Modélisation, analyse mathématique et simulations numériques de quelques problèmes aux dérivées partielles multi-échelles

Amélie Rambaud

Thèse de doctorat N° d'ordre :
Université Claude Bernard Lyon 1
Institut Camille Jordan
UMR 5208 CNRS-UCBL

Modélisation, analyse mathématique et simulations numériques de quelques problèmes aux dérivées partielles multi-échelles

---

THÈSE

présentée et soutenue publiquement le

pour l'obtention du diplôme de

Doctorat de L'Université Claude Bernard Lyon1
(Spécialité Mathématiques)

par

Amélie Rambaud

Sous la direction de : Francis FILBET et Pascal NOBLE

Devant le jury composé de :

Sylvie BENZONI-GAVAGE	Examinatrice
Francis FILBET	Directeur de thèse
Thierry GOUDON	Rapporteur
Fréderic LAGOUTIERE	Examinateur
Roberto NATALINI	Rapporteur
Pascal NOBLE	Directeur de thèse
Jean-Paul VILA	Examinateur

Résumé

Nous étudions dans cette thèse plusieurs aspects d'équations aux dérivées partielles multi-échelles. Pour trois exemples distincts, la présence de multiples échelles, spatiales ou temporelles, motive un travail de modélisation mathématique ou constitue un enjeu de discrétisation.
La première partie est consacrée à la construction et l'étude d'un système multicouche de type Saint-Venant pour décrire un fluide à surface libre (océan). Son obtention s'appuie sur une analyse des échelles spatiales mises en jeu, en particulier sur l'hypothèse dite "eau peu profonde" , classiquement utilisée dans le cas des fluides géophysiques. Nous justifions donc nos équations, et montrons un résultat d'existence locale de solution. Puis nous proposons un schéma volumes finis et des simulations numériques en vue de valider notre modèle.
Dans la deuxième partie, nous étudions un problème hyperbolique de relaxation, inspiré de la théorie cinétique des gaz. La différence entre l'échelle temporelle du mécanisme de transport et celui de la relaxation constitue un enjeu numérique crucial. Nous construisons donc un schéma numérique *via* une stratégie "préservant l'asymptotique" : nous montrons sa convergence pour toute valeur du paramètre de relaxation, ainsi que sa consistance avec le problème à l'équilibre local. Des estimations d'erreurs sont établies et des simulations numériques sont présentées.
La dernière partie traite un problème d'écoulement sanguin dans une artère avec stent, modélisé par un système de Stokes dans un domaine contenant une petite rugosité périodique, *i.e.* une géométrie double échelle. Pour éviter une discrétisation coûteuse du domaine rugueux (l'artère stentée), nous formulons un ansatz de développement de la solution type Chapman-Enskog, et obtenons une loi de paroi implicite sur le bord du domaine lisse (l'artère seule). Nous montrons alors des estimations d'erreurs et des simulations numériques.

Mots clés : analyse d'échelles, modèle multicouche de Saint-Venant, systèmes hyperboliques, lois de conservation, volumes finis, relaxation, terme source raide, flux à variation totale décroissante, schéma préservant l'asymptotique, couche limite, loi de paroi.

Abstract

This work is concerned with different aspects of multiscale partial differential equations. For three distinct problems, we address questions of modelling and discretization thanks to the observation of the multiplicity of scales, time or space.
So we propose in the first part a model of approximation of a fluid with a free surface, say an ocean. The derivation of our multilayer shallow water type model is based on an analysis of the different space scales generally observed in geophysical flows, in particular the so-called *shallow water* assumption. We obtain an existence and uniqueness result of local in time solution. Next we propose a finite volume scheme and numerical simulations in order to validate our model.
In the second part, we study a hyperbolic relaxation problem, initially motivated by the kinetic theory of gaz. Different time scales appear through the competition between a transport phenomenon and a relaxation one, to a local equilibrium. Adopting an *Asymptotic Preserving* strategy of discretization, we build and analyze a numerical scheme. The convergence is proved for any value of the relaxation parameter, as well as the consistency with the equilibrium problem, thanks to error estimates. Then we present some numerical simulations.
The last part deals with a blood flow model in a stented artery. We consider a Stokes problem stated in a multiscale space domain, that is a macroscopic box (the artery) containing a microscopic roughness (the stent). In order to avoid expensive simulations when discretizing the whole rough domain, we perform a Chapman-Enskog type development of the solution and derive an implicit wall law on the boundary of the smooth domain. Error estimates are shown and numerical illustrations of the results are presented.

Key words: multilayer shallow water model, hyperbolic systems, conservation laws, finite volumes, relaxation, source term, stiff, total variation diminishing, asymptotic preserving schemes, boundary layer, wall-laws.

Table des matières

Avant-propos

Si les trois problématiques adressées dans ce travail (modélisation mathématique, étude théorique, analyse et simulations numériques) constituent séparément un travail en soi, elles n'en sont pas moins intimement liées lorsque l'on souhaite comprendre un phénomène naturel.

La *modélisation mathématique* consiste à traduire un phénomène réel (donc complexe) à l'aide d'outils mathématiques. L'objectif de cet exercice scientifique est double : il s'agit d'une part d'obtenir un modèle qui *décrive* autant que possible la réalité et d'autre part, de pouvoir faire des *prévisions*, par exemple météorologiques dans le cas d'un modèle d'atmosphère. Evidemment, ce travail n'a rien de systématique et repose sur une succession d'observations et de simplifications. Le mathématicien doit faire des concessions afin de satisfaire ses deux ambitions. En effet, plus le modèle mathématique prend en compte de paramètres physiques, plus il offre une *description* proche de la réalité, mais plus il est complexe et de ce fait son étude théorique et sa mise en oeuvre *prédictive* (numérique) en deviennent plus difficiles.
Une fois le modèle mathématique établi, il reste de nombreuses questions auxquelles il faut tenter de répondre. Peut-on le *justifier*, formellement ou rigoureusement ? Possède-t-il une ou plusieurs solutions ? Cette ou ces solutions fournissent-elles un bonne description de la physique observée ? Peut-on les approcher en discrétisant le problème ? Comment ? Les résultats obtenus sont ils conformes à la réalité ? Peuvent-il être utilisés de manière prédictive ? ...
Cette liste de questions, loin d'être exhaustive, met déjà en lumière les multiples difficultés mathématiques auxquelles nous devons nous confronter pour comprendre un phénomène issu de la physique ou de la biologie. C'est à travers trois exemples distincts (un modèle de fluide à surface libre, un système hyperbolique de relaxation et un modèle d'écoulement sanguin) que nous tentons ici d'apporter quelques éléments de réponses sur les aspects modélisation, analyse théorique et simulations numériques. Les trois exemples abordés s'inscrivent dans des contextes très différents, mais sont néanmoins liés par une caractéristique commune : ils contiennent de *multiples échelles*, d'espace ou de temps, et ce sont précisément ces différentes échelles qui permettent et motivent les travaux effectués au cours de la thèse.

La première partie est ainsi consacrée à la modélisation d'un fluide géophysique à surface libre (typiquement un océan) pour lequel ce sont des échelles spatiales caractéristiques qui sont très différentes. En effet, en s'appuyant sur des observations physiques, nous pouvons mettre en évidence un nombre sans dimension $\varepsilon > 0$ très *petit*, à savoir le rapport entre deux grandeurs caractéristiques du problème : la profondeur typique du fluide et la longueur d'onde moyenne des mouvements horizontaux. Grâce à cette hypothèse, dite de *shallow water*, nous dérivons un nouveau modèle multicouche de type Saint-Venant à partir des équations primitives, que nous étudions ensuite [165].
Dans la deuxième partie de ce manuscrit, en collaboration avec F. Filbet [94], nous nous intéressons à un problème hyperbolique de relaxation. Motivés par la théorie cinétique des gaz, nous présentons ici un modèle jouet dans lequel les diverses échelles s'affrontant sont des échelles de temps : un phénomène de transport est accompagné d'un mécanisme de retour vers un équilibre local matérialisé par un terme source dans les équations. La vitesse de ce mécanisme de *relaxation*, représentée dans les équations par un coefficient sans dimension $\dfrac{1}{\varepsilon} > 0$, peut être très rapide ($\varepsilon \to 0$). Le terme source devient alors *raide* et constitue un enjeu important dans son traitement numérique : c'est la préoccupation majeure de notre travail, effectué dans le cadre des schémas *préservant l'asymptotique* (*Asymptotic Preserving*).
Enfin, la troisième partie est issue d'un travail en collaboration avec V. Milišić et K. Pichon Gostaf [150], initié au CEMRACS 2009 et concerne un modèle d'écoulement sanguin dans des artères avec stents. Dans ce cas, les différentes échelles apparaissent dans le domaine géométrique sur lequel sont posées les équations. En effet, si l'on symbolise grossièrement l'artère par un cylindre droit Ω, le stent forme alors une *rugosité* périodique et de petite taille $\varepsilon > 0$ au bord du domaine lisse : les équations sont donc posées dans un domaine rugueux Ω_ε, dont le maillage direct est peu envisageable car très coûteux si l'on souhaite rendre compte de l'influence effective du stent sur l'écoulement du sang dans l'artère. L'objectif est donc de surmonter cette difficulté numérique en modifiant les équations afin d'obtenir un système posé dans le domaine lisse, les informations de la rugosité étant contenues dans une *loi de paroi* implicite.

Chapitre 1

Introduction générale et présentation des travaux

Dans cette introduction générale, nous motivons les travaux effectués au cours de la thèse. Les trois parties s'inscrivent dans des contextes de recherche très différents et très riches. C'est pourquoi nous établissons d'abord un bref et non exhaustif état de l'art de chaque partie, avant de présenter les résultats obtenus en les confrontant (autant que possible) aux recherches actuelles.

1 Partie I : modélisation de fluides géophysiques à surface libre

Cette section vise à motiver et décrire les travaux de la Partie I, qui sont réunis au sein des Chapitres 2 et 3 [165]. Nous commençons par présenter plusieurs modèles classiques de fluides à surface libre (typiquement les océans) : de Navier-Stokes à Saint-Venant, en passant par les équations primitives. Nous évoquons les liens qui les unissent, leur justification mathématique (méthodes d'analyse dimensionnelle), les résultats d'existence de solutions, ainsi que leur traitement numérique. Puis nous situons notre nouveau modèle multicouche de type Saint-Venant au sein de cette hiérarchie et le comparons aux autres modèles multicouches existants. Enfin, nous résumons les résultats obtenus sur ce système, à savoir sa dérivation à partir des équations primitives et un théorème d'existence de solution forte locale (Chapitre 2), une étude de l'énergie du système (Annexe A), ainsi que la construction d'un schéma volumes finis et des simulations numériques (Chapitre 3).

1.1 Une hiérarchie de modèles

Lorsque l'on adresse la question de la description d'un fluide, plusieurs approches sont possibles. A l'échelle microscopique, on considère des « particules » dont on suit les trajectoires au cours du temps (vision Lagrangienne). A l'opposé, on peut adopter une vision macroscopique (Eulerienne) et considérer l'évolution de quantités hydrodynamiques telles la densité, la vitesse, ou la pression. C'est ce niveau d'observation que nous choisissons

dans la Partie I pour étudier un fluide géophysique tel que l'eau d'un océan, d'une mer, ou d'un fleuve. Nous aborderons une autre échelle de description à la Partie II, plus adéquate pour traiter les gaz raréfiés.
Concernant la description macroscopique des fluides géophysiques, il existe une littérature très riche. Citons par exemple les ouvrages classiques de P.L. Lions [137], de R. Lewandowski [136], ou de J. Pedlosky [159]. Précisons les hypothèses que nous faisons dans le présent travail, tant sur les caractéristiques du fluide que sur celles de l'écoulement. Nous considérons ici un fluide à surface libre :

- incompressible (son volume reste inchangé sous l'action d'une pression externe),
- visqueux de viscosité μ et newtonien (son taux de déformation est proportionnel aux forces de cisaillements appliquées, et nous pouvons en définir un tenseur des contraintes visqueuses Σ_μ, défini ci-après).

Donnons-nous un repère local à la surface de la terre, de direction verticale z supportant l'axe de rotation du fluide, et de direction horizontale $\mathbf{x} \in \mathbb{R}^{2\,(1)}$. Nous faisons l'hypothèse que la latitude reste constante ; ainsi le repère local devient un repère cartésien fixe et les forces s'appliquant au fluide sont :

- la force de gravité $\mathbf{g}$, parallèle à l'axe vertical,
- la « force » de Coriolis $-2\,\Omega \times \mathbf{U}$,

où Ω est dirigé suivant la direction z, l'axe de rotation. On suppose en effet ici que l'angle de Coriolis est constant égal à $\pi/2$ [159]. Par ailleurs $\mathbf{U} = (\mathbf{u}, w)^T \in \mathbb{R}^3$ désigne la vitesse du fluide, décomposée en vitesses horizontale et verticale, de sorte que la force de Coriolis s'écrive :

$$-2\,\Omega \times \mathbf{U} \;=\; \Big(-f\,\mathbf{u}^{\perp};\, 0\Big)^T \;,$$

avec $f > 0$ le paramètre de Coriolis, constant. On désigne également par p la pression et ρ la densité du fluide. L'équation de conservation de la masse s'écrit alors :

$$\partial_t \rho + \operatorname{div}(\rho\,\mathbf{U}) \;=\; 0\,. \tag{1.1}$$

Puis, en appliquant la deuxième loi de Newton, nous obtenons les équations de conservations de la quantité de mouvement :

$$\partial_t(\rho\,\mathbf{U}) + \operatorname{div}(\rho\,\mathbf{U}\otimes\mathbf{U}) \;=\; -2\,\Omega\times\mathbf{U} + \rho\,\mathbf{g} - \nabla p + \rho\,\operatorname{div}\Sigma_\mu\,, \tag{1.2}$$

où les deux derniers termes de forces désignent respectivement la force de flottabilité et la force de friction liée à la viscosité du fluide par le tenseur des contraintes :

$$\Sigma_\mu \;=\; \mu\,\Big(\nabla\otimes\mathbf{U} + (\nabla\otimes\mathbf{U})^T\Big)\,. \tag{1.3}$$

Tel quel, ce système n'est pas fermé. Nous devons donc ajouter des hypothèses et des conditions aux bords, voire des équations d'état sur la température, la salinité, *etc.* Une

[1] Pour la dimension $d = 2$, nous noterons x *la* variable horizontale.

approximation classique est celle de Boussinesq [(2)]. Ici, nous simplifions encore et supposons :

- le fluide est *homogène*, de densité $\rho \equiv 1$, de sorte que la conservation de la masse se réduit à la condition d'incompressibilité,
- nous ne considérons que les équations de conservation de la masse et de la quantité de mouvement.

Pour des équations plus générales, nous renvoyons à nouveau à [136, 137, 159]. Nous pouvons maintenant présenter le système qui sera le point de départ à nos approximations successives *via* l'adimensionnement.

Les équations de Navier-Stokes incompressibles à surface libre

Nous considérons un coefficient constant de viscosité μ (voir par exemple [75] pour une viscosité variable). Les équations sur la vitesse (u, w) et la pression p sont données, sous forme conservative par :

$$\left\{ \begin{array}{lcl} \mathrm{div}_{\mathbf{x}}\, \mathbf{u} + \partial_z w & = & 0\,, \\ \\ \partial_t \mathbf{U} + \mathrm{div}(\mathbf{U} \otimes \mathbf{U}) + \nabla p & = & -2\,\Omega \times \mathbf{U} - \mathbf{g} + \mathrm{div}\,\Sigma_\mu\,, \end{array} \right. \tag{1.4}$$

satisfaites pour

$$t > 0\,,\ \ (\mathbf{x}, z) \in \Omega_t = \left\{ (\mathbf{x}, z) \in \mathbb{R} \times \mathbb{R}^+ \,|\, z_b(\mathbf{x}) \leqslant z \leqslant \eta(t, \mathbf{x}) \right\}\,,$$

où z_b désigne la bathymétrie (supposée indépendante du temps) et $\eta(t, \mathbf{x})$ représente la surface libre. La hauteur du fluide est donc donnée par

$$H(t, \mathbf{x}) = \eta(t, \mathbf{x}) - z_b(\mathbf{x})\,.$$

On introduit le tenseur total des contraintes :

$$\Sigma_T = -p\, I_3 + \Sigma_\mu\,.$$

Ainsi, on peut compléter le système par des conditions aux bords. Notons $(\mathbf{u}_b, w_b)$ (resp. $(\mathbf{u}_s, w_s)$) la vitesse du fluide au fond (resp. à la surface libre) . De plus, $\mathbf{n}_s$ et $\mathbf{n}_b$ désignent respectivement les normales unitaires extérieure à la surface libre et intérieure au fond. Nous supposons d'une part la continuité des contraintes à la surface libre, traduite par

$$\Sigma_T\, \mathbf{n}_s = 0, \tag{1.5}$$

en considérant comme nulle la pression atmosphérique. D'autre part, nous imposons au fond la non pénétration, ainsi qu'une loi de type Navier, avec un coefficient de friction κ constant :

$$\left\{ \begin{array}{lcl} \mathbf{u}_b \cdot \nabla_{\mathbf{x}} z_b & = & w_b, \\ \\ \kappa\, \mathbf{u}_b & = & \mu\, \partial_z \mathbf{u}_b. \end{array} \right. \tag{1.6}$$

[2]Elle consiste à supposer la densité ρ constante dans l'équation sur la quantité de mouvement sauf dans le terme de pression $\nabla p/\rho$.

Le terme de friction considéré est simplement linéaire ; nous ne prenons pas en compte la friction turbulente, qui ajouterait un terme quadratique à ces équations (voir par exemple [159, 143]).
Même dans cette formulation assez simple du système de Navier-Stokes (1.4), l'étude mathématique reste complexe et les simulations numériques coûteuses, notamment en raison de la non-linéarité des équations et de la dépendance en temps du domaine spatial Ω_t. C'est pourquoi ingénieurs et mathématiciens ont établi toute une hiérarchie de systèmes simplifiés pour modéliser les fluides géophysiques, avec deux objectifs principaux :

- comprendre et *décrire* plus précisément les multiples dynamiques,
- et être capable de fournir des *prévisions* fiables de ces dynamiques.

La dérivation de modèles plus simples s'appuie sur l'analyse dimensionnelle, c'est-à-dire une étude des échelles typiques du problème. Nous introduisons donc des quantités caractéristiques :

- la profondeur caractéristique de l'océan H_0 et une longueur d'onde horizontale typique λ_0 (voir la Figure 1.1),
- les variations d'amplitudes typiques de la surface libre a_s et de la bathymétrie a_b (voir la Figure 1.1),
- des vitesses caractéristiques horizontales et verticales U et W.

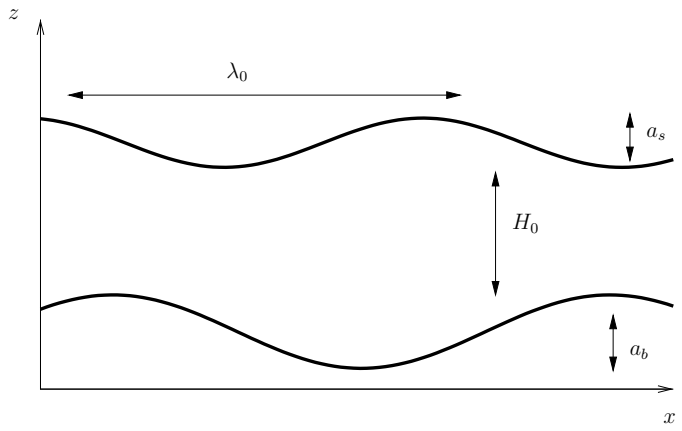


FIGURE 1.1 – Echelles spatiales caractéristiques.

En procédant à un adimensionnement des équations (voir par exemple [159] pour la description précise de cette analyse d'échelles), il apparaît en particulier deux nombres sans dimension, le nombre de Reynolds

$$\mathcal{R}e = \frac{U\,\lambda_0}{\mu},$$

et le rapport d'aspect

$$\varepsilon = \frac{H_0}{\lambda_0}\,.$$

Ainsi, en étudiant différentes asymptotiques de ces nombres [3], nous pouvons dériver une multitude de modèles de complexité réduite par rapport aux équations de Navier-Stokes. Pour arriver au modèle de la Partie I, nous choisissons un régime d'écoulement à $\mathcal{R}e$ intermédiaire, et nous intéressons plutôt à l'asymptotique

$$\varepsilon \ll 1\,, \tag{1.7}$$

appelée hypothèse *shallow water* (eau peu profonde). Cette hypothèse nous conduira à deux systèmes aujourd'hui largement validés mathématiquement et expérimentalement : les équations primitives, très souvent utilisées pour décrire l'atmosphère ou les océans, et les équations de Saint-Venant, particulièrement bien adaptées à la simulation de ruptures de barrage (en hydraulique) et à l'océanographie côtière. Notre nouveau modèle se situera entre les deux.

Approximation hydrostatique

Cette simplification classique des équations de Navier-Stokes, aussi appelée *modèle des équations primitives*, est historiquement due à des observations physiques. En effet, dans son *Traité de l'équilibre des liqueurs* (paru à titre posthume en 1663), Blaise Pascal énonçait une loi de *pression hydrostatique* pour les liquides : la pression du fluide décroît linéairement avec l'altitude. Cette hypothèse, aujourd'hui validée par les ingénieurs et les mathématiciens, se justifie grâce à une observation du rapport d'aspect ε : il satisfait l'hypothèse *shallow water* (1.7) [4]. Cette condition, avec l'incompressibilité, conduit à négliger dans les équations les termes d'ordres supérieurs à 1 en ε. En particulier, dans la conservation de la quantité de mouvement verticale, tous les termes sont négligés, le gradient de pression et la gravité. On obtient ainsi, après retour aux variables avec dimensions, l'approximation hydrostatique, aussi appelée système des *équations primitives* : pour $t > 0$ et $(\mathbf{x}, z) \in \Omega_t$,

$$\left\{\begin{array}{lcl} \mathrm{div}_{\mathbf{x}}\,\mathbf{u} + \partial_z w & = & 0\,, \\ \partial_t \mathbf{u} + \mathrm{div}_{\mathbf{x}}(\mathbf{u}\otimes\mathbf{u}) + \partial_z(w\,\mathbf{u}) + \nabla_{\mathbf{x}} p & = & -f\,\mathbf{u}^{\perp} + \mu\,\Delta\mathbf{u}\,, \\ \partial_z p & = & -g\,, \end{array}\right. \tag{1.8}$$

complété par des conditions aux bords. A la surface libre, l'advection de la surface libre et la continuité du tenseur des contraintes se récrivent :

$$\left\{\begin{array}{lcl} \partial_t \eta + \mathbf{u}_s\cdot\nabla_{\mathbf{x}}\eta & = & w_s\,, \\ \partial_z \mathbf{u}_s & = & \nabla_{\mathbf{x}}\mathbf{u}_s\cdot\nabla_{\mathbf{x}}\eta\,, \end{array}\right. \tag{1.9}$$

[3] Il en existe d'autres [159].

[4] Pour les fluides géophysiques, il est naturel de supposer les échelles verticales petites par rapport aux échelles horizontales : par exemple, la profondeur typique d'un océan est de quelques km, tandis qu'il peut s'étendre sur plusieurs milliers de km.

tandis que les conditions au fond, non pénétration et loi de paroi de type Navier [41] (avec coefficient de friction laminaire κ) sont toujours données par :

$$\left\{ \begin{array}{rcl} \mathbf{u}_b \cdot \nabla_{\mathbf{x}} z_b & = & w_b, \\ \\ \kappa\, \mathbf{u}_b & = & \mu\, \partial_z \mathbf{u}_b. \end{array} \right. \tag{1.10}$$

Des siècles après Blaise Pascal, on attribue à L. F. Richardson en 1922 dans [166] l'introduction des équations primitives de l'atmosphère, modèle que l'auteur établit avec l'ambition de fournir des prévisions météorologiques. Mais les premiers ordinateurs du milieu du $20^{\text{ème}}$ siècle n'avaient pas la puissance de calcul actuelle ; les équations primitives furent donc un temps mises de côté au profit de l'étude des modèles, plus simples, géostrophique et quasi-géostrophique (voir par exemple [159]). Les équations primitives reviennent au goût du jour avec l'amélioration des ordinateurs, dans la dernière partie du XXème siècle.

Concernant leur justification mathématique, basée sur des développements asymptotiques des équations de Navier-Stokes adimensionnées lorsque ε tend vers 0 (voir (1.7)), nous pouvons citer quelques travaux. Par exemple, les articles pionniers de J.-L. Lions, R. Temam et S. Wang [138, 139, 140] établissent formellement les équations primitives (ils étudient également la limite géostrophique) entre 1992 et 1995. Cette dérivation formelle est également décrite dans les ouvrages précédemment cités [136, 137, 159]. Par ailleurs, P. Azerad et F. Guillén [19, 20] prouvent rigoureusement entre 1999 et 2001 la validité de l'approximation hydrostatique pour les océans sous l'hypothèse d'une viscosité anisotropique et des conditions de Dirichlet homogènes au fond.

Remarque 1. Sur les études théoriques des équations primitives, nous renvoyons le lecteur une fois de plus aux articles [138, 139], où les auteurs établissent les premiers résultats d'existence globale de solutions faibles. Enfin, une revue précise des résultats d'existences pour les équations primitives (et d'autres modèles de fluides géophysiques) est réunie dans l'article de R. Temam et M. Ziane paru en 2004 [176].

Enfin, les équations primitives sont largement utilisées en météorologie et océanographie : elles interviennent dans plusieurs codes opérationnels aujourd'hui. Néanmoins, les simulations des équations primitives sont relativement coûteuses : deux difficultés numériques de Navier-Stokes, à savoir nonlinéarité et domaine spatial dépendant du temps, sont toujours présentes. C'est pourquoi une autre famille de modèles d'approximation des équations de Navier-Stokes a également connu un fort succès dès la fin des années 1970 : les modèles de Saint-Venant (ou *shallow water*). La force principale de ces systèmes est leur efficacité numérique, due essentiellement aux deux raisons suivantes :

- leur structure (partiellement) *hyperbolique*, que l'on précisera ultérieurement,
- la réduction manifeste de complexité numérique par rapport aux équations de Navier-Stokes : le système est posé dans un domaine spatial *fixe* (et non plus variable) et sa dimension est abaissée de un.

Modèles classiques de Saint-Venant

Le système de Saint-Venant homogène (sans terme source) unidimensionnel est introduit grâce à des observations physiques par A.J.C. Barré de Saint-Venant en 1871 [167]. Mais ce n'est que dans la deuxième moitié du XXème siècle que les mathématiciens ont étudié les liens entre ces équations et les autres modèles hydrodynamiques. Les modèles de Saint-Venant proviennent essentiellement d'une intégration dans la direction verticale des équations de Navier-Stokes, et décrivent l'évolution de la hauteur totale du fluide $H(t,x)$ et de la moyenne sur la colonne d'eau de la vitesse horizontale

$$\overline{\mathbf{U}}(t,\mathbf{x}) \;=\; \frac{1}{H(t,x)} \int_{z_b(\mathbf{x})}^{\eta(t,\mathbf{x})} \mathbf{u}(t,\mathbf{x},z)\,\mathrm{d}z\,.$$

Dans notre cas, la présence d'une bathymétrie non triviale et d'un terme de friction linéaire fournit la formulation suivante :

$$\left\{\begin{array}{lcl} \partial_t H + \mathrm{div}_{\mathbf{x}}(H\,\overline{\mathbf{U}}) & = & 0\,, \\ \\ \partial_t(H\,\overline{\mathbf{U}}) + \mathrm{div}_{\mathbf{x}}\left(H\,\overline{\mathbf{U}}\otimes\overline{\mathbf{U}}\right) + g\,H\,\nabla_{\mathbf{x}} H & = & -g\,H\,\nabla_{\mathbf{x}} z_b - \kappa\,\overline{\mathbf{U}}\,. \end{array}\right. \tag{1.11}$$

Justification des équations de Saint-Venant. D'une part, l'obtention formelle du système classique de Saint-Venant à partir des équations d'Euler, c'est-à-dire sans viscosité, est bien connue (voir par exemple Stoker [171] en 1958 ou Whitham [179] en 1999). D'autre part, les travaux plus récents de J.-F. Gerbeau et B. Perthame [101] (2001) dérivent une version visqueuse des équations de shallow water 1D à partir des équations de Navier-Stokes 2D dans le cas d'un fond plat, avec une loi de friction de type Navier au fond. Cette version étendue du système de Saint-Venant s'écrit :

$$\left\{\begin{array}{lcl} \partial_t H + \mathrm{div}_{\mathbf{x}}(H\,\overline{\mathbf{U}}) & = & 0\,, \\ \\ \partial_t(H\,\overline{\mathbf{U}}) + \mathrm{div}_{\mathbf{x}}\left(H\,\overline{\mathbf{U}}\otimes\overline{\mathbf{U}}\right) + g\,H\,\nabla_{\mathbf{x}} H & = & -g\,H\,\nabla_{\mathbf{x}} z_b + 4\,\mu\,\mathrm{div}_{\mathbf{x}}(H\,\nabla_{\mathbf{x}}\overline{\mathbf{U}}) - \tilde{\kappa}\,\overline{\mathbf{U}}\,, \end{array}\right. \tag{1.12}$$

où $\tilde{\kappa}$ est le coefficient de friction modifié, défini par :

$$\tilde{\kappa} = \frac{\kappa}{1+\frac{\kappa}{3\,\mu}\,H}\,.$$

Cependant, il est à noter que l'obtention de ce système nécessite, pour être rigoureuse, des hypothèses supplémentaires. En particulier, dans [101], les auteurs requièrent l'asymptotique suivante pour les coefficients de friction et de viscosité :

$$\mu = \varepsilon\,\mu_0\,, \quad \kappa = \varepsilon\,\kappa_0\,. \tag{1.13}$$

Alors, sous ces hypothèses, les systèmes (1.11) et (1.12) sont des approximations de Navier-Stokes en $O(\varepsilon)$ et $O\left(\varepsilon^2\right)$ respectivement. Plus tard, S. Ferrari et F. Saleri [87] (2004), puis F. Marche [143] (2007) généralisent le résultat de [101] : ils dérivent un système de Saint-Venant 2D à partir des équations de Navier-Stokes 3D, incluant les effets de Coriolis et avec

une topographie non triviale, soumise cependant à une autre restriction mathématique, à savoir la faible variation de la bathymétrie :

$$\nabla_{\mathbf{x}} z_b \,=\, O(\varepsilon). \tag{1.14}$$

Notons que dans [143], l'auteur considère également un terme de friction turbulente, et obtient un terme visqueux différent de celui de [87]. Dans un travail plus récent, L. Bonaventura, A. Decoene et F. Saleri [75] (2007) dérivent un autre modèle, avec une nouvelle correction des termes de friction. Citons également les travaux de J.F. Bouchut et M. Westdickenberg en 2004 [35], puis ceux de M. Boutounet, L. Chupin, P. Noble et J.P. Vila en 2008 [37] où les auteurs s'affranchissent de l'hypothèse sur le gradient de la bathymétrie. Mentionnons enfin l'article de 2007 de D. Bresch et P. Noble [44] dans lequel les auteurs proposent une justification mathématique rigoureuse de la dérivation formelle du système de Saint-Venant 1D à partir des équations de Navier-Stokes incompressibles 2D sur un plan incliné.

Energies, résultats d'existences. Beaucoup d'études théoriques ont été conduites sur les équations de Saint-Venant, en raison notamment de leur structure mathématique hyperbolique (nous y reviendrons un peu plus loin). Nous nous intéressons ici aux résultats d'existence de solutions, qui pourront être confrontés au théorème du Chapitre 2.
Rappelons d'abord l'énergie naturelle du système de Saint-Venant, qui fournit les estimations *a priori* de base lorsque l'on cherche des solutions faibles. Elle s'écrit :

$$E \,=\, \frac{1}{2}\, H\, |\overline{\mathbf{U}}|^2 + \frac{1}{2}\, g\, H^2 + g\, H\, z_b\,.$$

L'inégalité d'énergie s'obtient de manière classique (voir par exemple les ouvrages de D. Serre [169] sur les lois de conservation) et varie suivant le terme de viscosité choisi. Nous verrons à l'Annexe A que notre système Saint-Venant multicouche possède également une énergie consistante avec celle du modèle classique à une couche. Cependant, cette estimation n'est pas suffisante pour établir l'existence de *solutions faibles*. La difficulté majeure provient du manque de contrôle sur la hauteur H, pour les passages à la limite dans les termes non linéaires. Plusieurs solutions ont été trouvées, en ajoutant des termes de friction quadratique, de capillarité, en considérant différents termes visqueux ; le principe est d'établir des estimations *a priori* supplémentaires, contrôlant $\log H$ dans un espace adéquat. Citons deux résultats d'existence de solutions faibles ainsi obtenus.

