.

References

[1] Alinhac S 1995 Blowup for nonlinear hyperbolic equations *Progress in Nonlinear Equations and Their Applications* vol 17 (Boston: Birkhäuser)

[2] Alvarez-Samaniego B and Lannes D 2008 A Nash–Moser theorem for singular evolution equations. Application to the Serre and Green–Naghdi equations *Indiana Univ. Math. J.* **57** 97–131

[3] Benjamin T B 1966 Internal waves of finite amplitude and permanent form *J. Fluid Mech.* **25** 241–70

[4] Benjamin T B and Bridges T J 1997 Reappraisal of the Kelvin–Helmholtz problem: I. Hamiltonian structure *J. Fluid Mech.* **333** 301–25

[5] Bona J, Lannes D and Saut J-C 2008 Asymptotic models for internal waves *J. Math. Pures Appl.* **89** 538–66

[6] Bouchut F and Zeitlin V 2009 A robust well-balanced scheme for multi-layer shallow water equations *DCDC-S* to appear

[7] Camassa R, Holm D D and Levermore C D 1997 Long-time shallow-water equations with a varying bottom *J. Fluid Mech.* **349** 173–89

[8] Chandrasekhar S 1961 *Hydrodynamic and Hydromagnetic Stability* (Oxford: Clarendon)

[9] Choi W and Camassa R 1999 Fully nonlinear internal waves in a two-fluid system *J. Fluid Mech.* **396** 1–36

[10] Chumakova L, Menzaque F E, Milewski P A, Tabak E G and Turner C V 2009 Stability properties and nonlinear mappings of two and three-layer stratified flows *Stud. Appl. Math.* **122** 123–37

[11] Craig W and Groves M D 2000 Normal forms for wave motion in fluid interfaces *Wave Motion* **31** 21–41

[12] Craig W, Guyenne P and Kalisch H 2005 Hamiltonian long-wave expansions for free surfaces and interfaces *Commun. Pure. Appl. Math.* **58** 1587–641

[13] Craig W, Sulem C and Sulem P-L 1992 Nonlinear modulation of gravity waves: a rigorous approach *Nonlinearity* **5** 497–522

[14] Dafermos C 2005 *Hyperbolic Conservation laws in Continuum Physics* 2nd edn (Heidelberg: Springer)

[15] Duchêne V 2009 Asymptotic shallow water models for internal waves in a two-fluid system with a free surface arXiv:0906.0839

[16] Grisvard P 1974–1975 Quelques propriétés des espaces de Sobolev, utiles dans l'étude des équations de Navier–Stokes Exposé 4 *Problèmes d'Évolution Non Linéaires*, *Séminaire de Nice 1974–1975*

[17] Guyenne P 2006 Large-amplitude internal solitary waves in a two-fluid model *C. R. Mecanique* **334** 341–6

[18] Hörmander L 1997 *Lectures on Nonlinear Hyperbolic Differential Equations* (*Mathématiques Applications* vol 26) (Berlin: Springer)

[19] Hou T Y and Zhang P 1998 Growth rates for the linearized motion of 3-D fluid interfaces with surface tension far from equilibrium *Asian J. Math.* **2** 263–88

[20] Israwi S 2009 Large time existence for 1D Green–Naghdi equations (arXiv:0909.2232) submitted

[21] John F 1974 Formation of singularities in one-dimensional nonlinear wave propagation *Commun. Pure Appl. Math.* **27** 377–405

[22] Lannes D 2006 Sharp estimates for pseudo-differential operators with symbols of limited smoothness and commutators *J. Funct. Anal.* **232** 495–539

[23] Li Y A 2006 A shallow-water approximation to the full water wave problem *Commun. Pure Appl. Math.* **59** 1225–85

[24] Liu T-P 1979 Development of singularities in the nonlinear waves for quasilinear hyperbolic partial differential equations *J. Diff. Eqns* **33** 92–111

[25] Majda A 1984 *Incompressible Fluid Flow and Systems of Conservation Laws in Several Space Variables* (Heidelberg: Springer)

[26] Milewski P, Tabak E, Turner C, Rosales R and Menzaque F 2004 Nonlinear stability of two-layer flows *Commun. Math. Sci.* **2** 427–42

[27] Smoller J 1983 *Shock Waves and Reaction–Diffusion Equations* (Heidelberg: Springer)

[28] Taylor M E 1997 *Partial Differential Equations: III. Nonlinear Equations* (*Applied Mathematical Sciences* vol 117) (New York: Springer) (Corrected reprint of the 1996 original)

[29] Zakharov V E 1968 Stability of periodic waves of finite amplitude on the surface of a deep fluid *J. Appl. Mech. Tech. Phys.* **2** 190–4