- Par exemple, P. Orenga [157] (1995) établit un théorème d'existence de solutions faibles, pour un terme de viscosité $\mu\, H\, \triangle \overline{\mathbf{U}}$, avec des données initiales *suffisamment petites* et des conditions de Dirichlet au bord du domaine. La majoration essentielle est un contrôle de $H\, \log H$ dans $L^\infty\left(0, T; L^1\right)$.
- En considérant le terme visqueux div $\left(H\, \nabla \overline{\mathbf{U}}\right)$, plus « consistant » avec Navier-Stokes (c'est le cas que nous considèrerons dans la suite), il n'est pas possible de diviser l'équation de la quantité de mouvement par la hauteur, il faut donc procéder autrement. En 2003, les travaux de D. Bresch, B. Desjardins et C.K. Lin [42] préparent les

estimations nécessaires. Ainsi, D. Bresch et B. Desjardins montrent dans [43] l'existence globale de solutions faibles pour système de Saint-Venant bidimensionnel avec conditions aux bords périodiques, ainsi que des termes de capillarité et de frictions linéaire et quadratique. Le point crucial ici est l'introduction d'une énergie particulière pour le système, la *BD-entropie*, qui permet de compléter les estimations *a priori* et d'obtenir suffisamment de compacité sur la hauteur. Cette nouvelle énergie fait apparaître une nouvelle vitesse

$$V := \overline{\mathbf{U}} + \mu \nabla (\log H) \ .$$

Remarquons qu'un élément clé dans l'obtention de ces estimations est d'utiliser la conservation de la masse pour écrire une équation sur la fonction $\log H$, sous la forme bidimensionnelle :

$$\partial_t (H \nabla \log H) + \operatorname{div} \left(H \nabla \overline{\mathbf{U}} \right) + \operatorname{div} \left(H \, \overline{\mathbf{U}} \otimes \nabla \log H \right) = 0 \, .$$

Ainsi, si les techniques de P. Orenga, comme celles de D. Bresch et B. Desjardins ont été ensuite adaptées à des problèmes multicouches (voir ci-après), nous verrons à l'Annexe A qu'il est difficile de les appliquer à notre système multicouche qui ne possède pas de lois de conservations pour chaque couche séparément.
Concernant l'étude des solutions fortes, nous renvoyons aux références bibliographiques du Chapitre 2 p. **??**. Retenons simplement que les estimations *a priori* sont plus « classiques » et que la restriction fondamentale dans ces résultats est de considérer des conditions initiales *en dehors des zones sèches*, c'est-à-dire une hauteur initiale du fluide H^0 minorée par une constante strictement positive. Cette hypothèse aura son analogue dans le Théorème 5 où nous établissons l'existence locale de solution forte pour notre modèle multicouche.

Discrétisation. Abordons enfin la problématique de la discrétisation et des simulations numériques du modèle de Saint-Venant. La validité et l'efficacité numériques de ce système sont largement reconnues et vont au-delà des connaissances théoriques du système. En effet, des schémas simples peuvent traiter de manière réaliste des problèmes de rupture de barrage dans lesquels la condition de stricte positivité de la hauteur n'est plus satisfaite. Historiquement, les premières méthodes de discrétisation utilisées pour les équations de Saint-Venant sont les *éléments finis* (voir par exemple la thèse de J. Proft [163]). En effet, sous leur formulation non conservative (hauteur-vitesse), leur lien avec les équations de Navier-Stokes est mis en évidence et la discrétisation par éléments finis est donc naturelle. Cependant, nous privilégions ici une autre famille de méthodes, plus adéquates pour le traitement de solutions discontinues, et liées à la formulation conservative du système : les *Volumes Finis*. C'est la méthode que nous emploierons pour construire les schémas numériques du Chapitre 3, mais également dans la Partie II de la thèse. Nous donnons donc maintenant quelques détails sur sa mise en oeuvre dans le contexte du modèle classique de Saint-Venant, puisque le système introduit au Chapitre 2 aura une structure comparable.
Initialement utilisée pour la discrétisation des équations d'Euler, la méthode des volumes finis est particulièrement bien adaptée au système de Saint-Venant en raison de sa structure partiellement *hyperbolique*. En effet, le système *homogène*, qui s'écrit en dimension 1

d'espace :

$$\begin{cases} \partial_t H + \partial_x (H\,U) = 0\,, \\ \\ \partial_t (H\,U) + \partial_x \left(H\,U^2 + g\dfrac{H^2}{2} \right) = 0\,, \end{cases}$$

est un système *strictement hyperbolique* de deux lois de conservation, pourvu que la hauteur H reste strictement positive. Ses valeurs propres réelles distinctes sont alors

$$U - \sqrt{gH} \quad \text{et} \quad U + \sqrt{gH}\,.$$

Nous renvoyons aux ouvrages de D. Serre [169] pour l'étude théorique des problèmes hyperboliques. Ce qui nous intéresse ici, ce sont des particularités de ces systèmes utiles à la construction du schéma, à savoir la formulation *conservative*, et la *propagation à vitesse finie* de l'information (qui permet l'utilisation de schémas explicites). Pour une description précise des méthodes volumes finis en général, nous renvoyons le lecteur à quelques ouvrages classiques : ceux de R. J. LeVeque [133, 134], de R. Eymard, T. Gallouët et R. Herbin [84], ou encore d'E. Godlewski et P.A. Raviart [102]. Nous présentons succintement ici le principe général de la méthode pour une loi de conservation scalaire (pas de terme source) :

$$\partial_t V + \partial_x F(V) \;=\; 0\,. \tag{1.15}$$

Nous considérons un schéma aux différences finies explicite en temps (typiquement Euler) et nous donnons un maillage du domaine spatial $(x_{j+1/2})_{j\in\mathbb{Z}}$: les noeuds du maillage $x_{j+1/2}$ sont les interfaces des cellules de contrôle ou mailles $C_j = \big(x_{j-1/2}; x_{j+1/2}\big)$. En intégrant l'équation (1.15) sur chaque maille, nous approchons la solution non pas en valeurs ponctuelles aux noeuds du maillage, mais en valeurs moyennes sur chaque maille (d'où l'obtention d'une solution numérique constante par morceaux, foncièrement discontinue). Pour l'équation (1.15), le schéma s'écrit :

$$V_j^{n+1} - V_j^n + \frac{\Delta t^n}{\Delta x_j} \left(F_{j+1/2}^n - F_{j-1/2}^n \right) \;=\; 0\,,$$

où V_j^n désigne une approximation de la moyenne de V sur la cellule C_j au temps t^n. Enfin, le *flux* numérique $F_{j+1/2}^n$ représente une approximation du flux F à l'interface entre les cellules C_j et C_{j+1} au temps t^n. Parce que l'on n'a pas d'information sur la solution aux interfaces des mailles, la première difficulté dans la construction du schéma réside dans le choix du flux numérique. C'est la multiplicité de choix possibles pour ce flux, en fonction de la physique du problème, qui fait la diversité des schémas volumes finis.
Dans cette thèse (au Chapitre 3, mais également dans la Partie II), nous nous restreindrons aux schémas dits *à trois points*, pour lesquels le flux numérique s'écrit :

$$F_{j+1/2}^n = \mathcal{F}\left(V_j^n, V_{j+1}^n\right)\,,$$

où $\mathcal{F}$ est *consistant* avec le flux continu, c'est-à-dire

$$\mathcal{F}\,(V,V) \;=\; F(V)\,.$$

Plus précisément, nous utiliserons essentiellement des flux de *Lax-Friedrichs*. Ce flux s'écrit, pour l'équation (1.15) :

$$\mathcal{F}\left(V_j^n, V_{j+1}^n\right) = \frac{1}{2}\left[F\left(V_j^n\right) + F\left(V_{j+1}^n\right) - \frac{\Delta x}{\Delta t}\left(V_{j+1}^n - V_j^n\right)\right].$$

Outre la consistance, le schéma doit également satisfaire des propriétés de *stabilité*. Un critère primordial est la condition CFL, condition *nécessaire* (mais pas suffisante !) de stabilité. Elle s'écrit pour (1.15) et dans le cas d'un schéma explicite à trois points :

$$\lambda := a_\infty \frac{\Delta t}{\Delta x} \leqslant 1\,,$$

où $a_\infty = \max_{v\in\mathbb{R}} |F'(v)|$. Pour le flux de Lax-Friedrichs, cette condition satisfaite entraîne que pour tout temps n, la valeur V_j^{n+1} peut s'écrire comme une *combinaison convexe* de V_{j-1}^n et V_{j+1}^n, condition cruciale pour obtenir la stabilité du schéma. Cependant, ce flux provoquant de la diffusion numérique, nous utiliserons des limiteurs de pentes (voir le Chapitre 3).

Remarque 2. Ces deux conditions de consistance et de stabilité ne suffisent pas en général pour démontrer la convergence d'un schéma volumes finis, mais nous n'étudierons pas la convergence mathématique des schémas dans la Partie I. Nous verrons cependant à la Partie II une autre propriété du flux numérique qui sera utile dans la preuve de convergence, celle d'être *TVD* (pour *Total Variation Diminishing*) (voir la Section 2 de l'introduction générale).

Evidemment, toutes les équations ne ressemblent pas à (1.15) et la question du choix du flux n'est pas la seule difficulté. Par exemple, dans le cas de Saint-Venant, il y a des termes sources et des termes non conservatifs. Ainsi, en général, dans l'élaboration d'un schéma volumes finis, nous devons répondre aux interrogations suivantes :

- quel choix pour le *flux numérique* ? Est-il consistant ?
- que faire lorsqu'il y a des produits *non conservatifs* ?
- comment discrétiser les termes sources ?

Pour répondre à ces questions, nous nous basons sur la connaissance que l'on a du problème continu, c'est-à-dire ses propriétés de stabilité. Dans le cas de Saint-Venant, les enjeux discrets majeurs sont :

- la conservation de la *positivité* de la hauteur,
- la préservation des états stationnaires, notamment du « lac au repos » (on parle alors de schéma *well-balanced*),
- une inégalité d'entropie discrète.

Au Chapitre 3, nous nous intéresserons essentiellement aux deux premiers points, laissant de côté la question d'une entropie discrète. Evoquons par exemple deux articles qui proposent des schémas *well-balanced* pour le système unidimensionnel avec le seul terme source topographique. Dans [98], T. Gallouët, J-M. Hérard et N. Seguin proposent un schéma de type VFRoe-ncv, où le traitement du terme source se fait en ajoutant une équation au système, de façon à garder un problème homogène, mais non conservatif. Dans [10], E. Audusse, F. Bouchut, M.O. Bristeau, R. Klein et B. Perthame adoptent une stratégie différente : il s'agit de partir du solveur stable pour le problème homogène, et de discrétiser le terme source en s'appuyant sur l'état d'équilibre du « lac au repos » , *i.e.* pour lequel on a l'équilibre suivant :

$$\partial_x \left(g \frac{H^2}{2} \right) = -g\, H\, \partial_x z_b \,.$$

Au Chapitre 3, nous éviterons la difficulté liée à la discrétisation du terme de topographie en utilisant une topographie régulière. Par ailleurs nous considèrerons des schémas tout explicites et à une seule étape en temps dans cette partie, même si la condition CFL devient contraignante dans les cas où la viscosité μ est non nulle.

Remarque 3. Pourtant, les schémas explicites et à une seule étape ne sont pas toujours adéquats, notamment pour traiter les problèmes hyperboliques d'ordre 1 possédant un terme source *raide* (*stiff*). Ce type de problème nous intéressera à la Partie II, c'est pourquoi nous adopterons alors une stratégie de *splitting* (voir la Section 2 de l'introduction générale).

Enfin, il existe de multiples autres stratégies volumes finis de résolution numérique des équations de Saint-Venant, comme les schémas *cinétiques* (voir par exemple l'article de B. Perthame et C. Simeoni [160]). Nous renvoyons également le lecteur aux états de l'art sur le système de Saint-Venant présentés dans les thèses d'E. Audusse [9] et F. Marche [144].

Terminons cette partie sur les systèmes classiques de Saint-Venant par quelques commentaires. Il est dorénavant établi qu'ils sont très efficaces numériquement, robustes et peu coûteux, notamment pour simuler des problèmes hydrauliques ou côtiers. Ils présentent néanmoins quelques faiblesses. D'une part, le « bon comportement » numérique de ces équations n'est pas totalement compris ni justifié au niveau théorique. En effet, le système est mal posé dans le vide (lorsque H atteint 0, perte de l'hyperbolicité), son domaine de validité est restreint aux eaux peu profondes et nous avons vu que sa fermeture rigoureuse dans le cas visqueux nécessite des hypothèses bien particulières (1.13). D'autre part, la formulation même du système, en hauteur et vitesse moyenne sur la profondeur, entraîne une perte d'information sur le profil vertical de la vitesse dans le fluide.
Afin, d'un côté, de pallier à ce manque d'information sur la vitesse à l'intérieur du fluide, et de l'autre, de préserver la « formulation Saint-Venant » à l'efficacité numérique reconnue, des modèles *intermédiaires* entre les équations primitives et le système de Saint-Venant ont été introduits, les *modèles multicouches de type Saint-Venant* [12, 15].

Modèles classiques multicouches de type Saint-Venant

Dans les modèles classiques multicouches, l'objectif principal est de récupérer de l'information sur le profil vertical des vitesses dans la couche de fluide, mais sans s'affranchir des restrictions physiques du problème de Saint-Venant. Le principe général d'obtention d'un tel système est, comme pour le modèle classique de Saint-Venant, d'intégrer l'équation de conservation de la quantité de mouvement dans la direction verticale, mais cette intégration s'effectue après une *discrétisation* verticale du volume de fluide. En d'autres termes, il s'agit de découper la hauteur totale du fluide H en N couches :

$$H(t,\mathbf{x}) = \sum_{i=1}^{N} h_i \,.$$

La manière classique de procéder à ce « découpage » , illustrée à la Figure 1.2 (pour $N = 4$), consiste à *suivre la surface libre*, c'est-à-dire écrire la hauteur h_i de la couche i comme une fraction de la hauteur totale :

$$h_i(t,x) = l_i\, H(t,x), \;\; \text{avec} \;\; 0 \leqslant l_i \leqslant 1, \;\; \sum_{i=1}^{N} l_i = 1\,.$$

C'est avec cette approche-ci que les premiers modèles multicouches (à un seul fluide) ont

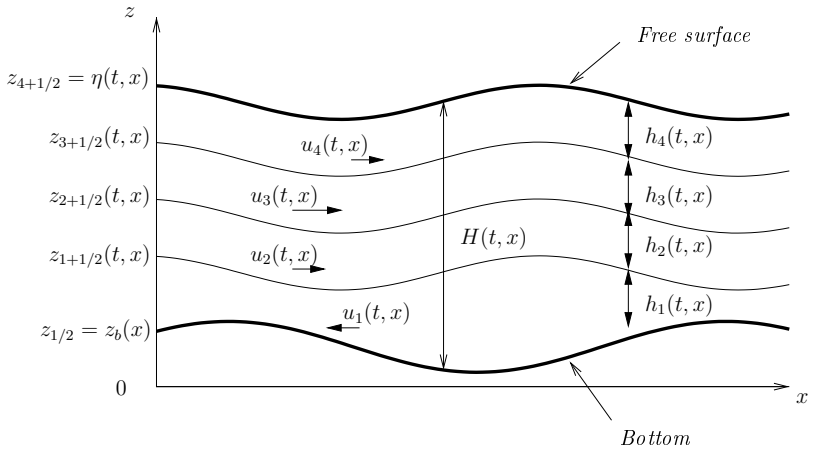


FIGURE 1.2 – *Approche multicouche classique.*

été introduits. D'abord, E. Audusse propose en 2005 dans [12] une version sans échange de masse entre les couches, c'est-à-dire que, pour chaque couche i, nous disposons de la loi de conservation :

$$\partial_t h_i + \partial_x \left(h_i\, u_i\right) \;=\; 0\,,$$

où u_i représente la vitesse du fluide dans la couche i, *i.e.* la moyenne sur la couche i de la vitesse horizontale :

$$u_i(t,x) \,=\, \frac{1}{h_i}\int_{z_{i-1/2}}^{z_{i+1/2}} u(t,x,z)\,\mathrm{d}z\,,\; 1\leqslant i\leqslant N\,. \tag{1.16}$$

Il apparaît que le système multicouche de [12] perd la propriété d'hyperbolicité, et s'apparente davantage à un modèle de N fluides immiscibles plutôt qu'à celui d'un seul fluide. C'est pourquoi E. Audusse, M.O. Bristeau, B. Perthame et J. Sainte-Marie introduisent en 2010 dans [15] un autre modèle multicouche basé sur la même discrétisation verticale (voir la Figure 1.2), mais avec un terme d'échange de masse entre les couches i et $i+1$, à savoir :

$$\partial_t h_i + \partial_x\left(h_i\,u_i\right) \,=\, w_{i+1/2} - w_{i-1/2}\,.$$

Ce nouveau système s'avère être hyperbolique et plus consistant avec la physique du problème. La stratégie pour l'obtenir est basée sur les mêmes hypothèses que [101] et [87] et s'applique rigoureusement (formellement) dans le cas non visqueux. Pourtant, il subsiste quelqu'incertitude dans le cas avec viscosité : avec la même hypothèse (1.13) que dans le cas à une couche [101], on retrouve dans les modèles [**?**, 15] le terme visqueux de (**??**) dans chaque couche, mais la justification mathématique de ce choix est plus délicate avec plusieurs couches qu'avec une seule.
Outre la dérivation formelle des systèmes multicouches de [12] et [15], les auteurs en proposent également une étude théorique : la question de l'hyperbolicité est soulevée et une inégalité d'entropie similaire à l'estimation d'énergie du système de Saint-Venant classique est établie.
Enfin, le traitement numérique proposé est un schéma cinétique et plusieurs propriétés du schéma sont démontrées. Mais citons également d'autres travaux, essentiellement numériques, qui démontrent l'efficacité de ces systèmes (en version bicouches au moins) pour modéliser des zones côtières ou de détroits [14, 13, 11, 59, 104].

Si les auteurs de [12, 15] n'étudient pas précisément l'existence de solutions, il existe cependant plusieurs résultats d'existence de solutions *faibles* pour des systèmes bicouches. Comme évoqué précédemment, ces résultats sont basés sur les techniques de P. Orenga si le terme visqueux est sous la forme $\mu h_i \triangle u_i$ ou celles de D. Bresch et B. Desjardins s'il est de la forme $\mu\,\mathrm{div}\,(h_i\,\nabla u_i)$. Citons par exemple les articles [64, 95, 151, 161] de 2003 à 2006 pour le premier cas, et l'article de 2009 [76] qui adapte la *BD-entropie* à un modèle bicouche. Il est important de retenir que ces travaux concernent des modèles de deux fluides *immiscibles*, qui possèdent donc deux lois de conservations de la masse, à savoir

$$\partial_t h_i + \mathrm{div}\,(h_i\,\mathbf{u}_i) \,=\, 0\,,$$

ce qui est crucial pour l'obtention des estimations d'énergies supplémentaires. Mais nous nous intéressons ici à un modèle multicouche pour un seul fluide : nous ne disposons plus de N lois de conservation mais d'une seule sur la hauteur totale du fluide, comme dans le modèle de [15].

Nous venons ainsi d'évoquer trois modèles largement utilisés en océanographie, plutôt « profonde » pour le modèle primitif et côtière pour les modèles de Saint-Venant, à une ou plusieurs couches. Le modèle que nous construisons dans la Partie I s'inscrit plutôt dans le contexte « eaux profondes » , bien qu'il ressemble, dans sa formulation, aux précédents modèles multicouches.

1.2 Travaux effectués : un autre modèle multicouche de Saint-Venant

Dans la Partie I, nous introduisons un nouveau modèle multicouche de type Saint-Venant, à partir des équations primitives de l'océan (avec dimension) (1.8). Ainsi, comme nous allons le voir ci-après, la formulation du système a des points communs avec ceux de [12, 15] mais le domaine d'application est fondamentalement différent puisque nous restons ici loin des zones côtières. De fait les modèles ne sont pas vraiment comparables, si ce n'est sur la forme, la méthodologie employée pour la construction.
Voyons d'abord comment la stratégie que nous adoptons pour obtenir notre modèle est similaire à celle de [12, 15]. Nous découpons aussi la hauteur du fluide en couches minces, mais ce « découpage » n'est pas fait de la même manière : nous ne suivons pas ici la surface libre, mais découpons le fluide en imposant les hauteurs des couches intermédiaires, comme illustré à la Figure 1.3 (pour 4 couches).

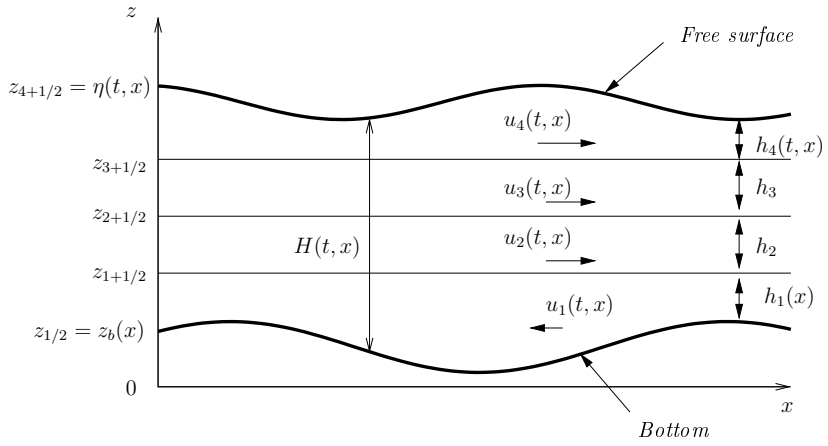


FIGURE 1.3 – *Une autre approche multicouche.*

Précisément, la hauteur totale est divisée comme suit.

$$\eta - z_b = H := \sum_{i=1}^{N} h_i, \quad \text{avec } h_i = z_{i+1/2} - z_{i-1/2} = O(\overline{h}), \quad 1 \leqslant i \leqslant N, \tag{1.17}$$

où la constante $\overline{h}$ est fixée et les noeuds $z_{i+1/2}$ de la discrétisation verticale sont donnés

par (voir la Figure 1.3) :

$$\begin{cases} z_{1/2} = z_b(x), \\ \\ z_{i+1/2} = i\,\overline{h}, \quad 1 \leqslant i \leqslant N-1, \\ \\ z_{N+1/2} = \eta(t,x). \end{cases} \tag{1.18}$$

Le modèle multicouche que nous dérivons au Chapitre 2 s'écrit, pour tout $(t, \mathbf{x})$ dans $\mathbb{R}^+ \times \mathbb{R}^2$:

$$\begin{cases} \partial_t H + \operatorname{div}_{\mathbf{x}} \left(\displaystyle\sum_{i=1}^{N} h_i\, \mathbf{u}_i \right) & = 0\,, \\ \\ \partial_t \left(h_N\, \mathbf{u}_N \right) + \operatorname{div}_{\mathbf{x}} \left(h_n\, \mathbf{u}_N \otimes \mathbf{u}_N + g\, \dfrac{h_N^2}{2} \right) & = \mu \left(\operatorname{div}_{\mathbf{x}} \left(h_N\, \nabla_{\mathbf{x}} \mathbf{u}_N \right) + \mathbf{DU}^z_{N+1/2} - \mathbf{DU}^z_{N-1/2} \right) \\ & \quad - g\, h_N\, \nabla_{\mathbf{x}} z_b + w_{N-1/2}\, \mathbf{u}_{N-1/2} - w_{N+1/2}\, \mathbf{u}_{N+1/2} \\ & \quad - f\, \left(h_N\, \mathbf{u}_N \right)^{\perp}\,, \\ \\ \partial_t \left(h_i\, \mathbf{u}_i \right) + \operatorname{div}_{\mathbf{x}} \left(h_i\, \mathbf{u}_i \otimes \mathbf{u}_i \right) + g\, h_i\, \nabla_{\mathbf{x}} h_N & = \mu \left(h_i\, \Delta_{\mathbf{x}} \mathbf{u}_i + \mathbf{DU}^z_{i+1/2} - \mathbf{DU}^z_{i-1/2} \right) \\ & \quad - g\, h_i\, \nabla_{\mathbf{x}} z_b + w_{i-1/2}\, \mathbf{u}_{i-1/2} - w_{i+1/2}\, \mathbf{u}_{i+1/2} \\ & \quad - f\, \left(h_i\, \mathbf{u}_i \right)^{\perp}\,, \quad 1 \leqslant i \leqslant N-1\,. \end{cases} \tag{1.19}$$

Dans ce système, le terme d'échange de masse $w_{i+1/2}$ désigne simplement la valeur de la vitesse verticale à l'interface entre les couches i et $i+1$, *i.e.* au point $z_{i+1/2}$. Il est défini par :

$$\begin{cases} w_{1/2} = \mathbf{u}_1 \cdot \nabla_{\mathbf{x}} z_b, \\ \\ w_{i+1/2} - w_{i-1/2} = -h_i\, \operatorname{div}_{\mathbf{x}} \mathbf{u}_i, \quad 1 \leqslant i \leqslant N-1. \end{cases} \tag{1.20}$$

De même, $\mathbf{u}_{i+1/2}$ (notation identique à celle de [15]) représente une approximation de la vitesse horizontale à l'interface $z_{i+1/2}$. Si les auteurs de [15] choisissent une reconstruction upwind de ce terme, nous considérons ici une moyenne :

$$\mathbf{u}_{i+1/2} = \begin{cases} 0 & \text{si} \quad i = 0\,,\, N\,, \\ \\ \left(h_i\, \mathbf{u}_{i+1} + h_{i+1}\, \mathbf{u}_i \right) / \left(h_{i+1} + h_i \right) & \text{si} \quad 1 \leqslant i \leqslant N-1. \end{cases} \tag{1.21}$$

Les termes $\mathbf{DU}^z_{i+1/2}$ représentent les dérivées verticales de la vitesse horizontale aux inter-

faces $z_{i+1/2}$ et viennent des termes visqueux verticaux. Ils sont donnés par :

$$\mathbf{DU}^z_{i+1/2} = \begin{cases} \kappa\,\mathbf{u}_1/\mu & \text{si} \quad i = 0\,, \\ 2\,(\mathbf{u}_{i+1} - \mathbf{u}_i)/(h_i + h_{i+1}) & \text{si} \quad 1 \leqslant i \leqslant N-1\,, \\ 0 & \text{si} \quad i = N\,. \end{cases} \tag{1.22}$$

Ce système est alors une approximation en $O\left(\overline{h}^2\right)$ de (1.8), où $\overline{h}$ désigne la hauteur des couches internes fixée préalablement. Précisément, nous montrons la proposition suivante.

Proposition 4 (Chapitre 2, p. **??**)**.**
Supposons que les variations de la bathymétrie vérifient :

$$\nabla_{\mathbf{x}} z_b \,=\, O\left(\overline{h}\right). \tag{1.23}$$

Alors le modèle multicouche (1.19), où h_i, $\mathbf{u}_{i+1/2}$ et $w_{i+1/2}$ sont donnés respectivement par (1.17), (1.21) et (1.20), est une approximation formelle des équations primitives (1.8) en $O\left(\overline{h}^2\right)$.

Dans le Chapitre 2, nous comparons brièvement ce modèle avec ceux de [12, 15], en soulignant la différence entre les termes visqueux : nous n'avons pas besoin ici d'hypothèse particulière sur les régimes de friction et viscosité pour les équations primitives, et nos termes de viscosité s'obtiennent formellement naturellement.
Nous établissons également un théorème d'existence de solution forte locale pour la version 1D de (1.19) :

$$\begin{cases} \partial_t H + \partial_x \left(\displaystyle\sum_{i=1}^{N} h_i\, u_i \right) & = 0\,, \\ \partial_t \left(h_N\, u_N\right) + \partial_x \left(h_n\, u_N^2 + g\, \dfrac{h_N^2}{2} \right) & = \mu\, \left(\partial_x \left(h_N\, \partial_x u_N \right) + DU^z_{N+1/2} - DU^z_{N-1/2} \right) \\ & \quad - g\, h_N\, \partial_x z_b + w_{N-1/2}\, u_{N-1/2} - w_{N+1/2}\, u_{N+1/2}\,, \\ \partial_t \left(h_i\, u_i\right) + \partial_x \left(h_i\, u_i^2 \right) + g\, h_i\, \partial_x h_N & = \mu\, \left(h_i\, \partial_{xx} u_i + DU^z_{i+1/2} - DU^z_{i-1/2} \right) - g\, h_i\, \partial_x z_b \\ & \quad + w_{i-1/2}\, u_{i-1/2} - w_{i+1/2}\, u_{i+1/2}\,, \quad 1 \leqslant i \leqslant N-1\,. \end{cases} \tag{1.24}$$

où les termes de Coriolis n'apparaissent plus car ils n'ont pas de sens en 1D. Introduisons quelques notations avant d'énoncer le théorème obtenu. Pour toute fonction f, on note $\|f\|$ (resp. $\|f\|_k$) sa norme L^2 (resp. H^k). Si $\mathbf{f} = (f_1, \ldots, f_n)$ est multidimensionnelle, on définit sa norme $\mathbf{H}^k$ coordonnée par coordonnée :

$$\|\mathbf{f}\|_k := \sum_{i=1}^{n} \|f_i\|_k\,.$$

Par ailleurs, si B désigne un espace de Banach, k un entier naturel et T une constante positive, on note $L^k_\infty(0,T;B)$ l'espace de Banach formé des fonctions f définies sur $[0,T]$ à valeurs dans B qui sont k fois différentiables par rapport à t et dont toutes les dérivées sont bornées dans B. Voici le théorème que nous obtenons.

Théorème 5 (Chapitre 2, p. **??**)**.**
Considérons le système (1.24) avec les conditions initiales

$$(\mathbf{U}, h_N)(0,x) = (\mathbf{U}^0(x), h_N^0(x)) \in \mathbf{H}^2(\mathbb{R}), \tag{1.25}$$

où $\mathbf{U} = (u_1 \dots u_N)^T$ *représente le vecteurs des vitesses. Supposons*

$$\inf_{x\in\mathbb{R}} h_N^0(x) \geqslant \eta_0 > 0,$$

pour une constante positive η_0 *et notons* $E = 2\,\|(\mathbf{U}^0, h_N^0)\|_2$*. Supposons également la régularité de la topographie* $z_b \in \mathcal{C}^2(\mathbb{R})$*. Alors, il existe un temps* $T > 0$ *tel que le problème de Cauchy (1.24)-(1.25) possède une unique solution* $(\mathbf{U}, h_N)$ *vérifiant :*

$$\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2\left(0,T;\mathbf{H}^3(\mathbb{R})\right),$$

$$h_N \in \mathcal{C}(0,T;H^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^1(\mathbb{R})).$$

En outre, pour tout t *de* $[0,T]$*,*

$$\forall\, x \in \mathbb{R},\ h_N(t,x) \geqslant \big(\inf_{x\in\mathbb{R}} h_N^0(x)\big)/2 > 0.$$

Enfin, on a les inégalités d'énergie suivantes :

$$\|(\mathbf{U}, h_N)(t)\|_2 \leqslant E, \quad \left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\, d\tau\right)^{1/2} \leqslant E.$$

Remarque 6. La preuve de ce résultat utilise une deuxième formulation du problème, à savoir un système de N équations paraboliques couplées avec une équation de transport sur la hauteur de la dernière couche h_N (voir le système (2.3.1), p. 54), et repose sur la méthode d'énergie de Nishida et Matsumura [146]. En revanche, nous n'obtenons pas de résultat d'existence de solution faible. Il se trouve en effet que les techniques employées par D. Bresch et B. Desjardins [43] sont difficiles à généraliser à notre problème multicouche, en particulier parce que nous ne disposons pas de N équations de conservations de la masse et que nous ne pouvons établir suffisamment d'estimations d'énergies (voir la Section A.1 de l'Annexe A).

Remarque 7. Ce théorème est seulement local en temps et surtout contraint à l'hypothèse de stricte positivité de la hauteur de la couche supérieure h_N. Mais cela ne constitue pas une faiblesse de notre modèle qui n'est pas destiné à décrire des zones sèches. Au contraire, cette contrainte permet de fournir un comportement dynamique à notre modèle, en lui ôtant ou ajoutant des couches si la hauteur h_N devient trop petite ou trop grande. Nous verrons dans les simulations numériques du Chapitre 3 comment le nombre de couches du modèle peut varier au cours du temps.

Nous nous intéressons donc au Chapitre 3 à la discrétisation par volumes finis du modèle multicouche (1.24). Pour cela, nous utilisons une troisième formulation du système multicouche (voir le système (3.1.1) p. 68). Nous construisons donc un schéma volume finis explicite en temps, avec un flux de Lax-Friedrichs et des limiteurs de pente *minmod*. Par ailleurs nous proposons également une discrétisation du même type pour le système primitif (1.8) sans viscosité ni friction, en utilisant une formulation lagrangienne du système.
Nous présentons ensuite des résultats numériques avec trois objectifs :

- montrer que notre modèle est au moins aussi bon que les modèles classiques de Saint-Venant et qu'il approche correctement les équations primitives sans viscosité,
- faire apparaître des recirculations à l'intérieur du fluide,
- et illustrer le comportement dynamique du modèle : on peut ajouter et retirer des couches par au-dessus.

Les Chapitres 2 et 3 sont réunis dans un article soumis [165]. Enfin, nous présentons en Annexe A de la Partie I quelques compléments sur notre modèle multicouche. Nous présentons quelques éléments sur son énergie en 1D à la Section A.1, ainsi que des simulations numériques supplémentaires à la Section A.2.

2 Partie II : analyse d'un schéma préservant l'asymptotique

Dans cette section, nous décrivons les motivations initiales qui ont conduit au travail de la Partie II, effectué en collaboration avec F. Filbet [94]. Nous en résumons ensuite les principaux résultats.
Le phénomène qui nous intéresse ici est celui de la relaxation, qui apparaît dans de nombreuses situations physiques [65, 142] : par exemple en théorie cinétique des gaz monoatomiques, si un état d'équilibre est perturbé, le système relaxe graduellement vers l'équilibre. Ce mécanisme de relaxation existe également dans les matériaux élastiques avec mémoire, ou dans les transitions de phases. Il est souvent représenté dans les équations par un coefficient $\varepsilon > 0$, soit grand (phénomène relativement lent par rapport aux échelles de temps caractéristiques), soit très petit (phénomène très rapide, quasiment instantané). Cela constitue un enjeu à la fois mathématique et numérique si l'on souhaite modéliser et simuler ce processus dans lequel plusieurs échelles de temps s'affrontent. Le centre de nos préoccupations est l'enjeu numérique, mais il est primordial de comprendre la physique du problème au niveau continu avant de le discrétiser.
Si les travaux réalisés ici concernent un modèle jouet hyperbolique, nous sommes néanmoins motivés par les équations cinétiques. C'est pourquoi nous commençons par évoquer brièvement l'équation de Boltzmann dans un adimensionnement particulier. Nous rappelons les résultats existants sur la limite singulière lorsque le *libre parcours moyen* tend vers 0 (c'est-à-dire lorsque le gaz relaxe très vite vers un équilibre thermodynamique), résultats qui se situent au niveau continu. Nous abordons ensuite la problématique du traitement numérique ainsi que les solutions précédemment apportées : afin de fournir une discrétisation du problème cinétique qui soit en adéquation avec la limite hydrodynamique lorsque

le paramètre de relaxation tend vers 0, le cadre choisi est celui des schémas *préservant l'asymptotique*. Nous en donnons une définition et nous décrivons plusieurs stratégies déjà développées et validées numériquement pour les équations cinétiques. Puis nous délaissons les modèles cinétiques pour nous concentrer sur les problèmes hyperboliques de relaxation. En effet, ils constituent un cadre cinétique simplifié (à vitesses discrètes) pour lequel une analyse mathématique rigoureuse peut être plus facilement conduite. Nous examinons donc plusieurs travaux effectués dans ce contexte, tant sur le plan continu que discret. Enfin, nous introduisons notre modèle simplifié issu de la famille des problèmes hyperboliques de relaxation et présentons notre schéma préservant l'asymptotique ainsi que nos résultats.

2.1 Motivation et état de l'art

Nous avons vu à la section précédente une description macroscopique d'un fluide. En se plaçant à un autre niveau d'observation, à l'échelle *mésoscopique*, on regarde plutôt l'évolution d'une *fonction de distribution* $f(t,x,v)$, dépendant du temps, de l'espace et de la vitesse des particules. C'est l'approche utilisée pour la description des gaz raréfiés. Les équations qui en découlent sont les équations cinétiques, comme par exemple l'équation de Boltzmann. Elle s'écrit, en version adimensionnée et pour un nombre de Mach $\nu \equiv 1$:

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \frac{1}{\varepsilon} Q(f,f). \qquad (\mathcal{P}^\varepsilon)$$

Cette équation traduit que les particules de gaz sont transportées à leur vitesse v et rentrent en collisions ; le mécanisme de collision est symbolisé par l'opérateur de collisions $Q(f,f)$, pondéré par un coefficient sans dimension $1/\varepsilon$. Le paramètre ε, strictement positif, est appelé le nombre de Knudsen. Il est défini comme le rapport entre le *libre parcours moyen* et une dimension spatiale caractéristique de l'écoulement et « mesure » la fréquence de collisions des particules, donc la raréfication du gaz : plus ε est proche de 0, plus il se produit de collisions, et plus le comportement du gaz est proche de celui d'un fluide, c'est-à-dire que la vision macroscopique devient plus pertinente pour le décrire : le gaz relaxe vers un équilibre local caractérisé par :

$$Q(f,f) = 0.$$

Précisément, le régime asymptotique $\varepsilon \to 0$ dans $(\mathcal{P}^\varepsilon)$ conduit aux équations d'Euler compressibles. Concernant l'étude mathématique de cette limite hydrodynamique [5], nous renvoyons par exemple aux travaux formels de C. Cercignani [60], et à une preuve rigoureuse pour des solutions régulières de R.E. Caflisch [53].

Cette « connaissance » du lien entre le problème cinétique et le problème hydrodynamique au niveau continu va s'avérer cruciale dans le développement de méthodes numériques pour le problème cinétique. En effet, une discrétisation $(\mathcal{P}_h^\varepsilon)$ (où h représente le paramètre de discrétisation) de l'équation $(\mathcal{P}^\varepsilon)$ est d'autant plus efficace et robuste qu'elle reste stable pour *toutes* les valeurs du paramètre ε. Aussi est-il naturel de souhaiter qu'elle

[5] Il existe d'autres limites hydrodynamiques : pour un nombre de Mach $\nu = \varepsilon$, on obtient à la limite $\varepsilon \to 0$ les équations de Navier-Stokes incompressibles [21, 22, 23]

soit consistante avec la limite hydrodynamique continue ($\mathcal{P}^0$) (connue !) lorsque $\varepsilon \to 0$. L'ambition des mathématiciens dans ce contexte est donc de construire un schéma pour le modèle cinétique possédant deux propriétés :

- la stabilité par rapport au paramètre ε,
- la consistance avec le problème de l'équilibre local à la limite $\varepsilon \to 0$.

Cependant, pour réaliser ces objectifs, nous sommes soumis à des contraintes purement numériques dont nous devrons tenir compte dans la suite :

- lorsque ε tend vers 0, le terme source devient raide, ce qui motive un traitement implicite en temps pour s'affranchir de la contrainte $\Delta t = O(\varepsilon)$,
- mais le terme source est (toujours) non linéaire et non local, ce qui nécessite une attention particulière lors de sa discrétisation pour éviter un coût de calcul prohibitif de la méthode implicite.

Afin de répondre à ces problématiques, S. Jin, L. Pareschi et G. Toscani [124, 119] définissent en 1998 la notion de schémas *préservant l'asymptotique* (ou "*Asymptotic Preserving*", abrégé dans la suite par AP) : c'est la construction d'un tel schéma qui nous intéresse ici.

La notion de schéma « préservant l'asymptotique ».

Nous utilisons la définition suivante [92, 119] :

Définition 8. Un schéma numérique $\mathcal{P}_h^\varepsilon$ pour ($\mathcal{P}^\varepsilon$) est dit AP si

1. il fournit une discrétisation stable du problème ($\mathcal{P}^\varepsilon$) pour toute valeur de $\varepsilon > 0$, et lorsque ε tend vers 0, à h fixé, il conduit à un schéma $\mathcal{P}_h^0$ consistant avec le problème limite (équilibre local) $\mathcal{P}^0$;
2. les termes implicites de collisions peuvent être implémentés explicitement.

Schématiquement, il s'agit de faire commuter le diagramme suivant (sans oublier les contraintes numériques) :

$$\begin{array}{ccc} \mathcal{P}_h^\varepsilon & \xrightarrow{\varepsilon \to 0} & \mathcal{P}_h^0 \\ h \downarrow 0 & & h \downarrow 0 \\ \mathcal{P}^\varepsilon & \xrightarrow[\varepsilon \to 0]{} & \mathcal{P}^0 \end{array}$$

Les schémas AP sont aujourd'hui largement employés pour la discrétisation des équations cinétiques, dans toute sorte d'asymptotique [(6)] . Pour des problèmes de limite diffusive,

[6] Quel que soit l'adimensionnement effectué et le régime asymptotique étudié (nombre de Mach, libre parcours moyen), la philosophie « préservant l'asymptotique » reste la même.

comme le transfert radiatif, beaucoup s'y sont intéressés depuis les années 1990. A la suite de [124, 125] en effet, les auteurs, S. Jin, L. Pareschi et G. Toscani, ont participé à plusieurs collaborations. Nous pouvons citer par exemple les articles de S. Jin avec C.D. Levermore [122], F. Golse [103], P. Degond [77] et J.-G. Liu [78], puis avec F. Filbet [92, 93]. L. Pareschi et G. Toscani ont aussi collaboré, ensemble ou séparément, avec L. Gosse [106, 107], F. Filbet [89], E. Gabetta [96], G. Dimarco [81, 80]. Evoquons également les travaux récents de J.-A. Carrillo, Th. Goudon, P. Lafitte et F. Vecil [56, 57, 105] et ceux de M. Bennoune, M. Lemou et L. Mieussens [25], ainsi que l'article de F. Filbet [91] et toutes les références bibliographiques de ces papiers.
Dans [25] par exemple, M. Bennoune, M. Lemou et L. Mieussens proposent une solution basée sur une décomposition microscopique/macroscopique de l'inconnue qui a l'avantage de fournir une méthode relativement systématique pour traiter différents types d'opérateurs de collisions puisque leur décomposition utilise seulement les propriétés basiques telles que les lois de conservations et les équilibres locaux. Dans [96], E. Gabetta, L. Pareschi et G. Toscani utilisent une pénalisation de l'opérateur de collision Q par une fonction linéaire de la fonction de distribution. Plus récemment dans [92, 93], F. Filbet et S. Jin pénalisent l'opérateur de Boltzmann par celui de BGK. C'est cette technique de pénalisation que nous appliquerons à notre exemple simple (voir ci-après).

Cependant, les résultats existants sur l'étude mathématique des schémas AP pour les équations cinétiques concernent seulement des cas particuliers. Par exemple, F. Golse, S. Jin et C.D. Levermore établissent dans [103] des estimations d'erreurs ainsi que la preuve que la convergence de leur schéma est uniforme par rapport à ε pour l'équation de transfert linéaire 1D : les auteurs utilisent en particulier la connaissance de la limite diffusive de l'équation de transport lorsque le libre parcours moyen tend vers 0. Mais les équations cinétiques non linéaires restent des problèmes de complexité élevée. C'est pourquoi nous nous intéressons à un contexte simplifié dans lequel une analyse mathématique rigoureuse et complète des schémas AP peut être conduite plus facilement : les problèmes hyperboliques comportant un terme source de relaxation.

Les problèmes hyperboliques de relaxation.

Il est naturel d'étudier de tels problèmes en raison de leur complexité réduite et de leur lien avec la théorie cinétique. Citons notamment deux exemples classiques de systèmes hyperboliques (linéaires) de relaxation souvent qualifiés de modèles cinétiques à vitesses discrètes. Le système de Broadwell [48, 135] (à 3 vitesses en 1D) d'une part possède deux lois de conservation (masse et moment) et un *théorème H*. Les simulations numériques de la Section 4.6 du Chapitre 4 concerneront ce modèle. D'autre part le modèle de relaxation semi-linéaire de S. Jin et Z.P. Xin [126] est un système cinétique à deux vitesses qui s'écrit :

$$\begin{cases} \partial_t u^\varepsilon + \partial_x v^\varepsilon = 0\,, \\ \\ \partial_t v^\varepsilon + a\,\partial_x u^\varepsilon = -\dfrac{1}{\varepsilon}\,\left(v^\varepsilon - A(u^\varepsilon)\right). \end{cases} \tag{2.1}$$

Etant donné que la stratégie de discrétisation *préservant l'asymptotique* s'appuie essentiellement sur la connaissance, au niveau continu, de l'équilibre local (c'est-à-dire la limite

$\varepsilon \to 0$ pour l'exemple (2.1)), nous rappelons maintenant quelques éléments théoriques sur les systèmes hyperboliques de relaxation à travers l'exemple simple de la relaxation semi-linéaire (2.1).

Une condition nécessaire de stabilité. Comme dans le domaine cinétique, la présence d'un (unique) équilibre local et la convergence du problème de relaxation vers cet équilibre lorsque ε tend vers 0 sont soumis à des conditions de stabilité, liées à la structure mathématique des équations. Un élément crucial de la théorie de ces systèmes est une condition de stabilité, dite *condition sous-caractéristique* ou *condition de stabilité de Shizuta-Kawashima* dans le cas des systèmes en dimension plus grande que 1 [29, 129]. Considérons l'exemple (2.1) pour l'illustrer.
D'une part, ce système, observé à $\varepsilon > 0$ fixé, contient deux phénomènes qui s'affrontent. Il possède une partie strictement hyperbolique, de vitesses caractéristiques $\sqrt{a}$ et $-\sqrt{a}$. Il est donc bien connu que ses solutions peuvent développer des discontinuités. Mais un autre phénomène entre en concurrence avec le transport, la relaxation, symbolisée par le terme source et qui est un mécanisme d'autant plus rapide que le paramètre $\varepsilon > 0$ est petit. Cet autre phénomène peut conférer au système un caractère dissipatif sous certaines conditions [29]. Il est même nécessaire de lui imposer une telle propriété si l'on souhaite obtenir l'existence globale en temps de solutions régulières.
D'autre part, en regardant (2.1) sous un autre angle, nous pouvons le voir comme une perturbation de la loi de conservation hyperbolique, suivante :

$$\partial_t u + \partial_x A(u) = 0\,. \tag{2.2}$$

En effet, en prenant formellement $\varepsilon = 0$ dans (2.1), il vient :

$$v - A(u) = 0\,,$$

on s'attend donc à ce que, à la limite $\varepsilon = 0$, u soit solution de (2.2), appelé le système à l'équilibre. Or, le problème de Cauchy pour (2.2) admet, sous certaines conditions, une unique solution entropique et sa vitesse caractéristique est $\sup |A'(u)|$. Il est donc évident que l'éventuelle convergence d'une solution $(u^\varepsilon, v^\varepsilon)$ de (2.1) vers $(u, A(u))$ où u est une solution de (2.2) est soumise à une condition de stabilité nécessaire, qui doit relier les valeurs propres des deux systèmes, tous les deux de nature hyperbolique : c'est la condition sous-caractéristique, qui demande que les valeurs propres du système à l'équilibre (2.2) soient « entrelacées » entre celles du problème perturbé (2.1). Cela se traduit donc ici par :

$$\forall\, u\,, \quad |A'(u)| \,<\, \sqrt{a}\,. \tag{2.3}$$

C'est une condition nécessaire pour établir la convergence (en un sens à définir) de $(u^\varepsilon, v^\varepsilon)$ (l'unique solution de (2.1)) vers $(u, A(u))$, où u est l'unique solution entropique de (2.2). Observons formellement le lien entre cette condition et le caractère dissipatif de (2.1) lorsque ε tend vers 0. Utilisons pour cela le développement de Chapman-Enskog, qui s'écrit ici

$$v^\varepsilon \,=\, A(u^\varepsilon) + \varepsilon\, v_1^\varepsilon + O\left(\varepsilon^2\right)\,.$$

En introduisant cette expression dans le système (2.1), nous obtenons en supprimant les termes d'ordres élevés en ε la correction du premier ordre de la loi de conservation (2.2) :

$$\partial_t u + \partial_x A(u) \;=\; \varepsilon\, \partial_x \left(\beta\, \partial_x u\right) \;, \tag{2.4}$$

où $\beta \;=\; a - \big(A'(u)\big)^2$. Ainsi le caractère parabolique de (2.4) est contraint à la condition (2.3).

Remarque 9. Il est important de préciser que la condition sous-caractéristique n'est pas toujours vérifée, ce qui peut entraîner des difficultés numériques. Cela se produit lorsque le transport est *non linéaire* : citons les travaux à ce sujet de C. Mascia et R. Natalini [145] ou l'article de S. Jin et M.A. Katsoulakis [121]. En réalité, dans le cadre qui nous intéresse ici, le transport est *linéaire* : nous pouvons donc facilement vérifier cette condition de stabilité, et même en déduire l'existence d'une entropie dissipative pour le système « perturbé ».

Pour la preuve rigoureuse de la convergence vers l'équilibre pour le modèle de Jin-Xin, nous renvoyons à l'article de R. Natalini [152]. Nous en verrons une adaptation à notre système à l'Annexe B de la Partie II. La preuve proposée dans [152] se situe dans le cadre mathématique des fonctions à variations bornées (BV) et des solutions faibles avec la régularité L^1_{loc} et repose sur des estimations de compacité uniformes en ε. Les estimations utilisent en particulier une propriété intéressante du système (2.1) : la quasi-monotonie (voir [108, 170], ainsi que le rappel de la définition à l'Annexe B, p. 130). Cette propriété permet d'obtenir un *principe de comparaison* utile pour l'uniformité des estimations. Enfin, pour des exposés détaillés sur les systèmes hyperboliques de relaxation en général, nous nous référons aux travaux de G.Q. Chen, T.P. Liu et C.D. Levermore [65], à l'article fondamental de T.-P. Liu [142], ainsi qu'à l'article de revue de R. Natalini [154] et les références de ces papiers.

Discrétisation. La littérature concernant la discrétisation de systèmes hyperboliques de relaxation généraux est, comme dans le cadre cinétique, très riche ! Nous tenons cependant à distinguer deux grandes familles de stratégies numériques. D'une part, les schémas dits « de relaxation » ont pour objectif premier de traiter un système de lois de conservations (le problème à l'équilibre) en introduisant une relaxation « artificielle » : dans ce cas, nous renvoyons par exemple aux travaux d'A. Chalabi [61, 62], et Y. Qiu [63]. D'autre part, et c'est le point de vue adopté dans cette thèse, il y a les schémas *AP*, visant à traiter le problème de relaxation lui-même, pout toutes les valeurs possibles du paramètre de relaxation. Nous renvoyons une fois de plus aux travaux de S. Jin, L. Pareschi et G. Toscani [124, 125], mais également ceux de S. Jin [120] avec C.D. Levermore [123], Z. Xin [126] et F. Filbet [92], ou encore les articles de D. Aregba-Driollet et R. Natalini [8] ou de L. Pareschi et G. Russo [158], ainsi que leurs références bibliographiques.
Cependant, même si l'efficacité de ces méthodes est aujourd'hui largement illustrée par les simulations numériques, peu de travaux proposent une analyse mathématique des schémas construits. D. Aregba-Driollet et R. Natalini [8] proposent et analysent un schéma AP pour le système de Jin et Xin (2.1). Les auteurs y adaptent les arguments de [152] au

niveau discret. A. Chalabi [61, 62, 63] obtient également la convergence de schémas semi-implicites de relaxation pour des lois de conservation scalaires ou des systèmes avec terme source quelconque pouvant être raide. Par ailleurs, F. Filbet et S. Jin, dans [92], appliquent leur méthode AP à un système hyperbolique non linéaire de relaxation et établissent des estimations sur le schéma semi-discret en temps.
Ainsi, à notre connaissance, les études théoriques concernent des systèmes très particuliers (par exemple (2.1)) ou sont partielles, montrant la convergence des schémas sans aborder précisément la question de la consistance avec le problème limite, c'est-à-dire la propriété de *préserver l'asymptotique* dans la définition 8. Parmi la profusion des travaux sur ce sujet, il apparaît toutefois quelques traits caractéristiques communs aux différentes stratégies de discrétisation employées. Nous citons ici trois propriétés que nous retiendrons pour construire notre schéma :

- La structure hyperbolique du problème étudié nous offre, comme nous l'avons déjà évoqué à la Section 1 de l'introduction, un cadre privilégié de méthodes numériques, celui des volumes finis [133, 134].

- Le terme source pouvant devenir raide et engendrer des contraintes numériques, nous adopterons, comme dans les articles précédemment cités, une stratégie de *splitting*. Précisément, il s'agit de traiter la partie transport lors d'une première étape (qui peut donc être explicite), puis de discrétiser la partie relaxation de manière implicite ou semi-implicite.

- Enfin, pour obtenir des estimations d'erreurs et la convergence pour notre schéma (voir la remarque 2), nous utiliserons pour le transport un flux *Total Variation Diminishing* (ou TVD) [133], c'est-à-dire qui fait diminuer la variation totale de la solution numérique. Notons que la variation totale est la version discrète de la semi-norme des fonctions à variations bornées : ce sera un outil crucial à la fois pour assurer la stabilité du schéma, mais également pour montrer la convergence au niveau discret vers l'équilibre local en adaptant les arguments du niveau continu.

Dans la Partie II, nous tentons de fournir une étude complète d'un schéma AP (possédant les trois propriétés ci-dessus) pour un modèle cinétique à deux vitesses généralisant le système de Jin-Xin (2.1). La description du modèle ainsi que le résumé des travaux fait l'objet de la section suivante.

2.2 Travaux effectués : résultats de convergences pour un modèle simple

L'analyse que nous proposons dans le Chapitre 4 concerne un modèle généralisant le modèle de Jin et Xin (2.1) qui s'écrit :

$$\begin{cases} \partial_t u^\varepsilon + \partial_x v^\varepsilon = 0\,, \\ \\ \partial_t v^\varepsilon + a\,\partial_x u^\varepsilon = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u^\varepsilon, v^\varepsilon), \end{cases} \tag{2.5}$$

où le paramètre de relaxation ε joue le rôle du nombre de Knudsen en théorie cinétique, tandis que le terme source $\mathcal{R}$ est l'analogue de l'opérateur de collision de l'équation de Boltzmann : nous le souhaitons le plus général possible. Ainsi, nous considèrerons une fonction non linéaire et possédant comme son analogue cinétique un unique équilibre local [7], à savoir :

$$\mathcal{R}(u,v) = 0 \;\Leftrightarrow\; v = A(u)\,. \tag{2.6}$$

Remarquons que l'équilibre local est le même que celui de (2.1), le problème de l'équilibre est donc aussi (2.2).

Stratégie de discrétisation. Comme annoncé précédemment, nous construisons un schéma de splitting dont la première étape traite la partie transport par un schéma de Lax-Friedrichs. Ensuite, la deuxième étape consiste à discrétiser le système différentiel ordinaire sur $(t^*; t^{n+1})$:

$$\left\{\begin{array}{l} \partial_t u = 0\,, \\ \\ \partial_t v = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u,v), \\ \\ u(t^*) = u^*\,, \quad v(t^*) = v^*\,, \end{array}\right. \tag{2.7}$$

où (u^*, v^*) représente la solution de l'étape de transport (nous avons enlevé les exposants ε par souci de clarté). Pour cette étape, nous utilisons une technique de pénalisation comme dans [96, 92]. Plus précisément, dans le même esprit que dans [92], où F. Filbet et S. Jin pénalisent l'opérateur de Boltzmann par l'opérateur BGK, nous pénalisons l'opérateur $\mathcal{R}$ avec l'opérateur semi-linéaire de Jin-Xin, à savoir :

$$\mathcal{L}(u,v) := v - A(u)\,.$$

En effet, en introduisant cette fonction dans (2.7), il vient :

$$\left\{\begin{array}{l} \partial_t u = 0\,, \\ \\ \partial_t v + \dfrac{1}{\varepsilon}\,v = -\dfrac{1}{\varepsilon}\,\big(\mathcal{R}(u,v) - \mathcal{L}(u,v)\big) + \dfrac{1}{\varepsilon}\,A(u), \\ \\ u(t^*) = u^*\,, \quad v(t^*) = v^*\,. \end{array}\right. \tag{2.8}$$

En utilisant cette formulation, nous pouvons donc intégrer exactement le membre de gauche de l'équation sur v, puis traiter de manière implicite seulement le dernier terme raide restant, $A(u)/\varepsilon$, qui a le bon goût de se calculer *explicitement* puisque sur l'intervalle (t^*, t^{n+1}) la fonction u est constante ! Ce schéma $\mathcal{P}_h^\varepsilon$ ainsi que sa version « relaxée » $\mathcal{P}_h^0$ sont décrits précisément à la section 4.2.1 p. 98 du Chapitre 4.

[7] Cela revient à demander à $\mathcal{R}$ de satisfaire le théorème des fonctions implicites.

Résultats théoriques. Pour les schémas ainsi construits, nous obtenons différents résultats de convergences, ainsi que des estimations d'erreurs qui justifient le diagramme commutatif de relaxation pour notre modèle. Ils sont énoncés à la section 4.2.2 et résumés ci-dessous.
Notons $h = (\Delta x, \Delta t)$ le paramètre de discrétisation. Le couple $(u^{\varepsilon,n}; v^{\varepsilon,n})$ désigne la solution du schéma $\mathcal{P}_h^\varepsilon$ et la déviation par rapport à l'équilibre local est donnée par :

$$\delta^{\varepsilon,n} \,=\, v^{\varepsilon,n} - A\left(u^{\varepsilon,n}\right) .$$

Théorème 10.

Sous certaines conditions sur le terme source, ainsi que la condition sous-caractéristique, les inégalités suivantes sont vérifiées.

(i) Contrôle de la déviation par rapport à l'équilibre :

$$\begin{cases} \|\delta^{\varepsilon,n}\|_1 \,\leq\, C\,\varepsilon + \left\|\delta^{\varepsilon,0}\right\|_1 \; e^{-\beta_0\, t^n/\varepsilon} & \forall n \geq 0,\, \varepsilon > 0, \\ \|\delta^{\varepsilon,n}\|_1 \,\leq\, e^{-\beta_0\, t^n/\varepsilon} \; \left\|\delta^{\varepsilon,0}\right\|_1 \,+\, C\,\Delta t\, e^{-\beta_0 \Delta t/\varepsilon} & si\, \varepsilon < \Delta t. \end{cases}$$

(ii) Convergence du schéma (non uniforme en ε) :

$$\int_{\mathbb{R}} |u_h^\varepsilon(t,x) - u^\varepsilon(t,x)| \,+\, |v_h^\varepsilon(t,x) - v^\varepsilon(t,x)|\, dx \\ \leq \frac{C}{\varepsilon}\left(\Delta t\left(\frac{\left\|\delta^{\varepsilon,0}\right\|_{L^1}}{\varepsilon} \,+\, 1\right) \,+\, \Delta x^{1/2}\right).$$

(iii) Consistance avec le problème limite (asymptotique $\varepsilon \to 0$) :

$$\|u_h^\varepsilon(t) - u_h(t)\|_1 + \|v_h^\varepsilon(t) - v_h(t)\|_1 \,\leq\, C_t\, e^{-\beta_0 \Delta t/\varepsilon} \,\left[1 + \left\|\delta^{0,0}\right\|_1\right].$$

Dans les preuves de ces estimations, nous utilisons tous les outils précédemment mentionnés. Nous imposons en particulier un certain nombre de contraintes sur le terme source afin de vérifier la condition de stabilité sous-caractéristique (section 4.2 p. 96). Nous établissons ensuite des estimations *a priori* dans L^∞ et BV, en s'appuyant sur un principe de comparaison comme dans le cas continu (section 4.3 p. 101). Les résultats de compacités ainsi établis permettent de montrer la consistance du schéma avec le problème limite (section 4.4 p. 106). Nous terminons le Chapitre 4 par l'analyse des erreurs de consistance en utilisant la formule des caractéristiques et en différenciant à chaque étape l'erreur venant de la partie transport de celle venant de la partie relaxation, ce qui nous permet de montrer la convergence du schéma (section 4.5, p. 110). Enfin, nous présentons à l'annexe B la preuve du théorème 1.1 (p. 95), c'est-à-dire le résultat de convergence vers l'équilibre local au niveau continu. Nous adaptons à notre cas les arguments de R. Natalini dans [152].

3 Partie III : un modèle d'écoulement sanguin

Cette dernière section introduit et présente la Partie III de la thèse. Ce travail en collaboration avec V. Milišić et K. Pichon Gostaf [150] a été initié au CEMRACS 2009, dans le cadre du projet *RUGOSITY*, proposé par E. Bonnetier, D. Bresch et V. Milišić. Comme dans la première partie, nous nous intéressons à un problème d'écoulement de fluide, dans une description macroscopique. Cependant le modèle qui nous préoccupe provient cette fois d'une problématique médicale : nous souhaitons mettre en évidence l'influence de la pose d'un stent (aussi appelé endoprothèse ou tuteur vasculaire) dans une artère sur l'écoulement sanguin. Nous commençons donc cette section en décrivant le contexte médical dans lequel intervient ce dispositif. Puis nous soulevons la question de la modélisation mathématique. Un aspect majeur de cette modélisation est la présence de deux échelles spatiales dans le problème. En effet, la petite taille du stent (disons ε) constitue une échelle « microscopique » relativement à la taille de l'artère, c'est-à-dire le domaine « macroscopique ». Nous décrivons alors les difficultés mathématiques et numériques liées à cet aspect multi-échelles et quelques moyens de les surmonter, à savoir la dérivation de lois de parois. Enfin, nous évoquons quelques résultats existants avant de résumer les contributions du Chapitre 5.

3.1 Motivation médicale

L'étude de la circulation sanguine dans le corps (hémodynamique) intéresse de nombreux scientifiques dès le $XVII$ème siècle. Le physicien et médecin J.L.M. Poiseuille [162] propose au milieu du XIXème siècle l'un des premiers modèles d'écoulement sanguin dans les artères et met en évidence un profil caractéristique laminaire et permanent d'écoulement : le *profil de Poiseuille*, sur lequel nous reviendrons plus tard.
Il existe plusieurs phénomènes pouvant affecter l'écoulement du sang ; celui qui nous préoccupe ici est *l'anévrisme*. C'est une dilatation localisée de l'artère, qui peut être provoquée par le dépôt de graisse dans l'artère (athérome). Cela engendre la formation d'une poche de taille variable sur l'artère, appelée *sac* (voir la Figure 1.6) Cette pathologie n'est pas sans risque. En effet, l'écoulement sanguin aux abords d'un sac anévrismal devient turbulent et peut engendrer la création d'un caillot ; en outre, une rupture d'anévrisme peut entraîner un accident vasculaire cérébral (AVC).
Une possibilité de traitement de l'athérome ou de l'anévrisme (non rompu) est une intervention chirurgicale, l'angioplastie, illustrée à la Figure 1.4 [8] : elle consiste à insérer un dispositif métallique maillé et tubulaire (le *stent* ou *tuteur vasculaire*) dans l'artère, afin de repousser contre la paroi ce qui obstrue le tube et empêche le sang de s'écouler normalement.
Plusieurs types de ces dispositifs métalliques existent, nous pouvons en observer deux exemples particuliers à la Figure 1.5 [9].

[8] Cette image provient de la page wikipédia **http ://fr.wikipedia.org/wiki/Stent**.
[9] Ce sont des modèles conçus et commercialisés par l'entreprise Cardiatis **www.cardiatis.com**.

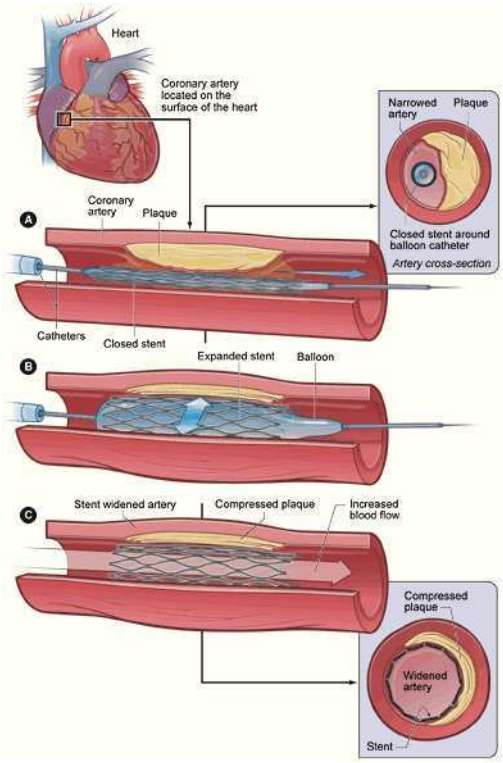


FIGURE 1.4 – *Angioplastie.*

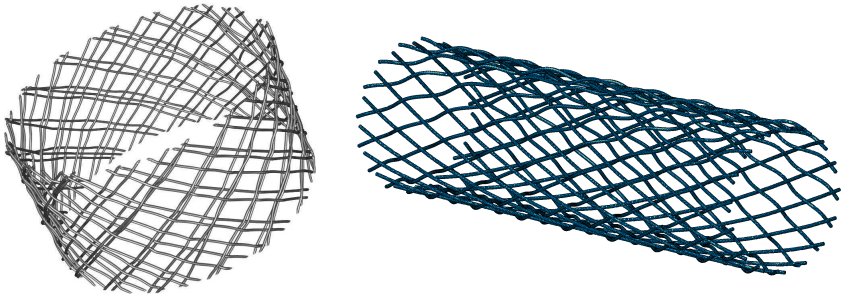


FIGURE 1.5 – *Deux modèles de stents.*

D'un point de vue expérimental, la pose d'un stent semble permettre de ralentir l'évolution (le grossissement) des sacs d'anévrisme, tout en y laissant circuler le sang et diminuant les effets turbulents de l'écoulement à leurs abords [16, 141]. Notre objectif est alors de fournir un modèle capable de rendre compte, par des simulations numériques, de telles observations. Or s'il existe de nombreux modèles mathématiques d'écoulement sanguin dans des artères sans stent, la présence de cette *rugosité* dans le domaine d'écoulement engendre des difficultés supplémentaires, aux niveaux mathématique et numérique, liées à

l'aspect multi-échelles du problème. En effet, la petite taille du stent constitue une échelle « microscopique » qui vient s'ajouter à l'échelle « macroscopique » globale de l'artère.
Donnons-nous quelques ordres de grandeurs des différentes échelles spatiales du problème :

- diamètre de l'artère fémorale :$\emptyset_{\mathrm{A}} = 6mm$
- épaisseur totale du stent : $\varepsilon = 0.25mm$
- épaisseur d'une spire du stent : $\varepsilon = 0.04mm$
- diamètre d'un globule rouge : $\emptyset_{\mathrm{RC}} = 0.008mm$.

Aussi sera-t-il fondamental de prendre en compte correctement la rugosité du domaine dans les équations afin d'étudier l'influence effective du stent. En outre, d'un point de vue numérique, nous devrons envisager une alternative au maillage direct du domaine rugueux, trop coûteux.

3.2 Modélisation mathématique : état de l'art

Commençons par remarquer que la situation qui nous intéresse ici appartient à une famille de problèmes mathématiques largement étudiés : les problèmes rugueux en mécanique des fluides. La littérature concernant l'étude des effets d'une rugosité sur l'écoulement d'un fluide est extrêmement riche. Nous pouvons citer par exemple les travaux d'Y. Achdou, O. Pironneau, F. Valentin, et P. Le Tallec [2, 3, 4], ceux d'Y. Amirat et J. Simon [6] ou encore ceux de D. Gérard-Varet et A. Basson [24], ainsi que les références de ces articles. Dans ce contexte, il apparaît un phénomène de *couche limite* au voisinage de la bordure rugueuse qui modifie le comportement du fluide. Cela constitue essentiellement un enjeu numérique, car en général, les grilles de discrétisations ne sont pas assez fines pour capturer correctement les rugosités de taille ε (très petit) du domaine. Une possibilité pour surmonter cette difficulté numérique est une modification des équations au niveau continu, à savoir la dérivation de *lois de parois*. Il s'agit de conditions aux limites artificielles sur le bord d'un domaine fictif « lisse » , à l'intérieur du domaine rugueux. Alors les effets de la rugosité seront contenues dans de nouvelles équations, elles-mêmes posées dans un domaine géométrique lisse, facile à discrétiser.
L'obtention de lois de parois constitue l'objectif principal du présent travail. Avant cela, il est nécessaire d'établir un modèle mathématique adapté à notre contexte particulier. Pour cela, nous devrons évoquer la question de la modélisation pour trois aspects du problème :

- l'artère, avec ou sans stent (domaine lisse ou rugueux),
- le fluide et ses propriétés,
- le régime de l'écoulement.

Précisons les hypothèses et notations pour la modélisation de notre domaine rugueux, l'artère stentée. Comme dit précédemment, il convient d'en décrire l'aspect double échelle, grâce au paramètre $\varepsilon > 0$, la taille caractéristique de la rugosité (le stent). Nous choisissons, pour simplifier, de ne pas prendre en compte l'élasticité des parois des artères et

de considérer des géométries 2D [10], qui peuvent cependant être vues comme des coupes longitudinales d'artères 3D (il est raisonnable de considérer un écoulement à symétrie cylindrique). La variable x_1 désigne la direction horizontale, x_2 la direction verticale.
Le type de stent que nous modélisons est ouvert des deux côtés, comme à la Figure 1.5 [11], tandis que nous nous contenterons de trois géométries simples d'artères, à savoir :

- un tube horizontal aux parois rigides, dans lequel l'écoulement se fait de la gauche vers la droite.
- un domaine d'écoulement principal comme ci-dessus, avec une bifurcation avec une artère collatérale verticale (Figure 1.6 à gauche),
- un domaine d'écoulement principal comme ci-dessus, avec cette fois la présence d'un sac d'anévrisme au-dessous (Figure 1.6 à droite).

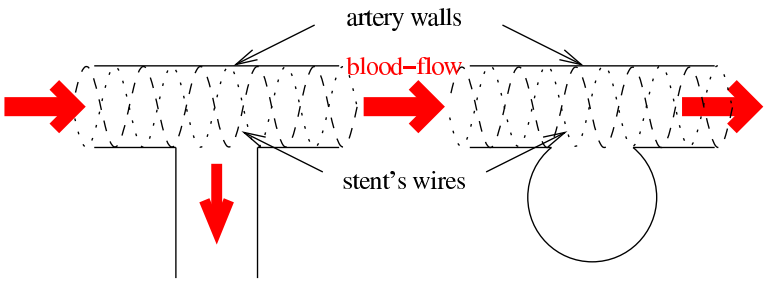


FIGURE 1.6 – *Deux géométries simples d'artères avec stent : avec une artère collatérale (gauche) ; avec un sac anévrismal (droite).*

Le stent est modélisé par le graphe d'une fonction périodique ou quasi-périodique, ou encore par une succession périodique de petites billes circulaires de taille ε (cela constitue le bord rugueux Γ_ε du domaine).
Nous utilisons les mêmes notations pour désigner les domaines, que ce soit dans le cas d'une seule artère ou d'une bifurcation, comme illustré aux Figures 1.7 et 1.8 : Ω_0 désigne le domaine macroscopique lisse (l'artère sans le stent, c'est le domaine fictif), Ω_ε est le domaine rugueux (aussi macroscopique), où ε représente la taille caractéristique du stent. En outre, le schéma de droite des Figures 1.7 et 1.8 résulte d'un zoom sur une cellule microscopique $Z^+ \cup P$, c'est-à-dire par passage de la variable spatiale « lente » x à la variable « rapide » $y = x/\varepsilon$.

[10] Les simulations de l'Annexe C concernent cependant des artères 3D.

[11] Il existe des modèles de stents fermés à une extrémité, ce qui peut se modéliser par une interface poreuse entre deux « boîtes » , comme dans les modèles étudiés par M.A. Fernández, J.-F. Gerbeau et V. Martin [86] (problème de Stokes), puis les mêmes auteurs avec A. Caiazzo [55] (problème de Navier-Stokes complet). Ces modèles de stents sont plutôt utilisés pour traiter les anévrismes intra-craniens.

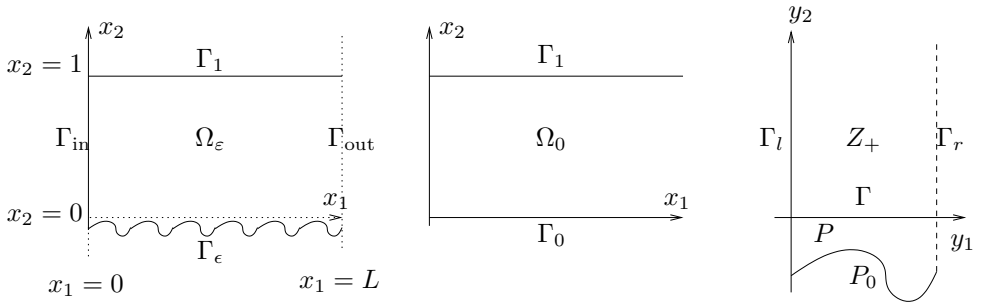


FIGURE 1.7 – *Domaine rugueux, domaine lisse, cellule microscopique (cas de la géométrie simple : une seule artère)*

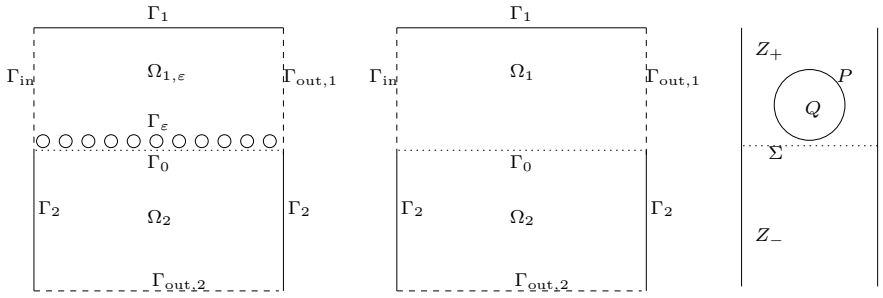


FIGURE 1.8 – *Domaine rugueux, domaine lisse, cellule microscopique (cas d'une bifurcation entre l'artère principale horizontale et une autre artère ou un sac anévrismal au-dessous)*

Dans ces géométries simplifiées, nous devons maintenant modéliser le sang et le régime d'écoulement dans lequel il se trouve. Là encore, de nombreux paramètres et propriétés biologiques sont en jeu et il convient de cibler correctement ceux que nous voulons effectivement représenter. Rappelons que l'objectif principal de ce travail est de fournir un modèle traduisant les effets du stent sur l'écoulement sanguin, tout en évitant un surcoût numérique dû au maillage de la rugosité. Nous nous contentons donc de présenter des modèles très simples. En particulier, dans les travaux que nous citons ci-après, nous employons une description macroscopique du problème :

- le sang est considéré comme un fluide newtonien, incompressible, visqueux et homogène [12] ;
- et l'écoulement du fluide est décrit par des équations dérivées du problème de Navier-Stokes complet, suite à différentes hypothèses simplificatrices (notamment la station-

[12] Nous avons bien conscience que c'est une hypothèse extrêmement simplificatrice de ne pas considérer le caractère fondamentalement non newtonien et non homogène du sang (mélange de plasma et de globules rouges). A défaut d'une réelle justification, la raison en est que notre objectif principal est la modélisation de la rugosité.

narité ou la linéarisation). Plusieurs types de conditions aux limites sont considérées. Nous nous restreindrons également à un problème scalaire, n'étudiant que le profil de la vitesse axiale $u_\varepsilon \in \mathbb{R}$ (et la pression p_ε).

Méthodologie générale d'obtention de lois de parois. Partant du système de Navier-Stokes dans le domaine rugueux par exemple, complété par des conditions aux limites *ad hoc*, la dérivation d'une loi de paroi dépend évidemment du type de condition au bord imposée (glissement, non-glissement, *etc*), mais on peut en dégager une méthodologie générale [99, 100], inspirée des techniques classiques d'homogénéisation [68, 168]. Cette théorie, a été introduite initialement pour décrire le comportement des matériaux composites, caractérisés par le fait de contenir des hétérogénéités, petites en comparaison avec la dimension globale du matériau : pour étudier l'influence au niveau macroscopique des hétérogénéités microscopiques, on introduit dans le modèle un petit paramètre $\varepsilon > 0$ représentant l'échelle des impuretés, puis on tente d'évaluer des approximations à divers ordres de la solution lorsque ε tend vers 0. Il s'agit de formuler un ansatz de développement asymptotique (type Chapman-Enskog) de la solution sous la forme :

$$u_\varepsilon(x) \;=\; u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon\, u_1\left(x, \frac{x}{\varepsilon}\right) + \dots\,, \tag{3.1}$$

pour injecter cela dans les équations, corriger les conditions aux limites et enfin rassembler les termes d'ordres égaux relativement à ε. Schématiquement, voici les différentes étapes qui nous intéressent :

1. point de départ : problème scalaire d'inconnues u_ε et p_ε dans le domaine rugueux Ω_ε. Formulation d'un ansatz du type (3.1) dans le domaine rugueux.

2. Approximation macroscopique d'ordre 0, (u_0, p_0), **dans le domaine lisse** Ω_0 (!) : ici, nous souhaitons récupérer le profil de Poiseuille, c'est-à-dire un profil parabolique de vitesse en la variable verticale x_2, typique de l'écoulement d'un fluide visqueux dans un tube droit en régime permanent et laminaire [162].

3. Approximation d'ordre 1 multi-échelles ($\mathcal{U}_\varepsilon$, $\mathcal{P}_\varepsilon$) **dans le domaine rugueux** Ω_ε, qui nécessite en particulier :

   - un choix de prolongement de (u^0, p_0) au domaine rugueux (nous verrons qu'il en existe plusieurs),
   - puis, suivant ce choix, l'introduction de correcteurs de couche limite (CCL) au niveau microscopique (β, π) (dans la cellule microscopique, en variable rapide $y = x/\varepsilon$), afin de corriger les erreurs commises à la fois sur la partie $\Omega_\varepsilon \backslash \Omega_0$, ainsi que sur les conditions au bord rugueux Γ_ε,
   - et enfin une correction macroscopique dans tout le domaine rugueux Ω_ε.

4. Enfin, on se ramène à une approximation dans le domaine lisse, en moyennant l'ansatz, afin de supprimer les oscillations microscopiques. On obtient ainsi $(\overline{\mathbf{u}}_\epsilon, \overline{p}_\epsilon)$, solution d'un système dans le domaine lisse Ω_0, complété par une loi de paroi implicite

sur la frontière Γ_0 contenant l'information de la rugosité. Suivant les conditions aux limites choisies initialement, cette loi de paroi possède différente dénominations : Navier, Beavers, Joseph, *etc* (voir par exemple [115]).

Un état de l'art de modèles d'artères avec stent. A notre connaissance, les premiers travaux proposant une analyse rigoureuse de lois de parois concernent le problème de Poisson avec des conditions aux limites homogènes de Dirichlet sur tout le bord d'un domaine rugueux : les articles d'Y. Achdou, O. Pironneau [1] avec F. Valentin [4] et P. Le Tallec [3]. Par ailleurs, W. Jäger et A. Mikelić [112, 114, 115] se sont intéressés au contact entre un fluide visqueux et un milieu poreux, dont la modélisation géométrique peut s'apparenter à l'illustration de la Figure 1.8. Les auteurs y étudient le même type de conditions aux limites que dans les travaux précédemment cités ; cependant les techniques de prolongement de la solution à l'ordre 0 du domaine lisse au domaine rugueux diffèrent (nous utiliserons à la Partie III la stratégie d'Y. Achdou, qui consiste en un prolongement *linéaire*, et non constant, dans la partie rugueuse). Pour autant, les deux stratégies conduisent aux mêmes lois de parois implicites moyennées. En effet, dans [39, 40], D. Bresch et V. Milišić unifient les deux approches, dérivent des lois de parois et établissent des estimations d'erreurs pour un modèle d'artère avec stent avec la géométrie de la Figure 1.7, périodique. Ils considèrent un problème de Poisson particulier pour la composante axiale $u_\varepsilon \in \mathbb{R}$ de la vitesse du fluide, avec des conditions sur les bords Γ_ε et Γ_1 de type Dirichlet, ainsi que des conditions entrantes et sortantes périodiques en vitesse sur les bords latéraux Γ_{in} et Γ_{out}. Le cas d'un écoulement dirigé en pression (plus réaliste dans le contexte de l'écoulement sanguin) est aussi étudié par W. Jäger et A. Mikelić [116, 117], toujours pour un écoulement de type Poiseuille. Pour le contexte des artères avec stent, ce sont les articles de D. Bresch, V. Milišić et E. Bonnetier [33, 34] qui sont à la base des travaux de la Partie III. Les auteurs adaptent leurs résultats précédents au cas non périodique, avec des conditions aux limites latérales de type Neumann. La présence de conditions de Neumann empêche la généralisation immédiate des résultats pour les conditions de Dirichlet et nécessite des estimations plus délicates, dites *très faibles* [155]. Les estimations *a priori* sont améliorées ensuite par V. Milišić [147]. Evoquons enfin que des résultats existent sur des rugosités aléatoires [24]. Pour la géométrie avec bifurcation (Figure 1.8), nous renvoyons à l'article de V. Milišić [148], dans lequel l'auteur étudie un problème de Stokes, toujours stationnaire et dirigé en pression.

Dans tous les travaux précédemment cités, le régime d'écoulement étudié est stationnaire. Bien que ces éléments bibliographiques soient loin d'être exhaustifs, peu de travaux concernent des écoulements instationnaires (citons néanmoins [113, 67] par exemple) et l'objectif principal de la Partie III est de généraliser les résultats de [40] au cas d'un problème de Stokes instationnaire, dirigé en pression.

3.3 Présentation des résultats

Le travail présenté dans la Partie III [150] généralise les résultats obtenus dans [40] et [148] au cas d'un problème de Stokes instationnaire, où l'écoulement est dirigé en pression.

L'idée est de se diriger vers un modèle d'écoulement plus réaliste, à savoir :

- une partie permanente établie : profil de Poiseuille,
- plus une perturbation périodique en temps due au pouls : profil de Wommersley.

Précisément, nous considérons la géométrie plus simple de [40], illustrée à la Figure 1.7. Le problème de départ s'énonce comme suit.
Trouver $\mathbf{u}_\varepsilon$ tel que :

$$\begin{cases} \partial_t \mathbf{u}_\varepsilon - \Delta \mathbf{u}_\varepsilon + \nabla p_\varepsilon = 0 \quad \text{dans } \Omega_\varepsilon \\ \text{div } \mathbf{u}_\varepsilon = 0 \quad \text{dans } \Omega_\varepsilon \\ \mathbf{u}_\varepsilon = 0 \quad \text{sur } \Gamma_1 \cup \Gamma_\varepsilon \\ p_\varepsilon = p_{\text{in}}(t) \quad \text{sur } \Gamma_{\text{in}} \\ p_\varepsilon = p_{\text{out}} = 0 \quad \text{sur } \Gamma_{\text{out}} \\ \mathbf{u}_\varepsilon \text{ est } x_1\text{-periodique} \end{cases} \tag{3.2}$$

On suppose que la pression est périodique en temps et ne dépend que de x_1 en espace : on peut se restreindre à la résolution d'un problème scalaire sur la vitesse horizontale u_ε. Avant de suivre les étapes classiques pour l'obtention de lois de paroi (présentées brièvement à la section précédente), le point de départ est un passage en série de Fourier en temps, possible grâce aux hypothèses de périodicités faites. Nous sommes ainsi ramenés à résoudre le problème suivant, pour tout mode $k \in \mathbb{Z}^*$: trouver $\hat{u}_{\epsilon,k}$ tel que

$$\begin{cases} (ik - \Delta)\, \hat{u}_{\epsilon,k} = \hat{C}_k & \text{dans } \Omega_\varepsilon\,, \\ \hat{u}_{\epsilon,k} = 0 & \text{sur } \Gamma_\varepsilon \cup \Gamma_1\,, \\ \hat{u}_{\epsilon,k} \ x_1 - \text{périodique} & \text{sur } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{3.3}$$

où $\hat{C}_k$ est le k-ième coefficient de Fourier de la pression entrante. Les étapes suivantes, à savoir ansatz sur le développement de la solution, approximation d'ordre 0, $\hat{u}_{0,k}$, introduction de correcteurs, approximation d'ordre 1, $\hat{\mathcal{U}}_{\epsilon,k}$ (contenant la variable lente et la variable rapide) dans le domaine rugueux, puis sa moyennisation par rapport à la variable rapide $\overline{u}_{\varepsilon,k}$, et enfin la solution $\hat{V}_{\varepsilon,k}$ de la loi de paroi implicite (de type Navier) dans le domaine lisse sont plus ou moins conservées. Il est toutefois intéressant de souligner deux faits :

- les modes de Fourier n'intéragissent pas avec la fréquence d'oscillation de la rugosité (heureusement !) ;
- et dans ce cadre de modèle jouet, on obtient toutes les estimations, à chaque étape du processus d'approximation, de manière directe, sans invoquer de résultats théoriques

abstraits : par exemple, les solutions très faibles *à la Nečas* [155] sont calculées ici explicitement, grâce à des hypothèses simplificatrices telles que la forme simple du domaine $\Omega_0 = [0,1]^2$ et à l'existence, notamment, d'une base hilbertienne de $L^2(\Omega_0)$.

Nous résumons dans la proposition suivante les résultats obtenus au Chapitre 5 (estimations d'erreurs successives et loi de paroi).

Proposition 11.
On a les estimations suivantes :

$$\|\hat{u}_{\epsilon,k} - \hat{u}_{0,k}\|_{H^1(\Omega_\varepsilon)} \leq c_1 \sqrt{\epsilon} \ , \quad \|\hat{u}_{\epsilon,k} - \hat{u}_{0,k}\|_{L^2(\Omega_0)} \leq c_2 \, \epsilon \, . \tag{3.4}$$

$$\|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\|_{H^1(\Omega_\varepsilon)} \leq c_3 \, \epsilon \ , \quad \|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\|_{L^2(\Omega_0)} \leq c_4 \, \epsilon^{3/2} \, . \tag{3.5}$$

$$\|\hat{u}_{\epsilon,k} - \overline{u}_{\varepsilon,k}\|_{L^2(\Omega_0)} \leq c_5 \, \epsilon^{3/2} \, . \tag{3.6}$$

Enfin, le problème d'approximation macroscopique dans le domaine lisse s'écrit :

$$\begin{cases} (ik - \Delta)\hat{V}_{\varepsilon,k} = \hat{C}_k & \text{dans } \Omega_0 \, , \\ \hat{V}_{\varepsilon,k} = 0 & \text{sur } \Gamma_1 \, , \\ \hat{V}_{\varepsilon,k} = \varepsilon \, \overline{\beta} \, \frac{\partial \hat{V}_{\varepsilon,k}}{\partial x_2} & \text{sur } \Gamma_0 \, , \\ \hat{V}_{\varepsilon,k} \text{ est } x_1\text{-periodique} & \text{sur } \Gamma_{\text{in}} \cup \Gamma_{\text{out}} \, . \end{cases} \tag{3.7}$$

De plus, on a les estimations d'erreurs suivantes :

$$\|\hat{u}_{\epsilon,k} - \hat{V}_{\varepsilon,k}\|_{L^2(\Omega_0)} \leq c_6 \, \epsilon^{3/2} \ \text{et} \quad \left\|\hat{u}_{\epsilon,k} - \hat{V}_{\varepsilon,k}\right\|_{H^1(\Omega_0)} \leq c_7 \sqrt{\epsilon}. \tag{3.8}$$

Toutes les constantes sont indépendantes du mode de Fourier k considéré.

La dernière section du Chapitre 5 vérifie numériquement les ordres d'approximations établis précédemment. Enfin, nous présentons en Annexe C des simulations directes d'une artère 3D avec un sac d'anévrisme, avec ou sans stent, afin de comparer la différence de coût de discrétisation suivant la taille ε du stent.

4 Conclusions, perspectives et travaux en cours

Dans cette section, nous commentons d'abord brièvement les résultats de la partie II et présentons une piste de recherche que nous explorons actuellement autour de ce schéma AP. Puis nous proposons d'autres perspectives de recherche dans le contexte de la dynamique des fluides à surface libre.

4.1 Sur l'analyse du schéma AP pour le système de Broadwell

Comme dit précédemment, le travail de la partie II trouve sa motivation initiale dans le domaine de la théorie cinétique des gaz. Nous tentons donc actuellement de conduire une analyse de notre schéma AP pour des modèles plus physiques que le modèle jouet à deux vitesses, tels que le modèle de Broadwell ou plus généralement des modèles cinétiques possédant un nombre fini quelconque de vitesses. Malheureusement, il apparaît très vite que les techniques employées pour l'analyse du schéma dans [94] (estimations uniformes dans L^∞ et BV) ne peuvent pas s'étendre au système de Broadwell. En effet, toutes les estimations *a priori* reposent fortement sur un principe de comparaison pour les systèmes quasi-linéaires possédant une propriété de *monotonie*. Le système de Broadwell ne bénéficie pas de cette propriété ! Il convient donc d'adopter une nouvelle approche pour traiter ce problème, en l'abordant sous un angle réellement cinétique. Nous avons à l'esprit les résultats de convergence vers l'équilibre hydrodynamique de l'équation de Boltzmann *via* des méthodes de dissipation d'entropie (voir par exemple l'ouvrage de C. Villani [177] et ses références). Afin de donner une idée générale de la méthodologie et des techniques que nous souhaitons appliquer à notre schéma numérique avec F. Filbet, faisons quelques commentaires sur le problème de Broadwell et citons quelques références bibliographiques. Pour l'essentiel, nous nous sommes basés sur l'ouvrage de C. Villani [177] pour les considérations générales, puis sur le cours de H. Cabannes, R. Gatignol et L.-S. Luo sur les modèles cinétiques à vitesses discrètes [52] et l'article de F. Berthelin, A.E. Tzavaras et A. Vasseur [27], ainsi que les références de ces travaux.

Deux formulations du système. Tout d'abord, la formulation « physique » du problème de Broadwell [48] est une version simplifiée de l'équation de Boltzmann. Précisément, le système rend compte de l'évolution au cours du temps de fonctions de distribution de particules f_i d'un gaz soumis à deux mécanismes : le transport des particules, à leurs vitesses v_i, et leurs éventuelles collisions modélisées par un opérateur de collision quadratique Q_i. Dans le cas de Broadwell, ou dans tout modèle dit « cinétique à vitesses discrètes » la simplification vient du fait que l'espace des vitesses des particules est discret (et borné dans notre cas), ce qui simplifie également grandement l'opérateur de collision. Dans une version 1D, nous partons ici du modèle à 4 vitesses, $+1$, 0 et -1 (deux types de particules se transportant à la vitesse nulle) :

$$\left\{ \begin{array}{lcl} \partial_t f_+ + \partial_x f_+ &=& \frac{1}{\varepsilon} Q_+ \,, \\ \partial_t f_0 &=& \frac{1}{\varepsilon} Q_0 \,, \\ \partial_t f_- - \partial_x f_- &=& \frac{1}{\varepsilon} Q_- \,, \end{array} \right. \tag{4.1}$$

où l'opérateur de collision est donné par :

$$\left\{ \begin{array}{l} Q_+ \,=\, f_0^2 - f_+ f_- \,, \\ Q_0 \,=\, -Q_+ \,, \\ Q_- \,=\, Q_+ \,. \end{array} \right.$$

La formulation « fluide » s'interprète alors comme le système composé des deux lois de conservation associées aux moments d'ordres 0 et 1 en vitesse de la fonction de distribution,

complété d'une équation sur le moment d'ordre 2 avec un terme source de relaxation. Précisément, rappelons le « changement de variables » qui correspond au calcul des-dits moments, d'ordre 0 pour ρ, 1 pour m et 2 pour z :

$$\left\{ \begin{array}{l} \rho \,=\, f_+ + 2\, f_0 + f_- \,, \\ m \,=\, \rho\, u \,=\, f_+ - f_- \,, \\ z = f_+ + f_- \,. \end{array} \right. \tag{4.2}$$

La formulation en moments de (4.1) s'écrit donc :

$$\left\{ \begin{array}{rcl} \partial_t \rho + \partial_x \left(\rho\, u\right) & = & 0 \,, \\ \\ \partial_t \left(\rho\, u\right) + \partial_x z & = & 0 \,, \\ \\ \partial_t z + \partial_x \left(\rho\, u\right) & = & -\frac{1}{\varepsilon}\, \rho\, \left(z - \frac{1}{2}\, \left(\rho + \rho\, u^2 \right) \right) \,. \end{array} \right. \tag{4.3}$$

Ainsi, nous disposons de deux formulations différentes du problème, mais qui ne sont pas vraiment « équivalentes » : nous sommes partis d'un modèle physique (4.1) qui représente l'évolution de fonctions distributions de particules, donc des fonctions *positives* ! Rien ne donne pour autant de conditions de signes si l'on fait le chemin inverse de (4.3) à (4.1). Une façon de voir cette non équivalence des formulations, ainsi que le fait que la « bonne » formulation pour ce qui nous intéresse est plutôt (4.1), est de regarder la condition de stabilité de Shizuta-Kawashima [29, 129, 179], évoquée précédemment et qui concerne les problèmes hyperboliques de relaxation [13].

Limite hydrodynamique, condition de stabilité. Considérons d'abord la forme (4.3) : c'est un système de lois de conservation avec un terme source de relaxation. Nous avons vu que la légitimité (stabilité) de la limite lorsque le paramètre de relaxation ε tend vers zéro est soumise à une condition de stabilité reliant les valeurs propres du système avec celles de sa limite de type « Euler » :

$$\left\{ \begin{array}{rcl} \partial_t \rho + \partial_x \left(\rho\, u\right) & = & 0 \,, \\ \\ \partial_t \left(\rho\, u\right) + \partial_x \left(\frac{1}{2}\, \left(\rho + \rho\, u^2 \right) \right) & = & 0 \,. \end{array} \right. \tag{4.4}$$

Cette condition peut s'observer, comme dans le modèle de Jin-Xin, en faisant formellement un développement de Chapman-Enskog de la solution de (4.3) autour de l'équilibre :

$$\left\{ z \,=\, \frac{1}{2}\, \left(\rho + \rho\, u^2 \right) \right\} \,.$$

[13] Notons néanmoins l'article de R. Natalini [153], dans lequel l'auteur se place du point de vue hyperbolique et propose une interprétation cinétique d'un système de relaxation quasi-linéaire pour approcher une loi de conservation multidimensionnelle non linéaire. Il montre alors la convergence vers la solution entropique du problème relaxé, sous la condition sous-caractéristique et avec une hypothèse de monotonie sur les Maxwelliennes. La démarche de ce travail est différente de la notre.

On peut aussi écrire (toujours formellement) l'entrelacement des valeurs propres de (4.4) entre celles de (4.3) (qui sont évidemment $+1$, 0 et -1). Il vient ainsi :

$$|u| < \sqrt{2}\,. \tag{4.5}$$

Or si l'on s'intéresse à la version « physique », c'est-à-dire en termes de distributions, alors la positivité requise des fonctions f_i implique en particulier :

$$|u| = \left| \frac{f_+ - f_-}{f_+ + 2f_0 + f_-} \right| \leq 1\,,$$

condition qui contient donc la condition (4.5) ! Nous ne nous intéresserons donc plus à ce critère de stabilité dans la suite, car nous nous concentrerons sur la vision cinétique.

Equilibres locaux. Avant de pouvoir passer à la limite hydrodynamique, lorsque le libre parcours moyen ($\sim \varepsilon$) tend vers 0, il nous faut caractériser les équilibre locaux de notre problème, à savoir les équivalents des Maxwelliennes locales de l'équation de Boltzmann, qui annulent l'opérateur de collision. Pour les calculer, on se donne les deux moments transportés au cours du temps, $\rho > 0$ et $\rho\, u \in \mathbb{R}$. Alors, on peut exprimer les Maxwelliennes locales de (4.1) associées à ces moments comme suit :

$$\left\{ \begin{array}{l} M_+ = \dfrac{\rho}{4}\,(1+u)^2\,, \\ \\ M_0 = \dfrac{\rho}{4}\,\left(1-u^2\right)\,, \\ \\ M_- = \dfrac{\rho}{4}\,(1-u)^2\,. \end{array} \right. \tag{4.6}$$

Entropie et convergence vers l'équilibre local. L'un des ingrédients essentiels des méthodes de dissipation d'entropie pour établir la limite hydrodynamique des équations cinétiques est le célèbre *Théorème H* de Boltzmann [32], qui affirme que la fonctionnelle H de Boltzmann, aussi appelée entropie, est décroissante au cours du temps [14]. Fort heureusement, cette fonctionnelle d'entropie possède son analogue dans les modèles cinétiques à vitesses discrètes comme le système de Broadwell :

$$H(f) := \sum_i f_i\, \ln(f_i)\,,$$

où dans notre cas $i \in \{+, 0, 0, -\}$. Le Théorème H s'écrit alors :

$$\frac{\mathrm{d}}{\mathrm{d}t} H(f(t,\cdot)) \leq 0\,.$$

[14] Cette entropie « mathématique » correspond en effet, au signe près, à l'entropie physique qui, elle, croît au cours du temps.

Remarque 12. Notons au passage que cette entropie (convexe !) est bien également une entropie « au sens hyperbolique » pour le système de relaxation (4.3). Elle est en outre bien *dissipative* au sens de l'article de G.Q. Chen, T.P. Liu, et C.D. Levermore [65] : elle confère en effet au terme source de relaxation le caractère dissipatif requis par la définition de [65] (grâce au Théorème H).

En fait, avec cette fonctionnelle d'entropie, tout reste à faire. Si l'on souhaite montrer la convergence vers l'équilibre hydrodynamique *via* une méthode d'entropie, le principe général est le suivant :

- La première idée est de mesurer la distance à l'équilibre, non pas en norme L^1 comme nous l'avons fait dans [94], mais plutôt « en entropie ». Il s'agit de montrer que *l'entropie relative*,

$$H\left(f|M\right) := H(f) - H(M) - H'(M)\,(f-M)$$

  tend vers 0 lorsque le paramètre ε tend vers 0 (ou lorsque t tend vers $+\infty$).

- Pour ce faire, un outil primordial est la fonctionnelle de dissipation d'entropie D définie par :

$$\frac{\mathrm{d}}{\mathrm{d}t} H(f(t,\cdot)) := -\frac{1}{\varepsilon}\, D(f(t,\cdot))\,.$$

  On souhaite alors montrer que cette fonctionnelle *contrôle*, en un sens, l'entropie relative. Il vient ensuite, en utilisant la définition de D que l'entropie relative vérifie une inégalité différentielle, ce qui permet d'établir sa convergence vers 0, avec parfois même un taux de convergence explicite.

- Enfin, on peut obtenir que la convergence en entropie implique la convergence en norme L^1, à l'aide d'inégalités de type Csisár-Kullback-Pinsker [74].

Ainsi, nous avons établi le programme qu'il nous faut suivre pour tenter de faire passer tous les arguments au niveau discret. Evoquons pour terminer un autre travail (parmi d'autres !) qui propose également d'appliquer ces techniques à un schéma exponentiel de Runge-Kutta pour des équations cinétiques, celui de G. Dimarco et L. Pareschi [81].

4.2 Autour des modèles de fluides géophysiques à surface libre

Dans le prolongement du travail de la partie I, je m'apprête à intégrer, à Toulouse sous la direction de J.-P. Vila, un groupe de recherche sur la modélisation et la prévision de circulations océaniques. Il s'agit essentiellement du modèle primitif, mais prenant en compte plus de paramètres physiques tels que la température et la salinité. L'objectif sera de travailler à la fois sur le modèle mathématique et sur le code opérationnel déjà existant.

En parallèle, je souhaiterais étudier les équations des rivières, ou problème des *roll waves*, plus particulièrement d'un point de vue numérique : serait-il possible d'appliquer à ce système un schéma AP du même type que celui de la partie II ? Il s'agit en effet du problème de Saint-Venant sur un plan incliné avec une loi de friction quadratique [121, 179].

Après un adimensionnement, le problème peut s'écrire sous la forme d'un système de deux lois de conservation avec un terme source de relaxation :

$$\left\{ \begin{array}{lcl} \partial_t h + \partial_x (h\,u) & = & 0\,, \\ \partial_t (h\,u) + \partial_x \left(h\,u^2 + g\dfrac{h^2}{2} \right) & = & \dfrac{1}{\varepsilon} \left(g\,h\,S - C_f\,u^2 \right)\,, \end{array} \right. \tag{4.7}$$

où S désigne la pente de la topographie, et C_f le coefficient de friction. Il apparaît que lorsque ce système viole la condition sous-caractéristique [121, 127], c'est-à-dire lorsque la pente est trop grande relativement à la friction (condition également reliée au nombre de Froude), on peut néanmoins voir s'installer un régime stable périodique : des roll waves. La question est donc de savoir si l'on peut adapter le schéma construit à la partie II, même si le critère de stabilité n'est pas vérifié, et si l'on peut observer numériquement ces ondes de surfaces, bien répandues dans la nature.

Première partie

Approximation des équations de Navier-Stokes incompressibles à surface libre : un modèle Saint-Venant multicouche dynamique

Chapitre 2

Derivation of a dynamic multilayer shallow water model of approximation of free surface Navier-Stokes equations ; existence of local in time strong solution

We propose a new simple approximation of the viscous primitive equations of the ocean including Coriolis force (2.1.1), by a multilayer shallow water type model. Using a finite volume type discretization in the vertical direction, we show that our system is a consistent approximation of (2.1.1) and we compare it briefly with other multilayer shallow water type existing models. Next, existence and uniqueness of local in time strong solution is proved for the new model.

2.1 Introduction and Main Result

The main goal of this work is to propose a simple and numerically efficient model of geophysical flows such as large-scale ocean circulations. Many of these flows are generally described by the incompressible Navier-Stokes equations with a free surface [137]. Due to the mathematical complexity of this system, different approximations are usually performed, which aim in particular at finding a compromise between physical consistency and reasonable computational cost. Going beyond the Boussinesq approximation [159] we start our study by considering an homogeneous fluid (water), with density equal to one. Moreover we use the so-called *hydrostatic approximation*, that is we assume the pressure is hydrostatic and is not an unknown of the problem. Precisely, the departure model consists in the primitive equations of the ocean, given in the conservative form below. We use bold characters to indicate vector valued functions or variables. Hence the 3D velocity of the fluid, for which we separate the horizontal component and the vertical one as

$\mathbf{U} = (\mathbf{u}, w)^T \in \mathbb{R}^3$, satisfies in a local frame $(\mathbf{x}, z)$ the set of equations:

$$\left\{ \begin{array}{lcl} \mathrm{div}_{\mathbf{x}}\, \mathbf{u} + \partial_z w & = & 0\,, \\ \partial_t \mathbf{u} + \mathrm{div}_{\mathbf{x}}(\mathbf{u} \otimes \mathbf{u}) + \partial_z(w\,\mathbf{u}) + \nabla_{\mathbf{x}} p & = & -f\,\mathbf{u}^{\perp} + \mu\,\Delta \mathbf{u}\,, \\ \partial_z p & = & -g\,, \end{array} \right. \tag{2.1.1}$$

considered for

$$t > 0\,,\ \ (\mathbf{x}, z) \in \Omega_t = \left\{ (\mathbf{x}, z) \in \mathbb{R} \times \mathbb{R}^+ \,|\, z_b(\mathbf{x}) \leqslant z \leqslant \eta(t, \mathbf{x}) \right\}\,,$$

where z_b is the topography (not depending on time) and η is the free surface. The fluid depth is given by

$$H(t, \mathbf{x}) = \eta(t, \mathbf{x}) - z_b(\mathbf{x})\,.$$

The constant $\mu > 0$ is the viscosity coefficient and $f > 0$ is the Coriolis parameter also chosen constant. Indeed, in this approximation we consider the latitude on the earth as a constant, and our local frame $(\mathbf{x}, z)$ can be seen as a fixed cartesian frame [159]. Hence the gravitational force is supported by the vertical direction, whose modulus is the gravity constant g. The hydrostatic pressure p is therefore given, for all t, $\mathbf{x}$, z by:

$$p\,(t, \mathbf{x}, z) \,=\, g\,\left(\eta\,(t, \mathbf{x}) - z\right)\,.$$

The system is completed with boundary conditions. We use the subscript s (*resp.* b) to indicate that the function is evaluated at the surface (*resp.* the bottom). On the one hand, it holds a kinematic equation and the continuity of stresses at the free surface:

$$\left\{ \begin{array}{lcl} \partial_t \eta + \mathbf{u}_s \cdot \nabla_{\mathbf{x}} \eta & = & w_s\,, \\ \partial_z \mathbf{u}_s & = & \nabla_{\mathbf{x}} \mathbf{u}_s \cdot \nabla_{\mathbf{x}} \eta\,, \end{array} \right. \tag{2.1.2}$$

when considering the atmospheric pressure equal to zero. At the bottom, we impose no penetration and a Navier type wall law [41], with a constant laminar friction coefficient κ, that is:

$$\left\{ \begin{array}{lcl} \mathbf{u}_b \cdot \nabla_{\mathbf{x}} z_b & = & w_b, \\ \kappa\,\mathbf{u}_b & = & \mu\,\partial_z \mathbf{u}_b. \end{array} \right. \tag{2.1.3}$$

This set of equations (or more complicated versions), though an approximation of Navier-Stokes, has been widely studied for decades. On the one hand, it is today used in many operational predictive models of ocean circulations. On the other hand, its mathematical justification, based on a scale analysis of the physical problem, has been done rigorously. See the pionner articles [139] for formal derivations and existence results for wind driven flows; [20] for rigorous justifications. Roughly speaking, this hydrostatic pressure approximation relies on an asymptotic expansion of the Navier-Stokes equations with respect to a small dimensionless parameter ε (aspect ratio), that is the *shallow water* assumption:

$$\varepsilon \,=\, \frac{H_0}{\lambda_0} \ll 1 \tag{2.1.4}$$

where H_0 and λ_0 are the characteristic depth and the typical horizontal wavelenght of the ocean.See [137, 159] for more details.

Although the primitive equations are simpler than the full Navier-Stokes system, they still contain two main difficulties: non linearity and time dependency of the spatial domain. Therefore many other model are built, either from the Navier-Stokes problem, or from the primitive equations. In particular, one classical way to dispense with the moving spatial domain is to perform an integration of the equations in the vertical direction. This leads to the classical shallow water (*Saint-Venant*) model (see for example the rigorous derivation with flat bottom [101], [75, 87, 143] for a small topography, [35, 37] for an arbitrary one). The main assets of such a model are the reduction of the spatial dimension of the problem and its mathematical properties, leading in particular to a very efficient numerical treatment. Indeed the hyperbolic formulation (away from vacuum) allows the use of robust finite volume schemes, even able to handle dambreak situations and wet/dry front for hydraulic or costal problems [98, 160].
But this system still has also some drawbacks. On the one hand, its good numerical behavior is not fully understood from a mathematical point of view. Indeed, although it is well justified when departing from the Euler equations, its viscous version requires additional assumptions to allow the closure of the system [101, 171, 179]. Moreover, it is not well posed neither in the vacuum nor for large variations of the free surface. On the other hand, considering the solution to this system, one can only reach the mean value of the horizontal velocity in the z direction. Therefore we loose information on the vertical profile of the velocity field.

In the present work, we stay in the context of deep water: we want to propose a consistant approximation of the primitive model (2.1.1)–(2.1.3), which reduces the mathematical complexity. Hence, in order to keep information on the vertical profile of the velocity field, while taking advantage of the numerical efficiency of the shallow water formulation, we will perform a vertical discretization of the fluid depth H, cut into N *thin* layers, and integrate the momentum equation on each layer. Let us emphasize here that the slicing is done in the most simple way, that is:

$$H(t,\mathbf{x}) = \sum_{i=1}^{N} h_i(t,\mathbf{x})\,,$$

where the intermediate layer heights h_i are all of constant size, say $\overline{h}$, except the lowest and the highest ones, which aims at catching somehow the boundary layers at the bottom and the top of the fluid. It is illustrated in Figure 2.1 for 4 layers. Hence we define the

Figure 2.1: *Vertical discretization.*

fluid velocity $\mathbf{u}_i$ in layer i by:

$$\mathbf{u}_i(t,\mathbf{x}) \,=\, \frac{1}{h_i}\int_{z_{i-1/2}}^{z_{i+1/2}} \mathbf{u}(t,\mathbf{x},z)\,\mathrm{d}z\,,\; 1\leqslant i\leqslant N\,. \tag{2.1.5}$$

Then, the N-layers model which will be investigated hereafter can be written in 2D as follows. For any $(t,\mathbf{x})$ in $\mathbb{R}^+\times\mathbb{R}^2$:

$$\left\{\begin{array}{ll}
\partial_t H + \mathrm{div}_{\mathbf{x}}\left(\displaystyle\sum_{i=1}^{N} h_i\,\mathbf{u}_i\right) & = 0\,,\\[2ex]
\partial_t\left(h_N\,\mathbf{u}_N\right) + \mathrm{div}_{\mathbf{x}}\left(h_n\,\mathbf{u}_N\otimes\mathbf{u}_N + g\,\dfrac{h_N^2}{2}\right) & = \mu\,\left(\mathrm{div}_{\mathbf{x}}\left(h_N\,\nabla_{\mathbf{x}}\mathbf{u}_N\right) + \mathbf{DU}^z_{N+1/2} - \mathbf{DU}^z_{N-1/2}\right)\\
& \quad -g\,h_N\,\nabla_{\mathbf{x}} z_b + w_{N-1/2}\,\mathbf{u}_{N-1/2} - w_{N+1/2}\,\mathbf{u}_{N+1/2}\\
& \quad -f\;\left(h_N\,\mathbf{u}_N\right)^{\perp}\,,\\[2ex]
\partial_t\left(h_i\,\mathbf{u}_i\right) + \mathrm{div}_{\mathbf{x}}\left(h_i\,\mathbf{u}_i\otimes\mathbf{u}_i\right) + g\,h_i\,\nabla_{\mathbf{x}} h_N & = \mu\,\left(h_i\,\Delta_{\mathbf{x}}\mathbf{u}_i + \mathbf{DU}^z_{i+1/2} - \mathbf{DU}^z_{i-1/2}\right)\\
& \quad -g\,h_i\,\nabla_{\mathbf{x}} z_b + w_{i-1/2}\,\mathbf{u}_{i-1/2} - w_{i+1/2}\,\mathbf{u}_{i+1/2}\\
& \quad -f\;\left(h_i\,\mathbf{u}_i\right)^{\perp}\,,\quad 1\,\leqslant\, i\,\leqslant\, N-1\,.
\end{array}\right. \tag{2.1.6}$$

The terms $w_{i+1/2}$, nothing but the values of the vertical velocity at the interfaces $z_{i+1/2}$, provide the mass exchange terms between layers i and $i+1$. They are computed thanks to the integration of the divergence free condition (see Section 2.2). Precisely, they are defined by:

$$\left\{\begin{array}{l}
w_{1/2} = \mathbf{u}_1\cdot\nabla_{\mathbf{x}} z_b,\\[2ex]
w_{i+1/2} - w_{i-1/2} = -h_i\,\nabla_{\mathbf{x}}\cdot\mathbf{u}_i,\quad 1\leqslant i\leqslant N-1.
\end{array}\right. \tag{2.1.7}$$

The terms $\mathbf{u}_{i+1/2}$ represent the approximate values of the horizontal velocity at the interfaces $z_{i+1/2}$, given by a centered reconstruction:

$$\mathbf{u}_{i+1/2} = \left\{\begin{array}{ll}
0 & \text{if}\quad i=0\,,\,N\,,\\[2ex]
\left(h_i\,\mathbf{u}_{i+1} + h_{i+1}\,\mathbf{u}_i\right)/\left(h_{i+1}+h_i\right) & \text{if}\quad 1\leqslant i\leqslant N-1.
\end{array}\right. \tag{2.1.8}$$

Finally, the terms $\mathbf{DU}^z_{i+1/2}$ are the z-derivatives of the horizontal velocity, evaluated at the interfaces $z_{i+1/2}$ and coming from the vertical viscosity. We choose:

$$\mathbf{DU}^z_{i+1/2} \,=\, \left\{\begin{array}{ll}
\kappa\,\mathbf{u}_1/\mu & \text{if}\quad i=0\,,\\[2ex]
2\,(\mathbf{u}_{i+1}-\mathbf{u}_i)/(h_i+h_{i+1}) & \text{if}\quad 1\leqslant i\leqslant N-1\,,\\[2ex]
0 & \text{if}\quad i=N\,.
\end{array}\right. \tag{2.1.9}$$

The formal derivation of this set of equations in 2D will be obtained in Section 2.2. Moreover, we will study in this paper the local in time existence of strong solution for the 1D version of the system, that is:

$$\begin{cases} \partial_t H + \partial_x \left(\sum_{i=1}^{N} h_i \, u_i \right) & = 0 \,, \\ \partial_t \left(h_N \, u_N \right) + \partial_x \left(h_n \, u_N^2 + g \, \dfrac{h_N^2}{2} \right) & = \mu \, \left(\partial_x \left(h_N \, \partial_x u_N \right) + DU^z_{N+1/2} - DU^z_{N-1/2} \right) \\ & \quad - g \, h_N \, \partial_x z_b + w_{N-1/2} \, u_{N-1/2} - w_{N+1/2} \, u_{N+1/2} \,, \\ \partial_t \left(h_i \, u_i \right) + \partial_x \left(h_i \, u_i^2 \right) + g \, h_i \, \partial_x h_N & = \mu \, \left(h_i \, \partial_{xx} u_i + DU^z_{i+1/2} - DU^z_{i-1/2} \right) - g \, h_i \, \partial_x z_b \\ & \quad + w_{i-1/2} \, u_{i-1/2} - w_{i+1/2} \, u_{i+1/2} \,, \quad 1 \leqslant i \leqslant N-1 \,. \end{cases} \tag{2.1.10}$$

where we drop the Coriolis terms which have no meaning in 1D. In order to state the result, we introduce the following notations.
For any function f, we note $\|f\|$ (resp. $\|f\|_k$) the L^2-norm (resp. H^k-norm) of f. If $\mathbf{f} = (f_1, \ldots, f_n)$ is multidimensional, we define its $\mathbf{H}^k$-norm by

$$\|\mathbf{f}\|_k := \sum_{i=1}^{n} \|f_i\|_k \,.$$

Let B be a Banach space, k a non-negative integer and T some positive constant. We denote by $L^k_\infty(0,T;B)$ the Banach space of functions f on $[0,T]$ which have their values in B and are k times differentiable with respect to t and all the derivatives are bounded in B. We can now state our main result.

Theorem 1.1. *Consider the system (2.1.10) where $w_{i+1/2}$, $u_{i+1/2}$ and $DU^z_{i+1/2}$ are defined by the 1D versions of (2.1.7), (2.1.8) and (2.1.9), with initial data*

$$(\mathbf{U}, h_N)(0,x) = (\mathbf{U}^0(x), h_N^0(x)) \, \in \, \mathbf{H}^2(\mathbb{R}) \,, \tag{2.1.11}$$

where $\mathbf{U} = (u_1 \, \ldots \, u_N)^T$ is the vector of velocities. Suppose

$$\inf_{x \in \mathbb{R}} h^0(x) \, \geqslant \, \eta_0 > 0 \,,$$

for some constant η_0, and note $E = 2 \, \|(\mathbf{U}^0, h_N^0)\|_2$. Assume the topography has the regularity $z_b \in \mathcal{C}^2(\mathbb{R})$. Then, there exists a positive constant T such that the Cauchy problem (2.1.10)-(2.1.11) has a unique strong solution $(\mathbf{U}, h_N)$ satisfying:

$$\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2\left(0,T;\mathbf{H}^3(\mathbb{R})\right) \,,$$

$$h_N \in \mathcal{C}(0,T;H^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^1(\mathbb{R})) \,.$$

Moreover, for any t in $[0,T]$,

$$\forall\, x \in \mathbb{R}\,,\, h_N(t,x) \geqslant \big(\inf_{x\in\mathbb{R}} h_N^0(x)\big)/2 > 0\,,$$

and the following energy estimates hold:

$$\|(\mathbf{U}, h_N)(t)\|_2 \leqslant E\,, \quad \left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\, d\tau\right)^{1/2} \leqslant E\,.$$

Theorem 2.3 is stated in 1D for sake of clarity in the computations but the proof, based on energy method of Matsumura and Nishida [146] can be adapted to the two dimensional problem [(1)], except that we have to choose initial data in $\mathbf{H}^3(\mathbb{R})$, and the solution $(\mathbf{u}_1, \ldots, \mathbf{u}_N, h_N)$ get the regularity:

$$\mathbf{u}_i \in \mathcal{C}(0,T;\mathbf{H}^3(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^1(\mathbb{R})) \cap L^2\left(0,T;\mathbf{H}^4(\mathbb{R})\right)\,,\, \forall\, i \in \{1,\ldots,N\}\,,$$

$$h_N \in \mathcal{C}(0,T;H^3(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^2(\mathbb{R}))\,.$$

Remark 1.2. This existence result is in good agreement with the ones already existing for classical shallow water systems, in particular the condition on the initial water height, bounded by below by a positive constant. Let us mention, without being exhaustive, for example the works [51, 131, 172, 173, 178], treating different kinds of solutions to the Cauchy problem.

Remark 1.3. Of course the existence is only local in time since the model (2.1.10) blows up when h_N reaches zero at one point. Therefore it gives a criterion to make the model dynamic by removing and adding layers. On the one hand if at a time t_1 the highest height h_N becomes too small (say under some non negative threshold), then one removes one layer at the top, and starts again with the model with $N-1$ layers. On the other hand, one can add a layer to the model when the height of the highest layer is large enough. We will see this dynamic behavior in some preliminary numerical simulations of Section 3.2.

The rest of the chapter is organized as follows. In Section 2.2 we first derive rigorously the multilayer system (2.1.6) from the three dimensional free surface primitive model (2.1.1). Then we briefly compare our model to some existing multilayer models [12, 15] and point out that we do not aim at modelling the same kind of geophysical problems. In Section 2.3 we prove Theorem 2.3, with the energy method of Matsumura and Nishida [146]. Finally, in Section 3.1, we design a simple numerical scheme in order to validate our model and perform some numerical experiments. We compare the multilayer model with classical shallow water models and the primitive system, and illustrate its dynamic behavior.

[1] The Coriolis term does not add any difficulty since it is a zeroth order term

2.2 Derivation of the model and comparison with other multilayer models

2.2.1 Derivation

As it was said in the introduction, we derive our multilayer model from the 3D viscous primitive system with friction and Coriolis terms (2.1.1)–(2.1.3) introduced in Section 1. We start by performing the vertical discretization of the water height illustrated in Figure 2.1:

$$\eta - z_b = H := \sum_{i=1}^{N} h_i, \quad \text{with } h_i = z_{i+1/2} - z_{i-1/2} = O(\overline{h}), \quad 1 \leqslant i \leqslant N, \tag{2.2.1}$$

where the small constant $\overline{h}$ is fixed and the nodes of discretization are chosen as:

$$\begin{cases} z_{1/2} = z_b(\mathbf{x}), \\ \\ z_{i+1/2} = i\,\overline{h}, \quad 1 \leqslant i \leqslant N-1, \\ \\ z_{N+1/2} = \eta(t,\mathbf{x}). \end{cases} \tag{2.2.2}$$

Using this vertical discretization and the definition of the velocities (2.1.5), we claim

Proposition 2.1. *Assume the variations of the bathymetry are controled as:*

$$\nabla_{\mathbf{x}} z_b \,=\, O\left(\overline{h}\right). \tag{2.2.3}$$

Then the multilayer formulation (2.1.6), where h_i, $\mathbf{u}_{i+1/2}$, $w_{i+1/2}$, are given by (2.2.1), (2.1.8) and (2.1.7), is a formal asymptotic approximation in $O\left(\overline{h}^2\right)$ of the primitive equations (2.1.1)-(2.1.2)-(2.1.3).

Proof. On the one hand, the integration through each layer $1 \leqslant i \leqslant N$ of the momentum equation gives:

$$\partial_t(h_i\,\mathbf{u}_i) - \Big[\partial_t z\,\mathbf{u}\Big]_{z_{i-1/2}}^{z_{i+1/2}} + \nabla_{\mathbf{x}}\cdot(h_i\,\mathbf{u}_i\otimes\mathbf{u}_i) - \Big[\left(\nabla_{\mathbf{x}} z\cdot\mathbf{u}\right)\,\mathbf{u}\Big]_{z_{i-1/2}}^{z_{i+1/2}} + \Big[w\,\mathbf{u}\Big]_{z_{i-1/2}}^{z_{i+1/2}} + g\,h_i\,\nabla_{\mathbf{x}}\eta$$

$$= -f\,\left(h_i\,\mathbf{u}_i\right)^{\perp} + \mu\left\{\Big[\partial_z\mathbf{u}\Big]_{z_{i-1/2}}^{z_{i+1/2}} + \nabla_{\mathbf{x}}\cdot\left(\int_{z_{i-1/2}}^{z_{i+1/2}} \nabla_{\mathbf{x}}\mathbf{u}\,\mathrm{d}z\right) - \Big[\nabla_{\mathbf{x}}\mathbf{u}\cdot\nabla_{\mathbf{x}} z\Big]_{z_{i-1/2}}^{z_{i+1/2}}\right\}. \tag{2.2.4}$$

Let us notice here that most of the terms between square-brackets will cancel since the inside layer sizes are constant in time and space. On the other hand, by integrating the divergence free condition, we get:

$$w(z_{i+1/2}) - w(z_{i-1/2}) = -\int_{z_{i-1/2}}^{z_{i+1/2}} \nabla_{\mathbf{x}}\cdot\mathbf{u}\,\mathrm{d}z, \quad 1 \leqslant i \leqslant N.$$

It is therefore sufficient to apply Taylor expansions in the vertical direction. Namely, assuming the velocities are smooth enough, we have the following approximations: for all $1 \leqslant i \leqslant N-1$, for all $z \in [z_{i-1/2}, z_{i+1/2}]$:

$$\left\{\begin{array}{lcl} \mathbf{u}(z) & = & \mathbf{u}_i + O\left(\overline{h}\right) , \\ \mathbf{u}(z_{i+1/2}) & = & \mathbf{u}_{i+1/2} + O\left(\overline{h}^2\right) , \\ \partial_z \mathbf{u}(z_{i+1/2}) & = & 2\,\dfrac{\mathbf{u}_{i+1} - \mathbf{u}_i}{h_i + h_{i+1}} + O\left(\overline{h}^2\right) , \\ \displaystyle\int_{z_{i-1/2}}^{z_{i+1/2}} \mathbf{u} \otimes \mathbf{u}\,\mathrm{d}z & = & h_i\,\mathbf{u}_i \otimes \mathbf{u}_i + O\left(\overline{h}^2\right) , \\ \displaystyle\int_{z_{i-1/2}}^{z_{i+1/2}} \nabla_{\mathbf{x}} \mathbf{u}\,\mathrm{d}z & = & h_i\,\nabla_{\mathbf{x}} \mathbf{u}_i + O\left(\overline{h}^2\right) . \end{array}\right. \tag{2.2.5}$$

Next, by the use of the boundary conditions at the bottom (2.1.3) and the order of magnitude of the variations of the bathymetry (2.2.3), the definition of the approximate reconstructions of the vertical velocity at the interfaces between layers (2.1.7) yields, for all $0 \leq i \leq N-1$:

$$w_{i+1/2} \,=\, w(z_{i+1/2}) + O\left(\overline{h}^2\right).$$

Let us look at the viscous terms with the previous approximations:

$$\nabla_{\mathbf{x}} \cdot \left(\int_{z_{i-1/2}}^{z_{i+1/2}} \nabla_{\mathbf{x}} \mathbf{u}\,\mathrm{d}z \right) = \left\{\begin{array}{ll} h_1\,\Delta_{\mathbf{x}} \mathbf{u}_1 - \nabla_{\mathbf{x}} \mathbf{u}_b \cdot \nabla_{\mathbf{x}} z_b + O(\overline{h}^2) & \text{if } i = 1\,, \\ h_i\,\Delta_{\mathbf{x}} \mathbf{u}_i + O(\overline{h}^2) & \text{if } i = 2, \ldots, N-1\,, \\ \nabla_{\mathbf{x}} \cdot (h_N\,\nabla_{\mathbf{x}} \mathbf{u}_N) + O(\overline{h}^2) & \text{if } i = N\,. \end{array}\right.$$

Hence, in the equation (2.2.4) for the lowest layer, the non zero boundary terms of the viscous part cancel each other and we get, using the boundary conditions at the bottom:

$$\begin{array}{rcl} \partial_t\,(h_1\,\mathbf{u}_1) + \nabla_{\mathbf{x}} \cdot (h_1\,\mathbf{u}_1 \otimes \mathbf{u}_1) + g\,h_1\,\nabla_{\mathbf{x}} h_N & = & -g\,h_1\,\nabla_{\mathbf{x}} z_b - w_{3/2}\,\mathbf{u}_{3/2} - \kappa\,\mathbf{u}_1 - f\,(h_1\,\mathbf{u}_1)^{\perp} \\ & & +\mu\,h_1\,\Delta_{\mathbf{x}} \mathbf{u}_1 + 2\,\mu\,\dfrac{\mathbf{u}_2 - \mathbf{u}_1}{h_1 + h_2} + O\left(\overline{h}^2\right) . \end{array}$$

For the inside layers, we use again the approximations (2.2.5): the result is obtained easily because the intermediate layer heights are *constant* in time and space. Finally, there is another term in equation (2.2.4) for the highest layer, coming from the time dependency of the highest layer. It is simplified thanks to the boundary conditions at the free surface

(2.1.2). It leads to:

$$\begin{aligned}\partial_t (h_N \mathbf{u}_N) + \nabla_{\mathbf{x}} \cdot (h_N \mathbf{u}_N \otimes \mathbf{u}_N) + g h_N \nabla_{\mathbf{x}} h_N &= -g h_N \nabla_{\mathbf{x}} z_b - w_{N+1/2} \mathbf{u}_{N+1/2} \\ &\quad + \mu \nabla_{\mathbf{x}} \cdot (h_N \nabla_{\mathbf{x}} \mathbf{u}_N) - 2\mu \frac{\mathbf{u}_N - \mathbf{u}_{N-1}}{h_N + h_{N-1}} \\ &\quad - f\ (h_N \mathbf{u}_N)^{\perp} + O\left(\overline{h}^2\right).\end{aligned}$$

To conclude, we drop the $O\left(\overline{h}^2\right)$ and obtain the system (2.1.6)-(2.1.7) as a formal approximation of system (2.1.1)–(2.1.3) in $O\left(\overline{h}^2\right)$. This ends the proof. □

2.2.2 Comparison with other multilayer models

Let us now briefly compare our model to other multilayer shallow water models, that is the ones introduced by E. Audusse and coauthors [12, 15]. First, we want to point out that if the general framework is somehow similar, the models do not aim at modelling the same phenomena. We focus here on deep water, while the models of [12, 15] mainly treat costal area [14, 13, 15, 59] [(2)]. Actually, we can see our model as an intermediate step for discretization of the primitive model, with an adaptative mesh in the vertical direction. Indeed, we will see in the preliminary numerical results that wecan change the number of layers as time goes. Moreover, when we approach a costal zone, we can imagine a coupling between our multilayer model in the deep area with a classical shallow water model in the shallow one.
Second, the way of cutting the water height H is different. In [12, 15], the authors "follow" the free surface inside the fluid, as illustrated in Figure 2.2 for 4 layers. Therefore, this vertical discretization allows to keep all the good properties of the classical shallow water system: the positivity of the total height immediately gives positivity for all the inside layers and the numerical treatment of the vacuum is also done as for the one layer case. Unfortunately, it keeps also the same mathematical weakness of the classical shallow water model, that is the closure of the system for the viscous terms [(3)].

2.3 Well-posedness of the multilayer model

In this section, we study the well-posedness of the 1D multilayer model (2.1.10) and prove Theorem 2.3. To do so, we rewrite the system under the form of a coupled parabolic-hyperbolic system with source terms. Since the only unknown layer height with our framework is the highest one h_N, we will denote it h for sake of clarity. Then, by dividing the

[2] Indeed these models are rather derived from the dimensionless Navier-Stokes equations and give a formal approximation in $O(\varepsilon^2)$, where ε is defined in (2.1.4).

[3] The viscous terms are chosen as the one in the classical shallow water system, but the derivation is justified in the zero viscosity case.

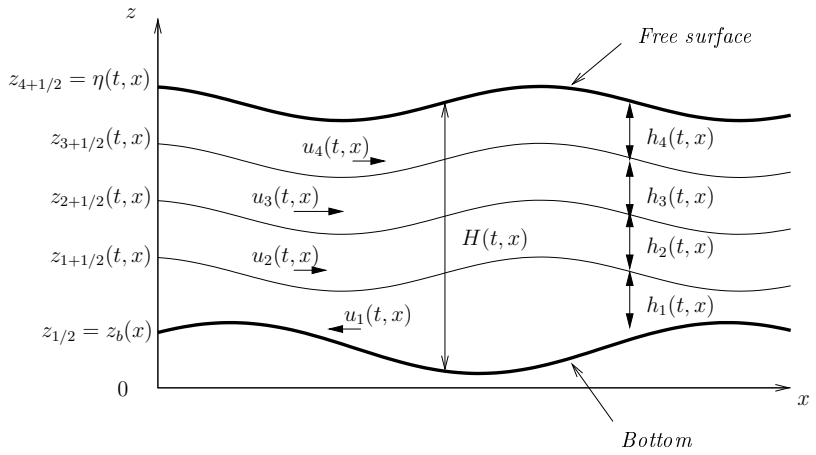


Figure 2.2: *Classical multilayer approach.*

equations by the heights, we rewrite the system on the unknown $(\mathbf{U}, h) = (u_1 \ \dots \ u_N, h)^T$ as follows.

$$\begin{cases} \partial_t \mathbf{U} - \mu \, \partial_{xx} \mathbf{U} = \mathbf{S} \, , \\ \\ \partial_t h + \partial_x (h \, u_N) = F \, , \end{cases} \tag{2.3.1}$$

where the source terms are described below.

$$\begin{cases} \mathbf{S} \ = \ \mathbf{S}_b + \mathbf{S}_l + \mathbf{S}_{nl} \, , \\ \\ F \ = \ w_{N-1/2} \ = \ - \displaystyle\sum_{i=1}^{N-1} \partial_x \left(h_i \, u_i \right) \, , \end{cases}$$

where $\mathbf{S}_b$ refers to the bottom source term

$$\mathbf{S}_b = -g \, \partial_x z_b \, (1, \dots, 1)^T \, ,$$

while $\mathbf{S}_l = \left(S_l^1, \dots, S_l^N\right)^T$ and $\mathbf{S}_{nl} = \left(S_{nl}^1, \dots, S_{nl}^N\right)^T$ are respectively the linear and the non linear sources, that is:

$$\left\{\begin{array}{rcl} S_l^i & = & -g\,\partial_x h\,, \quad 1 \leqslant i \leqslant N \\ \\ S_{nl}^1 & = & 2\,\mu\,\dfrac{u_2-u_1}{h_1\,(h_1+h_2)} - \dfrac{\kappa}{h_1}\,u_1 \\ \\ & & -\partial_x\left(u_1^2\right) - \dfrac{1}{h_1}\,\left(u_{3/2}\,w_{3/2} - u_1\,w_{1/2}\right)\,, \\ \\ S_{nl}^i & = & 2\,\mu\,\dfrac{u_{i+1}-u_i}{h_i\,(h_i+h_{i+1})} - 2\,\mu\,\dfrac{u_i-u_{i-1}}{h_i\,(h_i+h_{i-1})} \\ \\ & & -\partial_x\left(u_i^2\right) - \dfrac{1}{h_i}\,\left(u_{i+1/2}\,w_{i+1/2} - u_{i-1/2}\,w_{i-1/2}\right)\,, \quad 2 \leqslant i \leqslant N-1\,, \\ \\ S_{nl}^N & = & -2\,\mu\,\dfrac{u_N-u_{N-1}}{h\,(h+h_{N-1})} + \mu\,\dfrac{\partial_x h\,\partial_x u_N}{h} \\ \\ & & -\dfrac{1}{2}\,\partial_x\left(u_N^2\right) + \dfrac{1}{h}\,\left(u_{N-1/2}-u_N\right)\,w_{N-1/2}\,. \end{array}\right.$$

Hence, we can sum up by considering that the source term $\mathbf{S}_{nl}$ is roughly composed of three kinds of non linearities, that is

$$u/h\,, \quad u\,\partial_x u/h\,, \quad \partial_x h\,\partial_x u/h\,.$$

Consequently, in order to simplify the next calculations, we will only consider the simpler hyperbolic-parabolic problem

$$\left\{\begin{array}{l} \partial_t \mathbf{U} - \mu\,\partial_{xx}\mathbf{U} = \mathbf{S}_b + \mathbf{S}_l + \mathbf{S}_{nl}\,, \\ \\ \partial_t h + \partial_x(h\,u_N) = F\,, \end{array}\right. \tag{2.3.2}$$

where F, $\mathbf{S}_b$, $\mathbf{S}_l$ are not changed, while the nonlinear source is simplified as

$$\mathbf{S}_{nl} = \sum_{k=1}^{3} \mathbf{S}_k\,,$$

where

$$\mathbf{S}_1 = \frac{\mathbf{U}}{h}\,, \;\; \mathbf{S}_2 = \frac{u_N}{h}\,\partial_x \mathbf{U}\,, \;\; \mathbf{S}_3 = \frac{1}{h}\,\partial_x h\,\partial_x \mathbf{U}\,.$$

The proof of Theorem 2.3 is divided into three parts. In the first subsection we perform some estimates on the source terms for the simpler problem (2.3.2). Next we solve a linearized version of system (2.3.2) and derive energy estimates. Finally, we build a recursive sequence of solutions of linear systems, and show a convergence to the strong solution we are looking for.

2.3.1 Estimates on the source terms

We first recall a classical lemma of analysis which will be useful to estimate the source terms [174].

Lemma 3.1 (Moser estimate)**.** *Let $k \in \mathbb{N}$ and $n \in \mathbb{N}^*$. Suppose $f,\, g \in H^k(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. Then $f,\, g \in H^k(\mathbb{R}^n)$ and there exists a positive constant C such that*

$$\|fg\|_k \leqslant C\left(\|f\|_k \|g\|_\infty + \|g\|_k \|f\|_\infty\right).$$

We are now ready to state the estimations of the source terms.

Lemma 3.2. *Let $\mathbf{U}(t,\cdot),\, h(t,\cdot) \in \mathbf{H}^2(\mathbb{R})$ such as, $h(t,x) \geq \eta_0 > 0$, for some constant η_0. Then it holds:*

- *$(\mathbf{S}, F) \in H^1(\mathbb{R})$ and we have the following estimates.*

$$\|\mathbf{S}\|_1 \leqslant C(\eta_0)\,\|(\mathbf{U},h)\|_2 \left(1 + \|(\mathbf{U},h)\|_2\right), \tag{2.3.3}$$

$$\|F\|_1 \leqslant C_b\,\|\mathbf{U}\|_2, \tag{2.3.4}$$

  where $C(\eta_0)$, C_b are positive constants depending respectively, only on η_0 and the topograhy z_b.

- *If moreover $\mathbf{U} \in \mathbf{H}^3(\mathbb{R})$, then $F \in H^2(\mathbb{R})$ and there exists some constant C_b such that:*

$$\|F\|_2 \leqslant C_b\,\|\mathbf{U}\|_3.$$

- *Let $(\mathbf{U},h), (\mathbf{U}',h') \in \mathbf{H}^2(\mathbb{R})$ such that, $\forall\,(t,x) \in \mathbb{R}^+ \times \mathbb{R}$*

$$\|(\mathbf{U},h)\|_2,\, \|(\mathbf{U}',h')\|_2 \leqslant E, \quad h,\, h' \geq \eta_0 > 0, \tag{2.3.5}$$

  for some constants E and η_0. Then it holds

$$\|\mathbf{S}(\mathbf{U},h) - \mathbf{S}(\mathbf{U}',h')\|_1 \leqslant C(\eta_0)\left(1 + E + E^2\right)\|(\mathbf{U}-\mathbf{U}', h-h')\|_2, \tag{2.3.6}$$

$$\|F(\mathbf{U}) - F(\mathbf{U}')\|_1 \leqslant C_b\,\|\mathbf{U}-\mathbf{U}'\|_2, \tag{2.3.7}$$

  where $C(\eta_0)$, C_b are positive constants independent of E.

Proof. The estimate on F is directly obtained from the definition

$$F = \partial_x z_b\, u_1 - \sum_{i=1}^{N-1} h_i\, \partial_x u_i.$$

We have, for $k = 1$ or 2:

$$\|F\|_k \leqslant C_b\,\|\mathbf{U}\|_{k+1}.$$

Next the linear sources are estimated as

$$\begin{cases} \|\mathbf{S}_b\|_1 \leqslant C\,\|z_b\|_2\,, \\ \\ \|\mathbf{S}_l\|_1 \leqslant C\,\|h\|_2\,. \end{cases}$$

Then we estimate the non linear source part $\mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3$. It follows from Lemma 3.1 and the classical Sobolev embedding

$$H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})\,,$$

that we have the estimates:

$$\|\mathbf{S}_1\|_1 \leqslant \frac{C}{\eta_0}\,\|\mathbf{U}\|_1\,,\ \ \|\mathbf{S}_2\|_1 \leqslant \frac{C}{\eta_0}\,\|\mathbf{U}\|_2^2\,,\ \ \|\mathbf{S}_3\|_1 \leqslant \frac{C}{\eta_0}\,\|h\|_2\,\|\mathbf{U}\|_2\,.$$

It gives immediately (2.3.3). Now we note $\mathbf{S} = \mathbf{S}(\mathbf{U}, h)$ and $\mathbf{S}' = \mathbf{S}(\mathbf{U}', h')$. We compute the difference

$$\begin{cases} \mathbf{S}_l^i - \mathbf{S}_l'^i &=& -g\,\partial_x(h - h')\,\forall\, i\,, \\ \\ \mathbf{S}_1 - \mathbf{S}_1' &=& \dfrac{1}{h}\,\big(\mathbf{U} - \mathbf{U}'\big) + \dfrac{h' - h}{h\,h'}\,\mathbf{U}'\,, \\ \\ \mathbf{S}_2 - \mathbf{S}_2' &=& \dfrac{u_N}{h}\,\partial_x(\mathbf{U} - \mathbf{U}') + \dfrac{u_N - u_N'}{h}\,\partial_x\mathbf{U}' + \dfrac{u_N'}{h\,h'}\,(h' - h)\,\partial_x\mathbf{U}'\,, \\ \\ \mathbf{S}_3 - \mathbf{S}_3' &=& \dfrac{\partial_x h}{h}\,\partial_x(\mathbf{U} - \mathbf{U}') + \dfrac{u_N - u_N'}{h}\,\partial_x\mathbf{U}' + \dfrac{u_N'}{h\,h'}\,(h' - h)\partial_x\mathbf{U}'\,. \end{cases}$$

Then, applying Lemma 3.1 and using (2.3.5), we obtain

$$\begin{cases} \|\mathbf{S}_l - \mathbf{S}_l'\|_1 &\leqslant& C\,\|h - h'\|_2\,, \\ \\ \|\mathbf{S}_1 - \mathbf{S}_1'\|_1 &\leqslant& \dfrac{C}{\eta_0}\,\|\mathbf{U} - \mathbf{U}'\|_1 + \dfrac{C}{\eta_0^2}\,E\,\|h' - h\|_1\,, \\ \\ \|\mathbf{S}_2 - \mathbf{S}_2'\|_1 &\leqslant& \dfrac{C}{\eta_0}\,E\,\|\mathbf{U} - \mathbf{U}'\|_2 + \dfrac{C}{\eta_0^2}\,E^2\,\|h - h'\|_1\,, \\ \\ \|\mathbf{S}_3 - \mathbf{S}_3'\|_1 &\leqslant& \dfrac{C}{\eta_0}\,E\,\|\mathbf{U} - \mathbf{U}'\|_2 + \dfrac{C}{\eta_0}\,E\,\|h - h'\|_2 + \dfrac{C}{\eta_0^2}\,E^2\,\|h' - h\|_1\,. \end{cases}$$

Adding these inequalities, we get (2.3.6). Finally, the inequality (2.3.7) is straight forward since F is linear with respect to $\partial_x\mathbf{U}$. □

Next we give estimate of the commutator of the transport operator $\partial_t + u_N\,\partial_x$ and the second order space differential operator ∂_{xx}.

Lemma 3.3. *We assume $u_N \in H^2(\mathbb{R})$ with*

$$\|u_N\|_2 \leqslant E$$

for some positive constant E, and define the differential operator

$$\mathcal{L}_{u_N} := \partial_t + u_N\,\partial_x\,.$$

Then, for any $h \in L^0_\infty(0,T;H^2(\mathbb{R}))$, we have:

$$\Big\|\partial_{xx}\Big(\mathcal{L}_{u_N}\,(h)\Big) \,-\, \mathcal{L}_{u_N}(\partial_{xx}h)\Big\| \leqslant C\,E\,\|h\|_2\,,$$

Proof. We only compute:

$$\partial_{xx}\Big(\mathcal{L}_{u_N}\,(h)\Big) \,-\, \mathcal{L}_{u_N}\,(\partial_{xx}h) \,=\, 2\,\partial_x u_N\,\partial_{xx}h + \partial_{xx}u_N\,\partial_x h\,.$$

Hence, using Lemma 3.1 yields:

$$\Big\|\partial_{xx}\Big(\mathcal{L}_{u_N}\,(h)\Big) \,-\, \mathcal{L}_{u_N}\,(\partial_{xx}h)\Big\| \leqslant 2\,E\,\|\partial_{xx}h\| + E\,\|\partial_x h\|_1\,.$$

□

In the next subsection, we obtain energy estimates and study a linearized version of the multilayer system.

2.3.2 Study of the linearized problem

Let us introduce a linearized version of system (2.3.2):

$$\begin{cases} \partial_t \mathbf{U} - \mu\,\partial_{xx}\mathbf{U} \,=\, \mathbf{S}(\tilde{\mathbf{U}},\tilde{h},\partial_x\tilde{\mathbf{U}},\partial_x\tilde{h}) := \tilde{\mathbf{S}}, \\ \\ \mathcal{L}_{\tilde{u}_N}\,(h) \,=\, F - \tilde{h}\,\partial_x u_N := f. \end{cases} \tag{2.3.8}$$

In order to study the well-posedness of this coupled linear parabolic-hyperbolic problem, we first solve the parabolic system, and next the transport equation on h by considering the right hand side

$$f = -\tilde{h}\,\partial_x u_N - \sum_{i=1}^{N-1} \partial_x\,(h_i\,u_i)$$

as a known function. Thus, we will first study separately the following Cauchy problems, one parabolic system

$$\begin{cases} \left(\partial_t - \mu\,\partial_{xx}\right)\big(\mathbf{U}\big) \,=\, \tilde{\mathbf{S}}\,, \\ \\ \mathbf{U}(0,x) = \mathbf{U}^0 \,\in\, \mathbf{H}^2(\mathbb{R})\,, \end{cases} \tag{A}$$

and one hyperbolic scalar equation:

$$\begin{cases} \mathcal{L}_{\tilde{u}_N}\,(h) \,=\, f\,, \\ \\ h(0,x) = h^0 \,\in\, H^2(\mathbb{R})\,. \end{cases} \tag{B}$$

Proposition 3.4. *Let $\tilde{\mathbf{S}} \in \mathcal{C}(0,T;H^1(\mathbb{R}))$ for some $T>0$. Then the initial value problem (A) has a unique strong solution $\mathbf{U}$ which satisfies:*

$$\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2(0,T;\mathbf{H}^3(\mathbb{R}))\,.$$

Moreover, there exist two positive constants C_1 and C_2, depending only on the viscosity, such that for any t in $[0,T]$:

$$\|\mathbf{U}(t)\|_2^2 + C_1 \int_0^t \|\mathbf{U}(\tau)\|_3^2 \, d\tau \leqslant e^t \left(\|\mathbf{U}(0)\|_2^2 + C_2 \int_0^t \|\tilde{\mathbf{S}}(\tau)\|_1^2 \, d\tau \right). \tag{2.3.9}$$

Proof. First, the energy inequality is obtained in a classical way. Multiplying the system by $\mathbf{U}$ and integrating in space, one gets, for any $\alpha>0$:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{U}\|^2 + \mu\,\|\partial_x\mathbf{U}\|^2 \quad \leqslant \quad \frac{\alpha}{2}\|\mathbf{U}\|^2 + \frac{1}{2\,\alpha}\|\tilde{\mathbf{S}}\|^2\,.$$

Next, we differentiate with respect to x, multiply by $\partial_x\mathbf{U}$ and integrate in space, it gives, for any $\alpha>0$:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_x\mathbf{U}\|^2 + \mu\,\|\partial_{xx}\mathbf{U}\|^2 \quad \leqslant \quad \frac{\alpha}{2}\|\partial_x\mathbf{U}\|^2 + \frac{1}{2\,\alpha}\|\partial_x\tilde{\mathbf{S}}\|^2\,.$$

Finally, we compute the second order space derivative, multiply by $\partial_{xx}\mathbf{U}$ and integrate in space. Here, since we have too many derivatives on the source term $\tilde{\mathbf{S}}$, we integrate by parts the right hand side as follows: for any $\alpha>0$,

$$\left|\int_{\mathbb{R}} \partial_{xx}\tilde{\mathbf{S}}\,\partial_{xx}\mathbf{U}\,\mathrm{d}x\right| = \left|\int_{\mathbb{R}} \partial_x\tilde{\mathbf{S}}\,\partial_{xxx}\mathbf{U}\,\mathrm{d}x\right| \leqslant \frac{\alpha}{2}\|\partial_{xxx}\mathbf{U}\|^2 + \frac{1}{2\,\alpha}\|\partial_x\tilde{\mathbf{S}}\|^2\,.$$

Now we choose α such that $C_1 := 2\,\mu - \alpha > 0$, we add the previous inequalities and get:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{U}\|_2^2 + C_1\,\|\mathbf{U}\|_3^2 \quad \leqslant \quad \alpha\,\|\mathbf{U}\|_2^2 + \frac{1}{\alpha}\|\tilde{\mathbf{S}}\|_1^2\,.$$

We end the proof by applying the Gronwall Lemma.
This *a priori* estimate gives uniqueness of the solution. Concerning the proof of existence of solution, we introduce K^t the Green kernel of the operator $\partial_t - \mu\,\partial_{xx}$. Then, Duhamel's formula gives a solution $\mathbf{U} = (u_1,\ldots,u_N)^T$ of problem (A) defined by:

$$\forall\,(t,x) \in [0,T[\times\mathbb{R}\,,\; u_i(t,x) = K^t * u_i^0 + \int_0^t K^{t-s} * \tilde{\mathbf{S}}^i(s)\,\mathrm{d}s\,,\; i=1\,,\ldots,N\,.$$

We deduce immediately the smoothness of $\mathbf{U}$: it lies in $\mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R}))$. To get more regularity in time, we observe that:

$$\partial_t\mathbf{U} = \mu\,\partial_{xx}\mathbf{U} + \tilde{\mathbf{S}} \in \mathcal{C}(0,T;\mathbf{L}^2(\mathbb{R}))\,.$$

Hence $\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R}))$. □

We can now solve the Cauchy problem (B), considering the right hand side f as a known function.

Proposition 3.5. *Let $\tilde{u}_N \in \mathcal{C}(0,T;H^2(\mathbb{R}))$ and $f \in \mathcal{C}(0,T;H^k(\mathbb{R}))$ for $k = 1$ or 2, and $T > 0$. Denote*

$$E := \sup_{0 \leqslant t \leqslant T} \{\|\tilde{u}_N(t)\|_2\} .$$

Then the initial value problem (B) has a unique strong solution h which satisfies:

$$h \in \mathcal{C}(0,T;H^k(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^{k-1}(\mathbb{R})) .$$

Moreover, there exists a positive constant C_3, depending only on the dimension of the space, such that, for all t in $[0,T]$:

$$\|h(t)\|_k \quad \leqslant \quad e^{C_3\,E\,t}\left(\|h(0)\|_k + \int_0^t e^{-C_3\,E\,\tau}\,\|f(\tau)\|_k\,d\tau\right), \quad k = 1 \text{ or } 2 . \tag{2.3.10}$$

Proof. As in Proposition 3.4, we first obtain the energy estimate. Multiplying the equation by h and integrating in space, we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|h\|^2 \;=\; -\int_{\mathbb{R}} \tilde{u}_N\,\partial_x\left(\frac{h^2}{2}\right)\mathrm{d}x + \int_{\mathbb{R}} f\,h\,\mathrm{d}x .$$

We apply an integration by parts on the first term of the right hand side, and estimate the second term with Hölder inequality. It yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|h\|^2 \;\leqslant\; \frac{1}{2}\,E\,\|h\|^2 + \|f\|\,\|h\| .$$

Remove the square on the L^2-norms and get, for some constant $C > 0$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|h\| \;\leqslant\; C\,E\,\|h\| + \|f\| . \tag{2.3.11}$$

Next, we differentiate the equation and multiply by $\partial_x h$. We note that

$$\int_{\mathbb{R}} \partial_x\big(\tilde{u}_N\,\partial_x h\big)\,\partial_x h\,\mathrm{d}x = \frac{1}{2}\int_{\mathbb{R}} \partial_x\tilde{u}_N\,\big(\partial_x h\big)^2\,\mathrm{d}x ,$$

and get, as previously, the estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_x h\| \;\leqslant\; C\,E\,\|\partial_x h\| + \|\partial_x f\| . \tag{2.3.12}$$

Adding (2.3.11) and (2.3.12), it gives (2.3.10) for $k = 1$ thanks to the Gronwall Lemma. For $k = 2$, we need the estimate of commutator between the transport operator and ∂_{xx}, already proved in Lemma 3.3:

$$\Big\|\partial_{xx}\Big(\mathcal{L}_{\tilde{u}_N}(h)\Big) \;-\; \mathcal{L}_{\tilde{u}_N}\big(\partial_{xx}h\big)\Big\| \;\leqslant\; C\,E\,\|h\|_2 .$$

Hence, by differentiating again the equation, multiplying by $\partial_{xx}h$ and using the previous estimate, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_{xx}h\| \leqslant C\,E\,\|\partial_{xx}h\| + \|\partial_{xx}f\|\,. \tag{2.3.13}$$

Finally, adding (2.3.13) with (2.3.11) and (2.3.12) and applying the Gronwall Lemma, we obtain (2.3.10) for $k = 2$.

Next, we study the existence of solution for problem (B). We define the characteristic curve X associated to the equation:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}X = \tilde{u}_N(t, X)\,, \\ \\ X(t = t_0) = x_0\,. \end{cases}$$

Then, solution of (B) reads:

$$h(t, X(t, x)) = h(0, X(0, x)) + \int_0^t f(s, X(s, x))\,\mathrm{d}s \quad \forall t \in [0, T]\,.$$

We thus deduce that $h \in \mathcal{C}(0, T; H^1(\mathbb{R})) \cap \mathcal{C}^1(0, T; L^2(\mathbb{R}))$. For $k = 2$, we differentiate (B) with respect to x. Then $\phi := \partial_x h$ is solution of

$$\begin{cases} \partial_t\phi \,+\, \tilde{u}_N\,\partial_x\phi = \partial_x f \,-\, \partial_x\tilde{u}_N\,\phi\,, \\ \\ \phi(0, x) = \partial_x h^0 \,\in\, H^1(\mathbb{R})\,. \end{cases}$$

We solve this initial value problem by the iteration:

$$\phi^{(0)}(t, x) = \partial_x h^0(x)\,,\ \forall\,(t, x) \,\in\, [0, T] \times \mathbb{R}\,,$$

and $\phi^{(j)}$, for $j \geq 1$ is the solution of

$$\begin{cases} \partial_t\phi^{(j)} \,+\, \tilde{u}_N\,\partial_x\phi^{(j)} &=& \partial_x f \,-\, \partial_x\tilde{u}_N\,\phi^{(j-1)}\,, \\ \\ \phi^{(j)}(0, x) &=& \partial_x h^0(x) \quad \forall\, x \,\in\, \mathbb{R}\,. \end{cases}$$

Since

$$\|\partial_x f \,-\, \partial_x\tilde{u}_N\,\phi^{(j-1)}\|_1 \,\leqslant\, \|f\|_2 + C\,E\,\|\phi^{(j-1)}\|_1\,,$$

the approximation $\phi^{(j)}$ lies in $\mathcal{C}(0, T; H^1(\mathbb{R}))$. To get the convergence of $\left(\phi^{(j)}\right)_j$ to $\partial_x h$, we observe that

$$\mathcal{L}_{\tilde{u}_N}\left(\phi^{(j+1)} - \phi^{(j)}\right) = \partial_x\tilde{u}_N\,\left(\phi^{(j)} - \phi^{(j-1)}\right)\,,$$

and apply j times the energy estimate (2.3.10) to get

$$\begin{aligned} \left\|\phi^{(j+1)} - \phi^{(j)}\right\|_1 &\leqslant e^{C_3\,E\,t}\int_0^t e^{-C_3\,E\,\tau}\,C_3\,E\|\phi^{(j)} - \phi^{(j-1)}\|_1\,\mathrm{d}\tau \\ &\leqslant \cdots \leqslant e^{C_3\,E\,t}\,\frac{(C_3\,E\,t)^j}{j!}\,\left(2\,\|\partial_x h^0\|_1 + \int_0^t e^{-C_3\,E\,\tau}\,\|\partial_x f(\tau)\|_1\,\mathrm{d}\tau\right)\,, \end{aligned}$$

which tends to zero as j goes to $+\infty$. This gives the convergence of $(\phi^{(j)})_j$ to $\partial_x h$ in H^1, and then the H^2-regularity of h. □

Combining the previous propositions, we obtain existence for the full linearized problem (2.3.8).

Proposition 3.6. *Let $\tilde{\mathbf{S}} \in \mathcal{C}(0,T;H^1(\mathbb{R}))$ and $\tilde{u}_N, \tilde{h} \in \mathcal{C}(0,T;H^2(\mathbb{R}))$ for some $T>0$. Then the initial value problem*

$$\begin{cases} \partial_t \mathbf{U} - \mu\, \partial_{xx}\mathbf{U} = \tilde{\mathbf{S}}\,, \\ \\ \partial_t h + \tilde{u}_N\, \partial_x h = F - \tilde{h}\, \partial_x u_N\,, \end{cases}$$

has a unique strong solution $(\mathbf{U},h)$ which satisfies:

$$\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2(0,T;\mathbf{H}^3(\mathbb{R}))\,,$$
$$h \in \mathcal{C}(0,T;H^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^1(\mathbb{R}))\,.$$

Moreover, there exist two positive constants K and C, only depending on the topography and the viscosity coefficient, such that, for all t in $[0,T]$:

$$\|(\mathbf{U},h)(t)\|_2\,,\ \left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\, d\tau\right)^{1/2} \leqslant K\, e^{C\,(1+E)^2\,t} \left\{ \|(\mathbf{U}^0,h^0)\|_2 + \left(\int_0^t \|\tilde{\mathbf{S}}(\tau)\|_1^2\, d\tau\right)^{1/2} \right\}, \tag{2.3.14}$$

where $E := max\left\{ \underset{0\leqslant t\leqslant T}{sup}\ \{\|\tilde{u}_N(t)\|_2\}\ ,\ \underset{0\leqslant t\leqslant T}{sup}\ \left\{\|\tilde{h}(t)\|_2\right\}\right\}$.

Proof. We first obtain the energy estimate (2.3.14). On the one hand, we observe that the right hand side of the transport equation verifies

$$f = -\tilde{h}\,\partial_x u_N - \sum_{k=1}^{N-1} \partial_x\left(h_i\, u_i\right) \in \mathcal{C}(0,T;H^1(\mathbb{R})) \cap L^2\left(0,T;H^2(\mathbb{R})\right)\,,$$

and we have the estimate

$$\|f\|_2 \leqslant C_b\,(1+E)\,\|\mathbf{U}\|_3\,.$$

Therefore, applying Cauchy-Schwarz inequality:

$$\int_0^t e^{-C_3\,E\,\tau}\,\|f(\tau)\|_k\,\mathrm{d}\tau \leqslant e^{C_4\,(1+E)^2\,t}\left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\,\mathrm{d}\tau\right)^{1/2}, \tag{2.3.15}$$

for some constant C_4 depending on C_b, C_3. On the other hand, from the inequality (2.3.9) we deduce that there exist two constant depending only on the viscosity C_1', C_2' such that

$$\|\mathbf{U}(t)\|_2,\ \left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\,\mathrm{d}\tau\right)^{1/2} \leq C_1'\, e^{C_2' t}\left[\|\mathbf{U}^0\|_2 + \left(\int_0^t \|\tilde{\mathbf{S}}(\tau)\|_1^2\,\mathrm{d}\tau\right)^{1/2}\right]. \tag{2.3.16}$$

Injecting this estimate in (2.3.15) yields

$$\int_0^t e^{-C_3\,E\,\tau}\,\|f(\tau)\|_k\,\mathrm{d}\tau \leqslant C_1'\, e^{C_5\,(1+E)^2\,t}\left[\|\mathbf{U}^0\|_2 + \left(\int_0^t \|\tilde{\mathbf{S}}(\tau)\|_1^2\,\mathrm{d}\tau\right)^{1/2}\right], \tag{2.3.17}$$

where $C_5 = C_2' + C_4$. Finally, we add (2.3.10) and (2.3.16), and control the exponentials to obtain (2.3.14).

This gives the uniqueness for the solution. Let us now prove the existence of solution. On the one hand, the existence of $\mathbf{U}$ follows from Proposition 3.4, and we have the regularity

$$\mathbf{U} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2(0,T;\mathbf{H}^3(\mathbb{R}))\,,$$

which gives

$$f \in \mathcal{C}(0,T;H^1(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2(0,T;\mathbf{H}^2(\mathbb{R}))\,.$$

So we can apply Proposition 3.5 for $k=1$ to obtain the existence of h in $\mathcal{C}(0,T;H^1(\mathbb{R}))$. To get more regularity in space, we differentiate the problem with respect to x:

$$\begin{cases} \partial_t\left(\partial_x \mathbf{U}\right) - \mu\,\partial_{xx}\left(\partial_x \mathbf{U}\right) = \partial_x \tilde{\mathbf{S}}\,, \\ \\ \partial_t\left(\partial_x h\right) + \tilde{u}_N\,\partial_x\left(\partial_x h\right) = \partial_x f - \partial_x \tilde{u}_N\,\partial_x h\,, \\ \\ \left(\partial_x \mathbf{U}^0,\,\partial_x h^0\right) \in H^1(\mathbb{R})\,. \end{cases}$$

Noticing that

$$\|\partial_x f\|_1 \leqslant C\,E\,\|\partial_x \mathbf{U}\|_1\,,$$

we can solve this problem by the same iteration process as in the lattest part of Proposition 3.5, this concludes the proof. □

In order to obtain the solution to the nonlinear initial value problem (2.3.2), we will build a convergent sequence, this is the last part of the proof of Theorem 2.3.

2.3.3 Iterative scheme

We construct a recursive sequence $(\mathbf{U}^{(j)}, h^{(j)}) = (u_1^{(j)}, \dots, u_N^{(j)}, h^{(j)})_{j\in\mathbb{N}}$ as follows.

$$\forall\,(t,x) \in [0,T]\times\mathbb{R}\,,\ (\mathbf{U}^{(0)}, h^{(0)})(t,x) = (\mathbf{U}^0, h^0)(x) \in \mathbf{H}^2(\mathbb{R})\,,$$

and for all $j\in\mathbb{N}$, $(\mathbf{U}^{(j+1)}, h^{(j+1)})$ solves the initial value problem:

$$\begin{cases} \left(\partial_t - \mu\,\partial_{xx}\right)\left(\mathbf{U}^{(j+1)}\right) = \mathbf{S}^{(j)}\,, \\ \\ \mathcal{L}_{u_N^{(j)}}\left(h^{(j+1)}\right) = F^{(j,j+1)}\,, \\ \\ (\mathbf{U}^{(j+1)}, h^{(j+1)})(t=0) = (\mathbf{U}^0, h^0)\,, \end{cases} \qquad (\mathcal{P}_{j+1})$$

where the sequence of source terms is given by, for any $j\in\mathbb{N}$:

$$\begin{cases} \mathbf{S}^{(j)} = \mathbf{S}\big(\mathbf{U}^{(j)}, h^{(j)}, \partial_x \mathbf{U}^{(j)}, \partial_x h^{(j)}\big)\,, \\ \\ F^{(j,j+1)} = -\displaystyle\sum_{i=1}^{N-1} \partial_x\left(h_i\,u_i^{(j+1)}\right) - \partial_x u_N^{(j+1)}\,h^{(j)}\,. \end{cases}$$

We define the constants

$$\begin{cases} E = 2\,\|(\mathbf{U}^0, h^0)\|_2\,, \\ \\ \eta_0 = \dfrac{1}{2}\inf\limits_{x\in\mathbb{R}} h^0(x)\,. \end{cases}$$

The following lemma gives the existence of the whole sequence.

Lemma 3.7. *For suitably small $T > 0$, the sequence $(\mathbf{U}^{(j)}, h^{(j)})_{j\in\mathbb{N}}$ is well defined and satisfies, for any $t \in [0,T]$ and any $j \in \mathbb{N}$:*

$$\begin{aligned} &\mathbf{U}^{(j)} \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;\mathbf{L}^2(\mathbb{R})) \cap L^2(0,T;\mathbf{H}^3(\mathbb{R}))\,, \\ &h^{(j)} \in \mathcal{C}(0,T;H^2(\mathbb{R})) \cap \mathcal{C}^1(0,T;H^1(\mathbb{R}))\,. \end{aligned} \tag{2.3.18}$$

Moreover, for all (t,x) in $[0,T]\times\mathbb{R}$ and all $j \in \mathbb{N}$, we have:

$$\|(\mathbf{U}^{(j)}, h^{(j)})(t)\|_2\,,\ \left(\int_0^t \|\mathbf{U}^{(j)}(\tau)\|_3^2\,d\tau\right)^{1/2} \leqslant E\,, \tag{2.3.19}$$

$$h^{(j)}(t,x) \geq \eta_0 > 0\,. \tag{2.3.20}$$

Proof. First we initialize the recursion. $(\mathbf{U}^{(0)}, h^{(0)})$ verifies the good conditions by definition. Applying Proposition 3.6, we obtain existence of $(\mathbf{U}^{(1)}, h^{(1)})$ in $\mathcal{C}(0,t;\mathbf{H}^2(\mathbb{R}))$ for any $t > 0$. Moreover, applying the characteristic formula to $h^{(1)}$, we get, for any $t > 0$:

$$\begin{aligned} h^{(1)}(t,y) &= h^0(X(0,t,y)) + \int_0^t F^{(0,1)}(s, X(s,t,y))\,\mathrm{d}s \\ &\geq 2\,\eta_0 + C(E)\,t \\ &\geq \eta_0 \quad \text{if } t \leqslant T_1 \text{ small enough}\,, \end{aligned}$$

where $T_1 = T_1(\eta_0, E)$, which yields (2.3.20) for $j = 1$. It remains to prove (2.3.19). To do so, we write the inequality (2.3.14) given by Proposition 3.6 for $(\mathbf{U}^{(1)}, h^{(1)})$, that is, for any $t \leqslant T_1$:

$$\|(\mathbf{U}^{(1)}, h^{(1)})(t)\|_2\,,\ \left(\int_0^t \|\mathbf{U}^{(j)}(\tau)\|_3^2\,\mathrm{d}\tau\right)^{1/2} \leqslant K\,e^{C\,(1+E)^2\,t}\left\{E/2 + \left(\int_0^t \|\mathbf{S}^{(0)}(\tau)\|_1^2\,\mathrm{d}\tau\right)^{1/2}\right\}.$$

Hence, applying Lemma 3.2 (2.3.3) to $\mathbf{S}^{(0)}$, we obtain

$$\|(\mathbf{U}^{(1)}, h^{(1)})(t)\|_2\,,\ \left(\int_0^t \|\mathbf{U}^{(j)}(\tau)\|_3^2\,\mathrm{d}\tau\right)^{1/2} \leqslant K\,e^{C\,(1+E)^2\,t}\left\{E/2 + C(\eta_0)\,E\,(1+E)\,\sqrt{t}\right\}.$$

Therefore, we can find $0 < T_2 = T_2(\eta_0, E) \leqslant T_1$ such that (2.3.19) is satisfied for any $t \leqslant T_2$. We choose $T := T_2$.

Next we pass from j to $j+1$. If for any j in $\mathbb{N}$, $(\mathbf{U}^{(j)}, h^{(j)})$ satisfies (2.3.18), (2.3.19) and (2.3.20) for any $t \leqslant T_2$, the existence of $(\mathbf{U}^{(j+1)}, h^{(j+1)})$ follows again from Proposition

3.6. Hence it remains to show (2.3.19) and (2.3.20) for $(\mathbf{U}^{(j+1)}, h^{(j+1)})$. As in the previous calculations, we rewrite the energy estimate (2.3.14) satisfied by $(\mathbf{U}^{(j+1)}, h^{(j+1)})$, that is for any $t \leqslant T$:

$$\|(\mathbf{U}^{(j+1)}, h^{(j+1)})(t)\|_2 \leqslant K\, e^{C\,(1+E)^2\,t} \left\{ E/2 + \Big(\int_0^t \|\mathbf{S}^{(j)}(\tau)\|_1^2 \,\mathrm{d}\tau \Big)^{1/2} \right\}.$$

Hence, applying again Lemma 3.2 (2.3.3) to $\mathbf{S}^{(j)}$ and using the bounds of $(\mathbf{U}^{(j)}, h^{(j)})$, it yields:

$$\|(\mathbf{U}^{(j+1)}, h^{(j+1)})(t)\|_2 \leqslant K\, e^{C\,(1+E)^2\,t} \left\{ E/2 + C(\eta_0)\, E \left(1+E\right) \sqrt{t} \right\} \leqslant E,$$

since the same constants as previously are involved. In the same way we get (2.3.20) fo the rank $j+1$, this ends the proof, with $T = T_2$. □

Now, to show that the sequence built above converges, we will prove that it is a Cauchy sequence in some function space. For this sake, for $j \geq 1$, we compute the difference between systems $(\mathcal{P}_{j+1})$ and $(\mathcal{P}_j)$:

$$\left\{ \begin{array}{lcl} (\partial_t - \mu\,\partial_{xx}) \left(\mathbf{U}^{(j+1)} - \mathbf{U}^{(j)} \right) & = & \mathbf{S}^{(j)} - \mathbf{S}^{(j-1)}\,, \\ \\ \mathcal{L}_{u_N^{(j)}} \left(h^{(j+1)} - h^{(j)} \right) & = & F^{(j,j+1)} - F^{(j-1,j)} - \partial_x h^{(j)} \left(u_N^{(j)} - u_N^{(j-1)} \right), \\ \\ \left(\mathbf{U}^{(j+1)} - \mathbf{U}^{(j)}, h^{(j+1)} - h^{(j)} \right)(t=0) & = & \mathbf{0}\,. \end{array} \right. \tag{$\mathcal{D}_j$}$$

Let us rewrite the right hand side of the transport equation, denoted by $\tilde{F}$:

$$\begin{array}{rcl} \tilde{F} & = & F^{(j,j+1)} - F^{(j-1,j)} - \partial_x h^{(j)} \left(u_N^{(j)} - u_N^{(j-1)} \right) \\ \\ & = & -\displaystyle\sum_{i=1}^{N-1} \partial_x \left[h_i \left(u_i^{(j+1)} - u_i^{(j)} \right) \right] - h^{(j)}\, \partial_x \left(u_N^{(j+1)} - u_N^{(j)} \right) \\ & & + \partial_x u_N^{(j)}\, (h^{(j)} - h^{(j-1)}) - \partial_x h^{(j)} \left(u_N^{(j)} - u_N^{(j-1)} \right). \end{array}$$

Thus, since the whole sequence is bounded by E, and by the use of the estimate (2.3.7) and Lemma 3.1 we get:

$$\|\tilde{F}\|_k \leqslant C\,E \left\| \left(\mathbf{U}^{(j)} - \mathbf{U}^{(j-1)}, h^{(j)} - h^{(j-1)} \right) \right\|_k + C_b\,(1+E) \left\| \mathbf{U}^{(j)} - \mathbf{U}^{(j-1)} \right\|_{k+1},$$

for $k = 1$ or 2. From Lemma 3.2 (2.3.6), we have also:

$$\left\| \mathbf{S}^{(j)} - \mathbf{S}^{(j-1)} \right\| \leqslant C(\eta_0)\,(1 + E + E^2) \left\| \left(\mathbf{U}^{(j)} - \mathbf{U}^{(j-1)}, h^{(j)} - h^{(j-1)} \right) \right\|_1.$$

Therefore, the solution to system $(\mathcal{D}_j)$ satisfies the following energy estimate, for any $t \leqslant T$ (where T is given by Proposition 3.7):

$$\left\| \left(\mathbf{U}^{(j+1)} - \mathbf{U}^{(j)}, h^{(j+1)} - h^{(j)} \right)(t) \right\|_1 \leqslant C(E)\, e^{C_b\,(1+E)^2\,t} \left(\int_0^t \left\| \left(\mathbf{U}^{(j)} - \mathbf{U}^{(j-1)}, h^{(j)} - h^{(j-1)} \right)(\tau) \right\|_1^2 \mathrm{d}\tau \right)^{1/2}.$$

Hence there exists a subsequence, still labelled $\left(\mathbf{U}^{(j)}, h^{(j)}\right)_j$ such as

$$\left(\mathbf{U}^{(j)}, h^{(j)}\right) \underset{j\to\infty}{\longrightarrow} (\mathbf{U}, h) \text{ strongly in } \mathcal{C}(0,T;H^1(\mathbb{R}))\,.$$

Moreover, Lemma 3.7 gives, up to a subsequence, the convergence:

$$\mathbf{U}^{(j)} \underset{j\to\infty}{\rightharpoonup} \mathbf{U} \text{ weakly in } L^2\Big(0,T;\mathbf{H}^3(\mathbb{R})\Big)\,,$$

while, for every fixed $t \leqslant T$:

$$\left(\mathbf{U}^{(j)}, h^{(j)}\right)(t) \underset{j\to\infty}{\rightharpoonup} (\mathbf{U}, h)(t) \text{ weakly in } \mathbf{H}^2(\mathbb{R})\,.$$

Thus we have a solution $(\mathbf{U}, h)$ to system (2.3.2), lying in $\mathcal{C}(0,T;H^1(\mathbb{R}))\cap L^\infty\Big(0,T;\mathbf{H}^2(\mathbb{R})\Big)$, satisfying for any t, x:

$$h(t,x) \geq \eta_0 > 0\,,$$
$$\|(\mathbf{U},h)(t)\|_2\,,\ \left(\int_0^t \|\mathbf{U}(\tau)\|_3^2\,\mathrm{d}\tau\right)^{1/2} \leqslant E\,.$$

Finally, we show that $(\mathbf{U}, h) \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R}))$ by regularizing: we consider $(\mathbf{U}^\varepsilon, h^\varepsilon) = (\rho_\varepsilon * \mathbf{U}, \rho_\varepsilon * h)$, where $\rho_\varepsilon *$ is the Friedrichs'mollifier with respect to x. Thus, applying $\rho_\varepsilon *$ to system (2.3.2) we obtain

$$\begin{cases} \partial_t \mathbf{U}^\varepsilon - \mu\,\partial_{xx}\mathbf{U}^\varepsilon = \mathbf{S}^\varepsilon + \mathbf{C}_0^\varepsilon\,, \\ \partial_t h^\varepsilon + u_N^\varepsilon\,\partial_x h^\varepsilon = F^\varepsilon + C_1^\varepsilon\,, \\ (\mathbf{U}^\varepsilon, h^\varepsilon)(t=0) = (\rho_\varepsilon * \mathbf{U}^0, \rho_\varepsilon * h^0) \in \mathcal{C}^\infty\,, \end{cases} \tag{2.3.21}$$

where $\mathbf{S}^\varepsilon = \rho_\varepsilon * \mathbf{S}$, $F^\varepsilon = \rho_\varepsilon * F$ and

$$\begin{cases} \mathbf{C}_0^\varepsilon = (\partial_t - \mu\,\partial_{xx})\,(\mathbf{U}^\varepsilon) - \rho_\varepsilon * (\partial_t - \mu\,\partial_{xx})\,(\mathbf{U})\,, \\ C_1^\varepsilon = \{\partial_t h^\varepsilon - \partial_x\,(h^\varepsilon\,u_N^\varepsilon)\} - \rho_\varepsilon * \{\partial_t h - \partial_x\,(h\,u_N)\}\,. \end{cases}$$

By classical arguments on mollifiers [146, 174], we have, as ε goes to zero:

$$\begin{cases} \mathbf{C}_0^\varepsilon\,,\ C_1^\varepsilon \ \to\ 0\,, \\ (\mathbf{U}^\varepsilon\,,\ h^\varepsilon) \ \to\ (\mathbf{U}, h)\,. \end{cases}$$

Therefore, at the uniform limit we have $(\mathbf{U}, h) \in \mathcal{C}(0,T;\mathbf{H}^2(\mathbb{R}))$. Uniqueness follows from the energy estimate and this concludes the proof of Theorem 2.3.

Chapitre 3

Finite Volume discretization and numerical simulations of the multilayer shallow water model

In this chapter, we propose a discretization of system (2.1.10) with a Finite Volume method in Section 3.1. Then we present some numerical results in Section 3.2: we compare the model to classical shallow water models as well as with the primitive system when there is no viscosity. Moreover, we illustrate the dynamic behavior of our model, with additions and substractions of layers.

3.1 Numerical scheme

We present in this section the discrete version of the system (2.1.10). Several strategies are possible. For example, in [12, 13, 58], the authors perform an upwind scheme based on approximate Riemann state solvers. In [14, 15], the authors use a kinetic formulation. Here, we will simply use a Finite Volume scheme [133, 134], by isolating an hyperbolic part of the system, for which we can evaluate exact eigenvalues, without computing eigenvectors. The lawfulness of this choice can be discussed but the results obtained are good, and the simplicity of the scheme makes it easy to implement, while the code is dynamic: we can add or remove layers when the uppest layer height becomes too large or too small.

In order to design a numerical scheme, we will consider a third formulation of the one dimensional multilayer problem (2.1.10). This time, the unknowns are denoted by $\mathbf{V}$, lying in $\mathbb{R}^{N+1}$, and $\mathbf{W}$ in $\mathbb{R}^N$:

$$= \big(V_i\big)_{0\leqslant i\leqslant N} = (h, h\,u_N, u_1, \dots, u_{N-1})^T \ , \quad \mathbf{W} = (w_{1/2}, \dots, w_{N-1/2})^T \, .$$

Hence, we separate the viscous terms: the horizontal one is included in the flux term with respect to x, and the vertical one is kept in the source term. The formulation reads as:

$$\begin{cases} \partial_t \mathbf{V} + \partial_x \mathbf{F}(\mathbf{V}) = \mathbf{S}(\mathbf{V}, \mathbf{W}) \,, \\ \\ w_{1/2} = u_1 \, \partial_x z_b \,, \\ \\ w_{i+1/2} - w_{i-1/2} = -h_i \, \partial_x u_i \,, \quad 1 \leqslant i \leqslant N-1 \,. \end{cases} \tag{3.1.1}$$

The flux term $\mathbf{F} \in \mathbb{R}^{N+1}$ then comprises two parts: a convective part $\mathbf{F}^C$ corresponding to the transport and a diffusive one $\mathbf{F}^D$ corresponding to the horizontal viscosity. Precisely, we write its ith coordinate, for $0 \leq i \leq N$, as follows:

$$F_i \, = \, F_i^C + F_i^D \,,$$

where

$$\begin{cases} F_0^C = h \, u_N \,, \\ \\ F_1^C = h \, u_N^2+ g \, h^2/2 \,, \\ \\ F_i^C = u_{i-1}^2 + g \, h \,, \ 2 \leqslant i \leqslant N \,. \end{cases} \qquad \begin{cases} F_0^D = 0 \,, \\ \\ F_1^D = -\mu \, h \, \partial_x u_N \,, \\ \\ F_i^D = -\mu \, \partial_x u_{i-1} \,, \ 2 \leqslant i \leqslant N \,. \end{cases}$$

The source term $\mathbf{S} = (S_i)_{0 \leqslant i \leqslant N}$ is composed of three parts, coming from different effects, namely $\mathbf{G} = \mathbf{S}^b + \mathbf{S}^v + \mathbf{S}^e$. First the topography source term $\mathbf{S}^b$ is given by

$$\begin{cases} S_0^b = 0 \,, \\ \\ S_1^b = -g \, h \, \partial_x z_b \,, \\ \\ S_2^b = \left(\dfrac{u_1^2}{h_1} - g \right) \partial_x z_b \,, \\ \\ S_i^b = -g \, \partial_x z_b \,, \ 3 \leqslant i \leqslant N \,. \end{cases}$$

Second, $\mathbf{S}^v$ represents the terms coming from the vertical viscosity and the friction:

$$\begin{cases} S_0^v = 0 \,, \\ \\ S_1^v = -2 \, \mu \, \dfrac{u_N - u_{N-1}}{h + h_{N-1}} \,, \\ \\ S_2^v = 2 \, \mu \, \dfrac{u_2 - u_1}{h_1 \, (h_1 + h_2)} - \dfrac{\kappa}{h_1} \, u_1 \,, \\ \\ S_i^v = 2 \, \mu \, \dfrac{u_i - u_{i-1}}{h_{i-1} \, (h_i + h_{i-1})} - 2 \, \mu \, \dfrac{u_{i-1} - u_{i-2}}{h_{i-1} \, (h_{i-1} + h_{i-2})} \,, \ 3 \leqslant i \leqslant N \,. \end{cases}$$

Finally $\mathbf{S}^e$ is the mass exchange term and is given by:

$$\begin{cases} S_0^e = w_{N-1/2}\,, \\ S_1^e = u_{N-1/2}\, w_{N-1/2}\,, \\ S_2^e = -\dfrac{1}{h_1}\, u_{3/2}\, w_{3/2}\,, \\ S_i^e = \dfrac{1}{h_{i-1}} \left(u_{i-1/2}\, w_{i-1/2} - u_{i-3/2}\, w_{i-3/2} \right)\,,\ 3 \leqslant i \leqslant N\,. \end{cases}$$

The system is completed with the initial condition

$$\mathbf{V}(0,x) = \mathbf{V}^0(x) = \left(h^0, h^0\, u_N^0, u_1^0, \ldots, u_{N-1}^0 \right)^T\,. \tag{3.1.2}$$

Therefore, using this formulation we base the construction of the scheme on the following remark dealing with hyperbolicity of some part of the system (3.1.1).

Remark 1.1. The system (3.1.1) when replacing the right hand side by 0 and taking $\mu = 0$ (that is only considering the transport flux $\mathbf{F}^C$), is hyperbolic and the eigenvalues of the Jacobian matrix J of the flux read as:

$$\mathcal{SP}(J) = \left\{ u_N + \sqrt{gh}, u_N - \sqrt{gh}, 2u_1, \cdots, 2u_{N-1} \right\}\,.$$

It is simply seen when we compute the Jacobian J:

$$\frac{\partial \mathbf{F}}{\partial \mathbf{V}} = J(\mathbf{V}) = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ gh - u_N^2 & 2u_N & 0 & 0 & \cdots & 0 \\ g & 0 & 2u_1 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & 2u_2 & 0 & \cdots \\ \vdots & \vdots & \cdots & \cdots & \ddots & \cdots \\ g & 0 & \cdots & \cdots & \cdots & 2u_{N-1} \end{pmatrix}\,.$$

The eigenvalues are seen immediately. Moreover, the eigenvectors are computed easily, and the Jacobian matrix does not degenerate in a non diagonalizable matrix when some velocities become equal, as long as we have $h > 0$.

Remark 1.2. This remark will be used to design the numerical scheme, but it is not really a property of hyperbolicity for the full system (2.1.10), since the source term of the formulation (3.1.1) contains derivatives of the unknowns u_i. Nevertheless the numerical results obtained with this formulation are totally lawful.

We now present the discrete version of the system (3.1.1). We use calligraphic letters to name the numerical scheme. We introduce a space-time discretization based on a uniform grid of points $x_{j+1/2}$ with space step Δx and on a grid of points $t_k = k\,\Delta t$ with a time

step Δt wich will be precised later through a CFL condition. The finite volume method consists in integrating the system on each control cell $C_j = (x_{j-1/2}; x_{j+1/2})$ of the mesh and each time step, and approximating the fluxes at the interfaces.

First, initial data $\left(\mathcal{V}_i^0\right)_{0\leqslant i\leqslant N}$ are computed, as usual in the finite volume framework, as the averaged values of (3.1.2) through each space cell, that is, for all $i \in \{0,\dots,N\}$, for all j in $\mathbb{Z}$,

$$\mathcal{V}_{i,j}^0 = \frac{1}{\Delta x}\int_{C_j} V_i^0(x)\,\mathrm{d}x\,.$$

Then, the semi discrete numerical scheme reads, for all i in $\{0,\dots,N\}$, for all j in $\mathbb{Z}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathcal{V}_{i,j} + \frac{1}{\Delta x}\left(\mathcal{F}_{i,\,j+1/2} - \mathcal{F}_{i,\,j-1/2}\right) = \mathcal{G}_{i,\,j}\,,$$

where $\mathcal{F}_{i,\,j+1/2}$ is the approximation of the ith coordinate of the flux at the cell interface $x_{j+1/2}$, and $\mathcal{G}_{i,\,j}$, $\mathcal{V}_{i,\,j}$ are the approximation of the mean value of the ith coordinate of $\mathbf{G}$, $\mathbf{V}$ on the cell C_j.

Let us now describe the numerical flux, sum of the convective part and the diffusive part, first without slope limiters. On the one hand we choose the Lax-Friedrichs flux for the "hyperbolic" part, that is for all i in $\{0,\dots,N\}$, for all j in $\mathbb{Z}$:

$$\mathcal{F}^C_{i,\,j+1/2} = \frac{1}{2}\left[F_i^C(\mathcal{V}_{j+1}) + F_i^C(\mathcal{V}_j) - a_\infty\left(\mathcal{V}_{i,\,j+1} - \mathcal{V}_{i,\,j}\right)\right],$$

where $a_\infty = \sup\left\{|\lambda|, \lambda \in \mathcal{SP}(J)\right\}$. On the other hand, the diffusive flux, essentially an approximation of the gradient of the velocities at the cell interfaces, is discretized classically. The coordinates of the discrete unknown $\mathcal{V}$ are expressed in terms of water height and velocities:

$$\mathcal{V}_0 = \mathcal{H}\,,\ \mathcal{V}_1 = \mathcal{H}\,\mathcal{U}_N\,,\ \mathcal{V}_i = \mathcal{U}_{i-1}\ \text{ for }\ 2 \leq i \leq N\,.$$

Therefore

$$\begin{cases} \mathcal{F}^D_{0,\,j+1/2} = 0\,, \\ \\ \mathcal{F}^D_{1,\,j+1/2} = -\mu\,\mathcal{H}_{j+1/2}\,\dfrac{\mathcal{U}_{N,\,j+1} - \mathcal{U}_{N,\,j}}{\Delta x}\,, \\ \\ \mathcal{F}^D_{i,\,j+1/2} = -\mu\,\dfrac{\mathcal{U}_{i-1,\,j+1} - \mathcal{U}_{i-1,\,j}}{\Delta x}\,,\ 2 \leqslant i \leqslant N\,, \end{cases}$$

where $\mathcal{H}_{j+1/2} = \left(\mathcal{H}_{j+1} + \mathcal{H}_j\right)/2$. Before treating the source terms, we shall impose the CFL stability condition. This necessary condition appears in the next calculations: we take the source term equal to zero and consider an explicit Euler scheme in time for sake of simplicity. Then, with the choice of three points explicit fluxes we made, we may write the numerical solution at time t^{n+1}, $\mathcal{V}_{i,j}^{n+1}$ as a combination of $\mathcal{V}_{i,j+1}^n$, $\mathcal{V}_{i,j}^n$ and $\mathcal{V}_{i,j-1}^n$. Precisely, we have, for $j \in \mathbb{Z}$, for $0 \leq i \leq N$:

$$\mathcal{V}_{i,j}^{n+1} \;=\; \alpha_{i,j}^n\,\mathcal{V}_{i,j-1}^n + \beta_{i,j}^n\,\mathcal{V}_{i,j}^n + \gamma_{i,j}^n\,\mathcal{V}_{i,j+1}^n\,,$$

where

$$\left\{\begin{array}{l} \alpha_{i,j}^n = \partial_i F_i^C(\xi_i^n)\,\dfrac{\Delta t}{2\,\Delta x} + a_\infty\,\dfrac{\Delta t}{2\,\Delta x} + \mu\,\dfrac{\Delta t}{\Delta x^2}\,\delta_{i,j-1/2}^n\,, \\ \\ \beta_{i,j}^n = 1 - a_\infty\,\dfrac{\Delta t}{\Delta x} - \mu\,\dfrac{\Delta t}{\Delta x^2}\,\left(\delta_{i,j-1/2}^n + \delta_{i,j+1/2}^n\right)\,, \\ \\ \gamma_{i,j}^n = -\partial_i F_i^C(\xi_i^n)\,\dfrac{\Delta t}{2\,\Delta x} + a_\infty\,\dfrac{\Delta t}{2\,\Delta x} + \mu\,\dfrac{\Delta t}{\Delta x^2}\,\delta_{i,j+1/2}^n\,, \end{array}\right.$$

$$\text{with}\quad \gamma_{i,j+1/2}^n = \left\{\begin{array}{ll} 0 & \text{if}\quad i = 0\,, \\ \mathcal{H}_{j+1/2}^n & \text{if}\quad i = 1\,, \\ 1 & \text{if}\quad 2 \le i \le N\,. \end{array}\right.$$

Now, with the definition of a_∞ and the fact that the size of the uppest layer is of order $O\left(\overline{h}\right)$ strictly smaller than 1, we can require the following CFL condition:

$$a_\infty\,\frac{\Delta t}{\Delta x} + 2\,\mu\,\frac{\Delta t}{\Delta x^2} \;<\; 1\,. \tag{CFL}$$

Actually, we will rather use an explicit Runge Kutta of order 4 scheme in time instead of an explicit Euler scheme, and the time step is recomputed after each step to statisfy the stability condition (CFL). Moreover, we introduce slope limiters in the fluxes to reconstruct the unknowns $\mathcal{V}_i$ on the right $(+)$ and on the left $(-)$ of the cell interfaces $x_{j+1/2}$. We use the classical minmod limiters $\sigma_{i,\,j}$, that is for $0 \le i \le N$, for $j \in \mathbb{Z}$:

$$\sigma_{i,\,j} = \;\text{minmod}\;\left(\mathcal{V}_{i,\,j+1} - \mathcal{V}_{i,\,j}, \mathcal{V}_{i,\,j} - \mathcal{V}_{i,\,j-1}\right).$$

Then the right and left reconstructed values at interface $x_{j+1/2}$ of the ith coordinate of $\mathcal{V}$ read:

$$\left\{\begin{array}{l} \mathcal{V}_{i,\,j+1/2}^+ = \mathcal{V}_{i,\,j+1} - \sigma_{i,\,j+1}\;\left(\mathcal{V}_{i,\,j+2} - \mathcal{V}_{i,\,j+1}\right)\,, \\ \\ \mathcal{V}_{i,\,j+1/2}^- = \mathcal{V}_{i,\,j} + \sigma_{i,\,j}\;\left(\mathcal{V}_{i,\,j+1} - \mathcal{V}_{i,\,j}\right)\,. \end{array}\right.$$

Hence the convective flux including the slope limiters reads, for $0 \le i \le N$, for $j \in \mathbb{Z}$:

$$\mathcal{F}_{i,\,j+1/2}^C = \frac{1}{2}\left[F_i^C(\mathcal{V}_{j+1/2}^+) + F_i^C(\mathcal{V}_{j+1/2}^-) - a_\infty\left(\mathcal{V}_{i,\,j+1/2}^+ - \mathcal{V}_{i,\,j+1/2}^-\right)\right].$$

We do the same to compute the diffusive flux $\mathcal{F}^D$. Let us now define the discrete $\mathcal{W}$ variable. Since the vertical velocity $\mathcal{W}_{i+1/2,j}$ is essentially an horizontal gradient of the velocity at the layer interface $z_{i+1/2}$ and in the horizontal cell C_j, we choose the following reconstruction for $j \in \mathbb{Z}$, for $1 \leqslant i \leqslant N-1$:

$$\left\{\begin{array}{l} \mathcal{W}_{1/2,j} = \mathcal{V}_{3,j}\,\partial_x z_b(x_j)\,, \\ \\ \mathcal{W}_{i+1/2,j} \;= \mathcal{W}_{i-1/2,j} - h_i\,\dfrac{1}{2\,\Delta x}\,\left\{\left(\mathcal{U}_{i,j+1/2}^+ + \mathcal{U}_{i,j+1/2}^-\right) - \left(\mathcal{U}_{i,j-1/2}^+ + \mathcal{U}_{i,j-1/2}^-\right)\right\}\,. \end{array}\right.$$

Finally we choose a smooth topographic source term, for which we can compute the derivative exactly. This avoids numerical difficulties that can occur with the discretization of the source term $\mathbf{S}^b$ (which can be handled by different methods, see for example [13, 98]).

Remark 1.3. We can mention that an explicit scheme in time may be very restrictive because of the viscous terms. Nevertheless, we consider a small viscosity coefficient, which reduces the constraint on the time step. Moreover, the instantaneous regularizing effect of the viscous terms makes the nonconservative product $h\,\partial_x u_N$ well defined at the first time step, even if we start with a discontinuous initial water height as in Section A.2.

Remark 1.4. We can note that this numerical scheme preserves at the discret level the conservation of the mass $\int h\,\mathrm{d}x$ and the constant steady states with periodic boundary condition. We will also see in the numerical results that it captures non constant steady states in the last test case performed (see Section 3.2.5).

3.2 Numerical experiments

In this section we present numerical simulations performed to validate numerically the multilayer system and its discretization. We also show its possible dynamic behavior, depending on the wanted accuracy in the vertical direction, which is given by the choice of the amplitude $\overline{h}$ of the inside fixed layers. In all the tests performed, the CFL number is equal to 0.95.

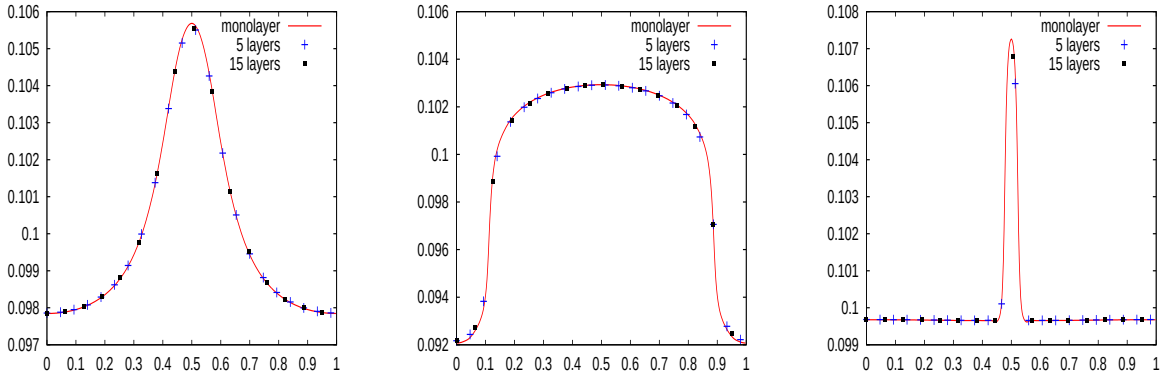
3.2.1 Test 1: perturbation of rest in height, comparison with classical Shallow Water model, flat bottom

In this section, we perturb the lake at rest in height. We take gravity constant $g = 1$, viscosity $\mu = 0.0001$, no friction, computational domain $[0, 1]$ and periodic boundary conditions. We compare first our model to existing models, and then investigate its dynamic behavior. We consider the initial condition

$$\begin{cases} \eta(t=0,x) = 0.1 + 0.01\ (1 + 0.8\ \cos(2\,\pi\,x/L))\ , \\ \\ u_i(t=0,x) \equiv 0 \ \text{ for } \ 1 \leqslant i \leqslant N\,, \end{cases} \tag{3.2.1}$$

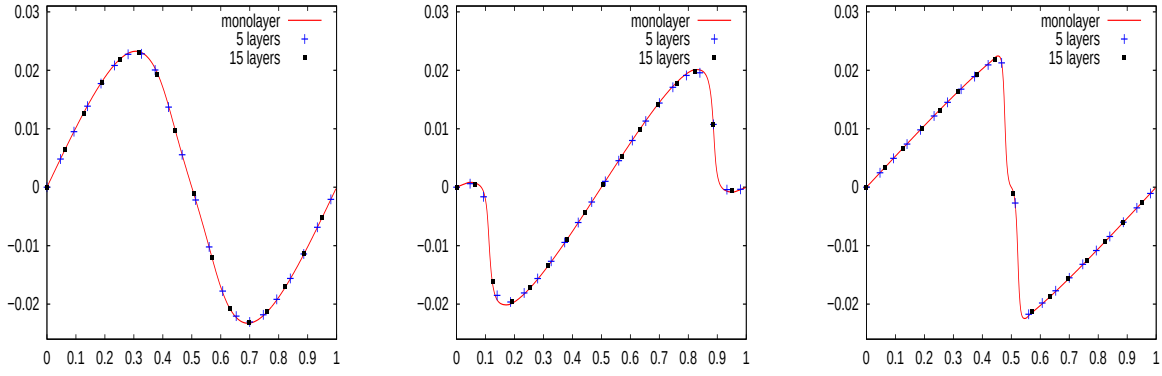
where the number of layers N will be either 5 ($\overline{h} = 0.02$) or 15 ($\overline{h} = 0.006$). Then we plot the solution to the multilayer scheme for 5 layers (+) and 15 layers (•), as well as the result obtained with a simple global Lax-Friedrichs scheme for the classical viscous shallow water system (**??**) (lines), for different times. Figure 3.1 shows the evolution of the free surface, and Figure 3.2 shows the evolution of the averaged horizontal velocity.

On the one hand, we observe that the results fit well with the classical one layer model, whatever the number of layers. Hence our multilayer model is at least as good as the classical Saint-Venant model. On the other hand, we do have an advantage when considering the multilayer model: information on the vertical velocity. Indeed, we see in Figure 3.3 the velocity field inside the water for 15 layers. This information can not be recovered by the classical shallow water model.



(a) (b) (c)

Figure 3.1: *Evolution of the free surface for initial condition (3.2.1),* $t = 5$ *(a),* $t = 10$ *(b),* $t = 20$ *(c)*



(a) (b) (c)

Figure 3.2: *Evolution of the mean velocity for initial condition (3.2.1) ,* $t = 5$ *(a),* $t = 10$ *(b),* $t = 20$ *(c)*

3.2.2 Test 2: perturbation of rest in height, zero viscosity, comparison with the full Euler hydrostatic model, flat bottom

In this test case, we compare the multilayer model with the hydrostatic system in the zero viscosity case, which reads, for all $t > 0$, for all (x, z) in Ω_t:

$$\begin{cases} \partial_t u + u\,\partial_x u + w\,\partial_z u + g\,\partial_x H = 0\,, \\ \\ \partial_t H + u(z = H)\,\partial_x H = w(z = H)\,, \end{cases}$$

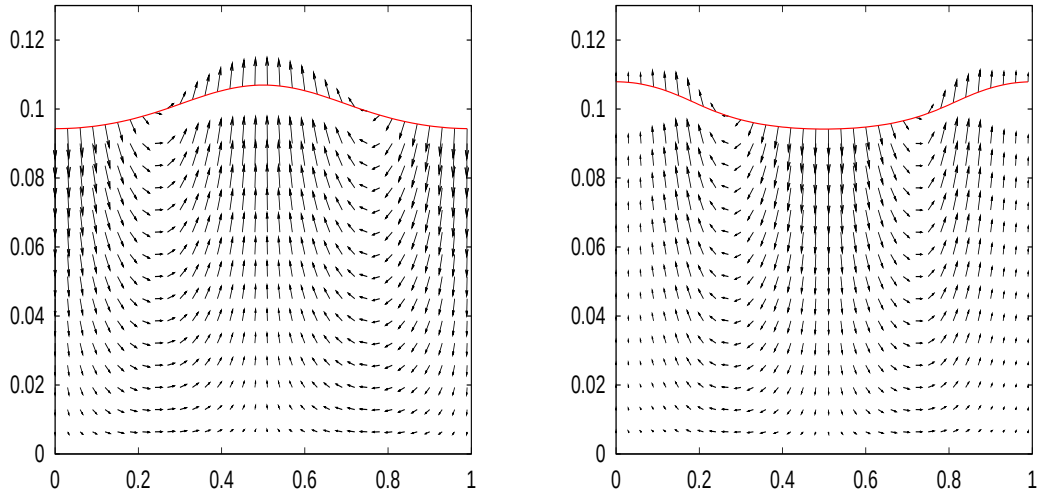


Figure 3.3: *Evolution of the velocity field for initial condition (3.2.1) and* 15 *layers, at times* $t = 1$ *(a) and* $t = 10$ *(b)*

completed with the non penetration condition at the bottom and the divergence free condition to reconstruct the vertical velocity:

$$w(t,x,0) = 0\,, \quad \forall\,(t,x)\ \in \mathbb{R}^+ \times \mathbb{R}\,,\ \partial_x u + \partial_z w = 0\,.$$

In order to design our numerical scheme for this system, we introduce a tracer parameter $0 \leqslant a \leqslant 1$ of particules of fluids in the vertical direction [38]. Thus the vertical position of one particle of fluid at time t and horizontal position x reads:

$$z = Z(t,x,a)\,,\ 0 \leqslant a \leqslant 1\,.$$

Then, we have $Z(t,x,a=0) = 0$ for particles at the bottom, and $Z(t,x,a=1) = H(t,x)$ at the free surface. Moreover we ask the advection equation at the free surface to be satisfied for any a, namely:

$$\partial_t Z(t,x,a) + u\left(t,x,Z(t,x,a)\right)\,\partial_x Z(t,x,a) = w\left(t,x,Z(t,x,a)\right)\,. \tag{3.2.2}$$

Hence, noting $c = \partial_a Z$ and injecting these new variables in the primitive equations, we get an hyperbolic system posed in a fixed domain $\Omega = \mathbb{R}^+ \times \mathbb{R} \times [0,1]$. It reads, for all $(t,x,a)\ \in \Omega$:

$$\partial_t \tilde{u}(t,x,a) + \partial_x \left(\frac{\tilde{u}^2}{2}\right)(t,x,a) = -g\,\partial_x Z(t,x,1)\,, \tag{3.2.3}$$

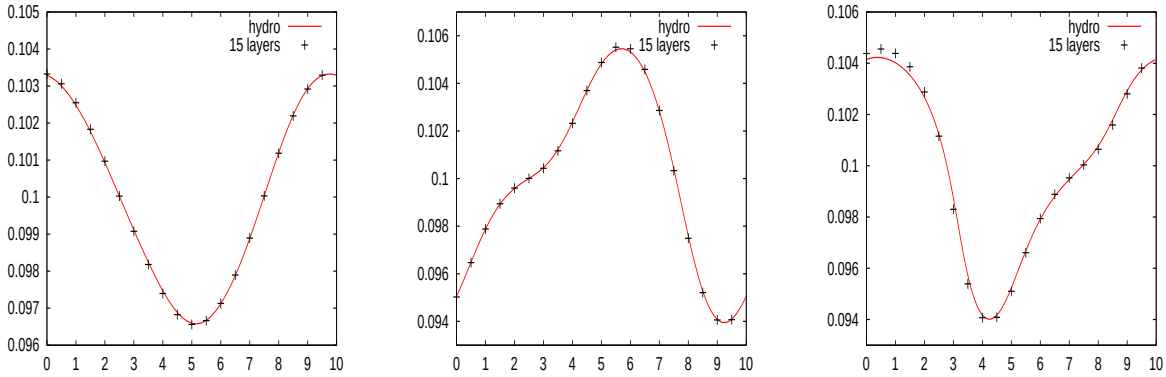
$$\partial_t c(t,x,a) + \partial_x\left(c\,\tilde{u}\right)(t,x,a) = 0\,, \tag{3.2.4}$$

where $\tilde{u}(t,x,a) = u\left(t,x,Z(t,x,a)\right)$ and $\tilde{w}(t,x,a) = w\left(t,x,Z(t,x,a)\right)$. Boundary conditions are obtained in the same way. This formulation allows to perform a finite volume scheme.

We take periodic boundary conditions, no friction and the computational domain is $[0,1]$. We run the multilayer code for 15 layers ($\overline{h} = 0.006$) and 40 layers ($\overline{h} = 0.002$), with the following initial conditions

$$\begin{cases} \eta(0,x) = 0.1 + 0.01\left(1 + 0.5\,\cos(2\,\pi\,x/L) + 0.2\,\sin(4\,\pi\,x/L)\right), \\ \\ u_i(0,x) \equiv 0 \;\text{ for }\; 1 \leqslant i \leqslant N\,, \end{cases} \tag{3.2.5}$$

and corresponding initial datum for the Euler system. In Figures 3.4 and 3.5 we output the profiles of the free surface and the mean velocity at different times. We can observe that curves fit well. Nevertheless, since the Lagrangian formulation of the Euler equations is only valid on short times, the hydrostatic code becomes unstable and starts to oscillate after around 100 iterations.



(a) (b) (c)

Figure 3.4: *Evolution of the free surface , $t = 10$ (a), $t = 40$ (b), $t = 80$ (c)*

The velocity fields are also plotted in Figure 3.6 at time $t = 50$ for both models. The profiles are in good agreement.

3.2.3 Test 3: perturbation of rest in velocity, flat bottom

In this section, we again consider a flat bottom, the spatial domain is $[0,1]$, viscosity $\mu = 0.0001$, no friction, periodic boundary conditions and we perturb the lake at rest in

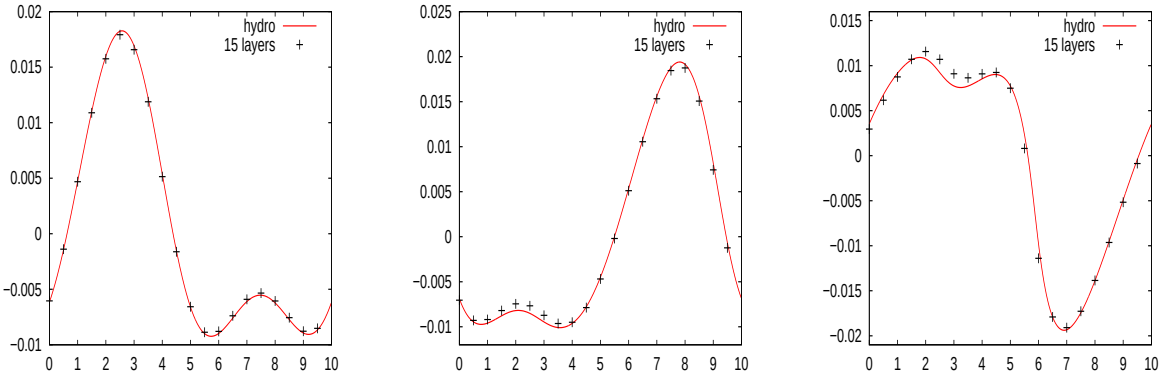


Figure 3.5: *Evolution of the mean velocity* , $t = 10$ *(a),* $t = 50$ *(b),* $t = 100$ *(c)*

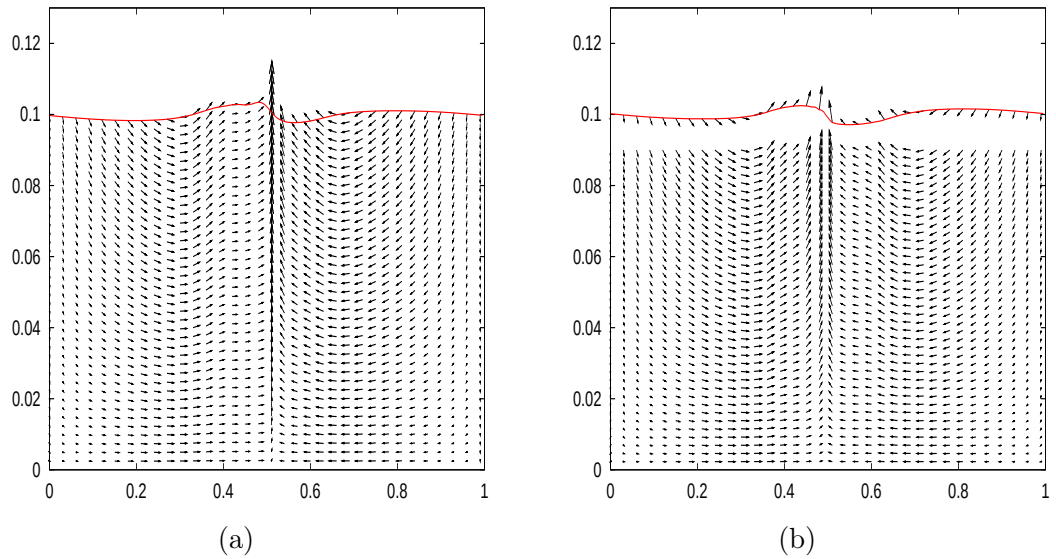


Figure 3.6: *Velocity field, hydrostatic (a),* 40 *layers (b)*

velocity, taking for initial conditions:

$$\begin{cases} \eta(0,x) \equiv 0.1\,, \;\; u_i(0,x) \equiv 0 \;\text{ for }\; 1 \leqslant i \leqslant N-1\,, \\ \\ u_N(0,x) = 0.2\,\sin(2\,\pi\,x)\,. \end{cases} \tag{3.2.6}$$

Then we run the multilayer code for different numbers of layers: 5 (amplitude $\overline{h} = 0.02$) and 15 layers ($\overline{h} = 0.006$). Thus, in Figure 4.1, we show the evolution in time of the

free surface and the velocity field inside the fluid, for these two choices. We observe the multilayer aspects of the velocity field, that is appearance of vorteces, becoming more visible when we take a larger number of layers.

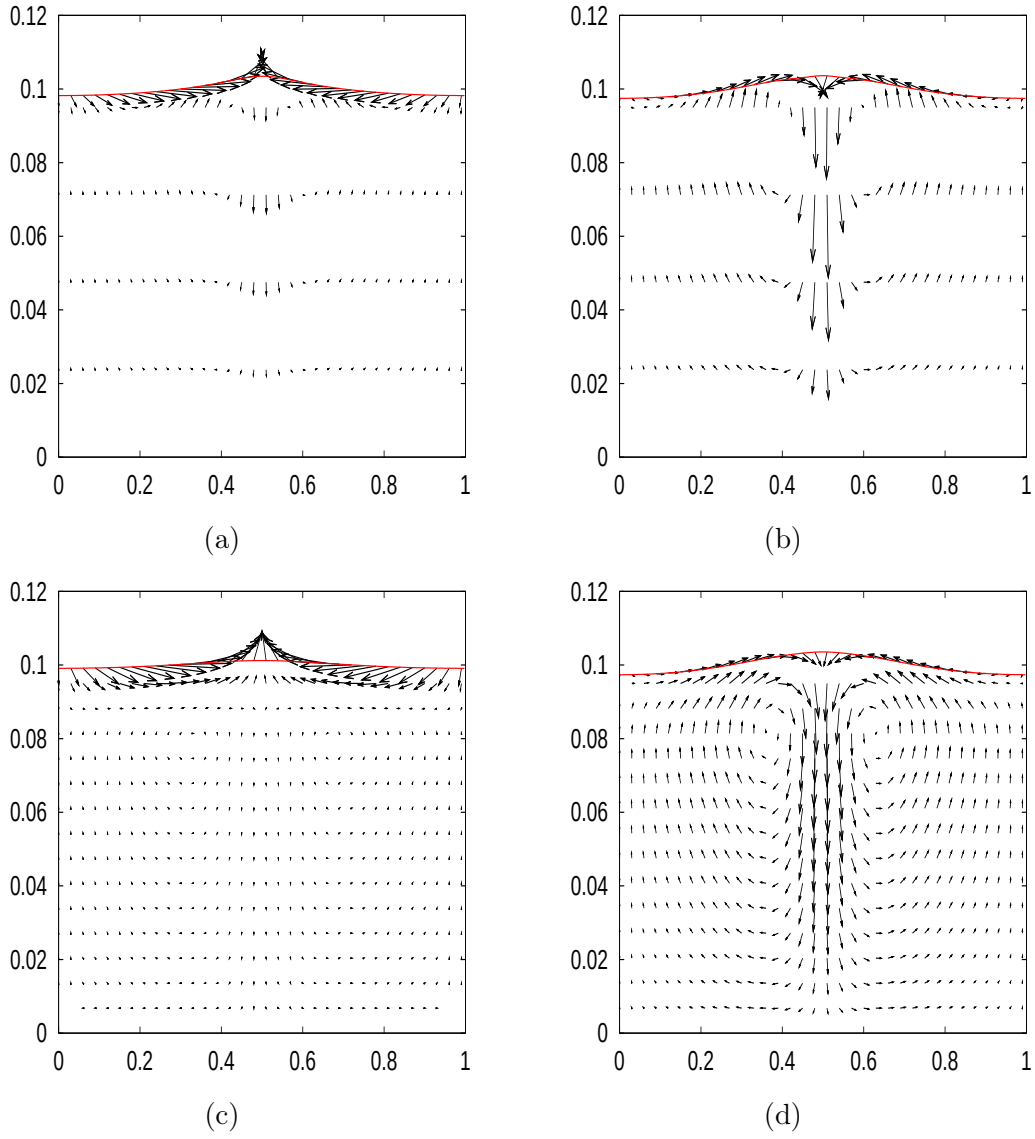


Figure 3.7: *Evolution of the velocity field for initial data (3.2.6)for* $t = 1$ *(left) and* $t = 10$ *(right):* 5 *layers (a),(b), and* 15 *layers (c), (d).*

3.2.4 Test 4: perturbation of rest, periodic bottom, dynamic behavior

For this test case we consider a spatial domain $[-10, 10]$, viscosity $\mu = 0.0009$, no friction, size of the inside layers $\overline{h} = 0.08$ and a smooth rapidly oscilatting bottom

$$z_b(x) = 0.1\ \sin(6\,\pi\,x/20)\,.$$

We perturb first the rest with a sinusoidal height, with a larger period, that is:

$$\begin{cases} \eta(0,x) = 1 + 0.07\ \sin(2\,\pi\,x/20)\,, \\ \\ u_i(0) \equiv 0 \ \text{ for } \ 1 \leqslant i \leqslant N\,. \end{cases} \tag{3.2.7}$$

We then output in Figure 3.8 the velocity field and the free surface at different times. We observe a transition period, during which occurs a lot of movement inside the fluid and variations of the amplitude of the free surface. This yields variations of the number of layers, initially equal to 11. Next, after 200 seconds, the behavior of the fluid seems to stabilize: the free surface tends to be symmetric to the topography and the velocities inside the fluid are getting smaller.

3.2.5 Test 5: subcritical flow over a bump

In this test case, we want to observe numerically the convergence to the steady state. For further information on this test as more general topography (that is discontinuous) we refer to [13, 11, 47, 49, 50, 82, 98]. In this work, since we avoid the difficulty of discretization of the topographic source term, we take a smooth bottom given by:

$$z_b(x) = -0.1\,\Big(\,\tanh(3\,(x-1)) + \tanh(-3\,(x+1))\Big)\,.$$

We consider the following initial and boundary conditions:

$$\begin{cases} \eta(0,x) = \eta_{out} \equiv 1\,, \ \ u_i(0,x) \equiv 0 \ \text{ for } \ 1 \leqslant i \leqslant N-1 \\ \\ h\,u_N(0,x) = (h\,u_N)_{in} \equiv 0.06\,. \end{cases} \tag{3.2.8}$$

We take for this test: spatial domain $[-10, 10]$, 200 points, $\mu = 0.0005$, no friction and two different numbers of initial layers: 15 (amplitude $\overline{h} = 0.06$) and 8 ($\overline{h} = 0.12$). We output the evolution of the free surface and the velocity field in Figures 3.9 (15 layers) and 3.10 (7 layers) at different times, the steady state is reached at $t = 45$.

We observe the number of layers reducing while the free surface, initially at rest, goes to the steady state, symmetric to the topography, whatever the size of the inside layers.

3.3 Conclusion

As a conclusion, in Chapters 2 and 3, we propose another approach to approximate free surface incompressible geophysical flows, by a multilayer shallow water type model, taking advantage of the numerical efficiency of such systems. This model is not aimed at dealing with wet/dry fronts, but rather deep water. It is derived from the hydrostatic primitive equations, which are widely used to model ocean motions. In particular no technical assumption on the regime of viscosity is made to make the derivation of the model rigorous.

Moreover, a local in time existence result is obtained, with a classical restriction of positivity on the uppest layer height. This result allows to get a dynamic behavior of the model in the numerical simulations. Indeed, when the uppest layer height becomes too small (resp. too large) we remove (resp. add) one layer, and so on.

At this time the coupling of this multilayer model with flat bottom, together with classical shallow water over a sloped plan, modelling the arrival of a small wave to the beach is under study.

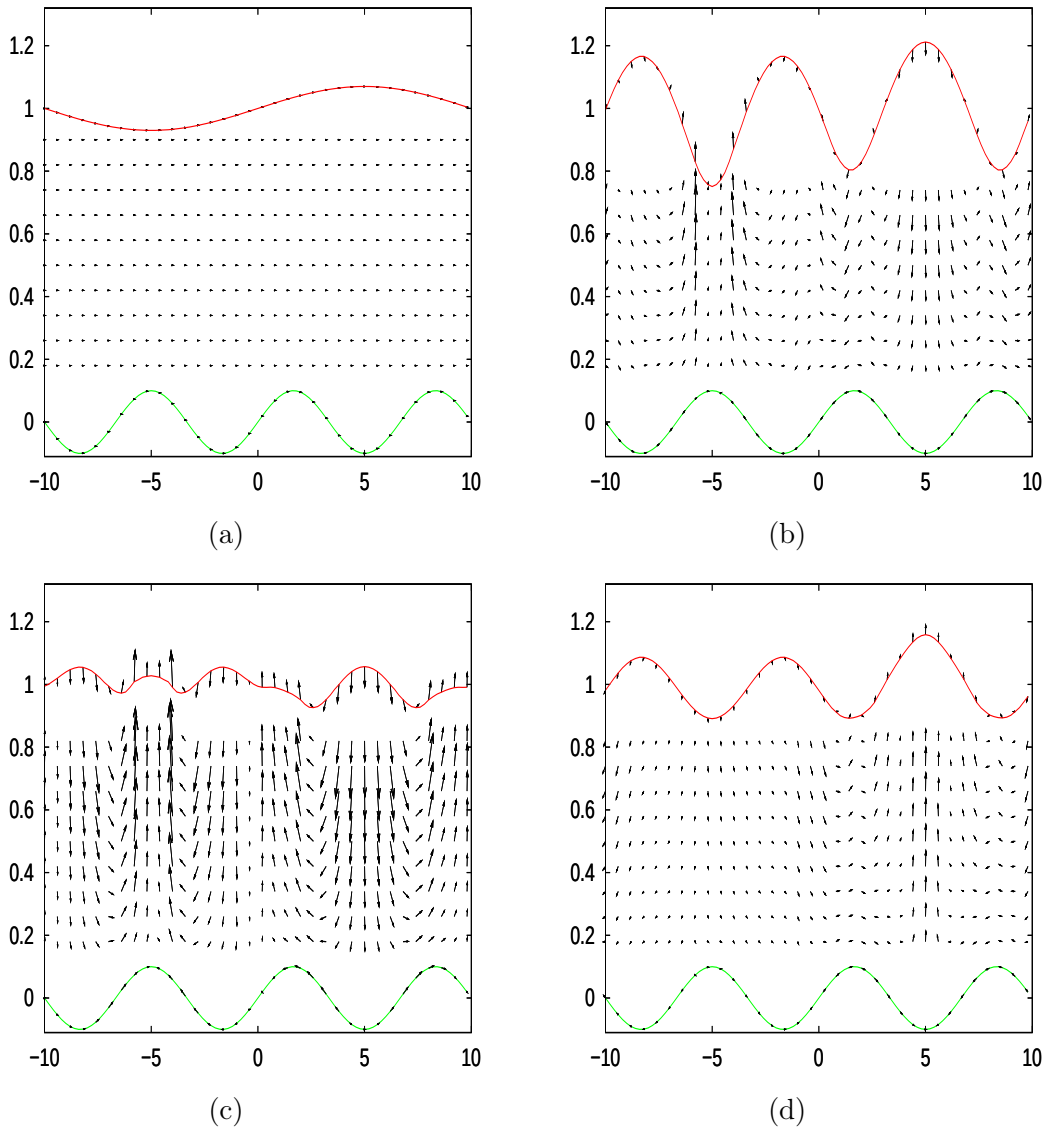


Figure 3.8: *Free surface and velocity field for initial data (3.2.7) at times* $t = 0$ *(11 layers)(a),* $t = 10$ *(9 layers) (b),* $t = 20$ *(10 layers) (c) and final time* $t = 100$ *(10 layers)(d).*

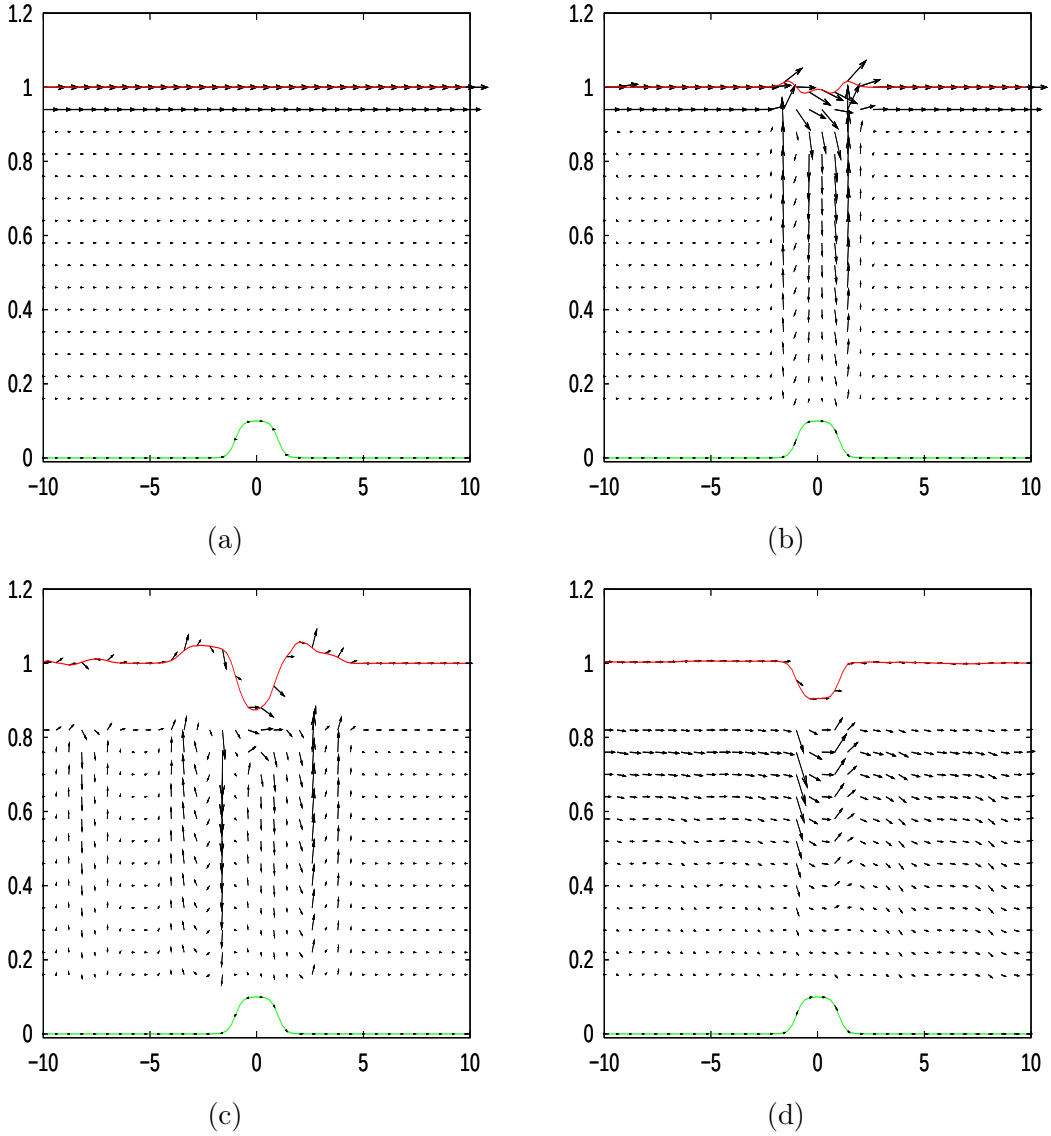


Figure 3.9: *Free surface and velocity field, subcritical flow, for in-out conditions (3.2.8),* 15 *layers initially, at times* $t = 0$ *(*15 *layers)(a),* $t = 0.4$ *(*15 *layers)(b),* $t = 2$ *(*13 *layers)(c), and final time* $t = 45$ *(*13 *layers)(d).*

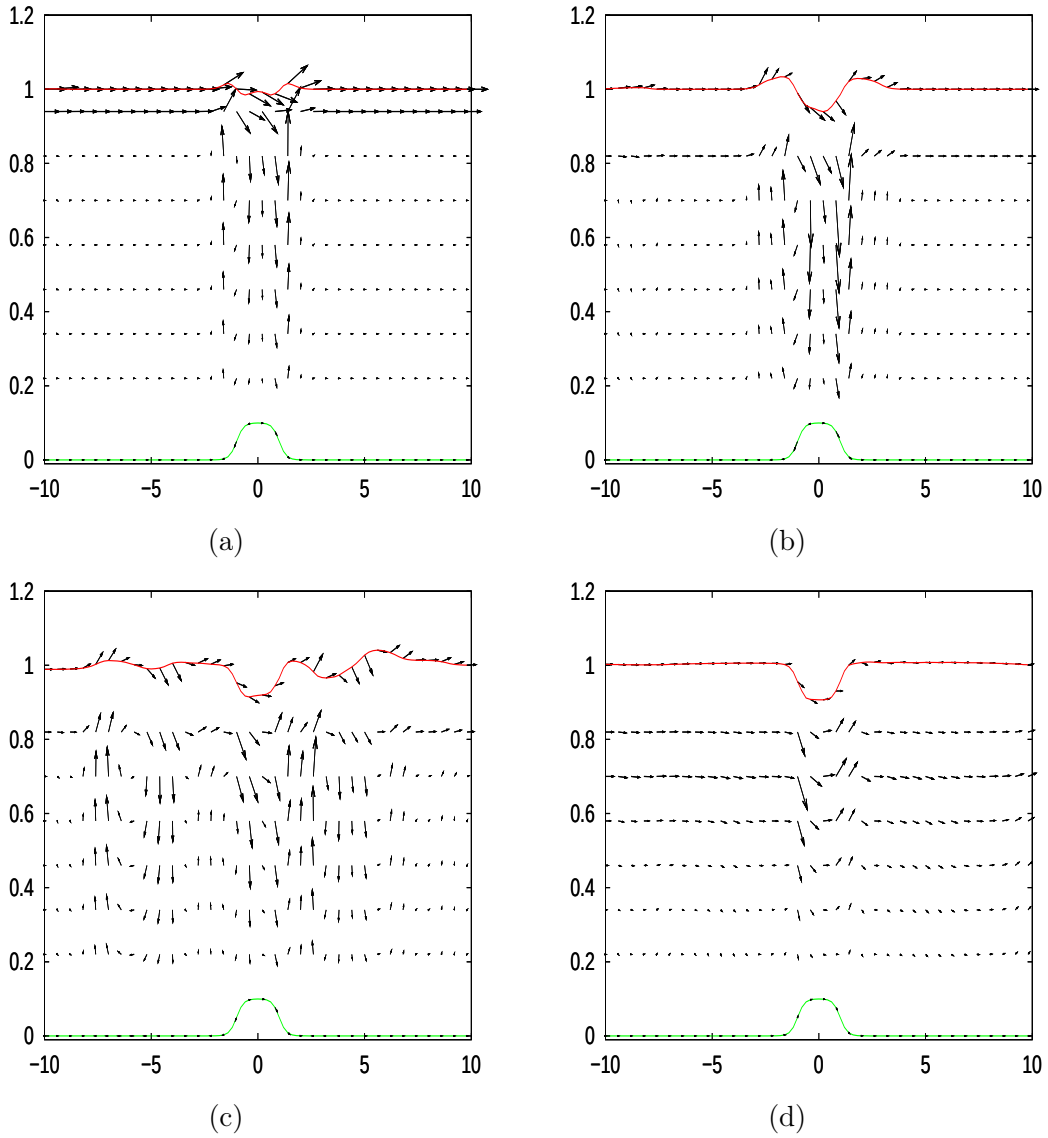


Figure 3.10: *Free surface and velocity field, subcritical flow, for in-out conditions (3.2.8), 8 layers initially, at times $t = 0.4$ (8 layers)(a), $t = 1$ (7 layers)(b), $t = 5$ (7 layers)(c) and final time $t = 45$ (7 layers)(d).*

Annexe A

Compléments sur l'étude du modèle multicouche

Dans cette annexe, nous regroupons quelques compléments sur l'étude de notre nouveau modèle multicouche de type Saint-Venant. Nous établissons à la Section A.1 une inégalité d'énergie naturelle associée à notre modèle (en 1D), et expliquons les difficultés à généraliser l'estimation de *BD-entropie* établie par D. Bresch et B. Desjardins [43] pour un modèle classique de Saint-Venant. Enfin, nous proposons quelques simulations numériques supplémentaires en 1D, notamment de solutions discontinues, à la Section A.2.

A.1 Sur l'énergie du système multicouche en 1D

Dans cette section, afin de simplifier l'écriture, nous revenons à notre modèle multicouche 1D et considérons une bathymétrie triviale ainsi que la définition des vitesses $u_{i+1/2}$ comme moyenne arithmétique de u_{i+1} et u_i (voir la remarque **??**). L'équation « générique » pour la couche i $(1 \leqslant i \leqslant N)$ s'écrit donc

$$\begin{aligned} \partial_t (h_i\, u_i) + \partial_x \left(h_i\, u_i^2\right) + g\, h_i\, \partial_x H &= \mu\, \partial_x (h_i\, \partial_x u_i) \\ &\quad + u_{i-1/2}\, w_{i-1/2} - u_{i+1/2}\, w_{i+1/2} \\ &\quad - 2\,\mu\, \frac{u_i - u_{i-1}}{h_i + h_{i-1}} + 2\,\mu\, \frac{u_{i+1} - u_i}{h_i + h_{i+1}}\,, \end{aligned} \tag{A.1.1}$$

avec les conventions :

$$\begin{cases} u_{1/2} = 0,\ u_{i+1/2} = \dfrac{u_{i+1} + u_i}{2} \ \text{ pour } 1 \leqslant i\, leqslant N-1,\ u_{N+1/2} = 0, \\ w_{1/2} = 0,\ w_{i+1/2} = -\overline{h} \displaystyle\sum_{k=1} i\, \partial_x u_k \ \text{ pour } 1 \leqslant i\, leqslant N-1,\ w_{N+1/2} = 0\,. \end{cases} \tag{A.1.2}$$

Pour les approximations des dérivées verticales, nous imposons :

$$\begin{cases} 2\,\mu\,\dfrac{u_1 - u_0}{h_1 + h_0} \;=\; \kappa u_1\,, \\[2ex] 2\,\mu\,\dfrac{u_{N+1} - u_N}{h_N + h_{N+1}} \;=\; 0\,. \end{cases}$$

Dans les deux sous-sections suivantes, nous recherchons des entropies pour notre système, qui pourront s'avérer utiles dans la recherche de solutions faibles. Cependant tous les calculs ici sont formels, il s'agit simplement de tenter d'adapter de manière heuristique les travaux effectués dans [43, 144].

A.1.1 Une estimation d'énergie naturelle

Nous obtenons d'abord une inégalité d'énergie *naturelle* pour notre système, qui est en adéquation avec l'estimation obtenue pour les modèles multicouches existants [12, 15]. En considérant des conditions aux limites périodiques (*i.e.* en écrivant le système pour $(t,x) \in \mathbb{R}^+ \times \mathbb{T}$), les solutions régulières satisfont l'estimation d'énergie suivante.

$$\partial_t \int_{\mathbb{T}} E\,\mathrm{d}x + 2\,\mu \int_{\mathbb{T}} \left(\sum_{i=1}^{N-1} \frac{(u_{i+1} - u_i)^2}{h_i + h_{i+1}} + \sum_{i=1}^{N} h_i\,(\partial_x u_i)^2 \right) \mathrm{d}x \;=\; -\kappa \int_{\mathbb{T}} u_1^2\,\mathrm{d}x\,, \quad \text{(A.1.3)}$$

où l'énergie E est définie par :

$$E \;=\; \frac{1}{2}\left(g\,H^2 + \sum_{i=1}^{N} h_i\,u_i^2 \right)\,.$$

Avant de montrer cette inégalité, remarquons que notre fonction énergie est la même que celle introduite dans [12, 15] : c'est la somme des énergies de chaque couche, qui correspondent chacune à l'énergie d'un système classique de Saint-Venant, à savoir :

$$E = \sum_{i=1}^{N} E_i \;\text{ avec }\; E_i = \frac{1}{2}\,h_i\,u_i^2 + g\,h_i\,\partial_x h_i\,.$$

Obtention de l'estimation d'énergie. Les calculs sont classiques : il s'agit de multiplier les équations (A.1.1) par u_i, de les intégrer sur le tore et de les sommer. Nous commençons par le membre de gauche de (A.1.1), et plus précisément le terme

$$A_i \;:=\; \partial_t\,(h_i\,u_i) + \partial_x\left(h_i\,u_i^2\right)\,.$$

En le multipliant par u_i et en intégrant sur $\mathbb{T}$, nous obtenons :

$$\begin{aligned}\int_{\mathbb{T}} A_i\, u_i &= \int_{\mathbb{T}} \left(u_i^2 \partial_t h_i + h_i\, u_i\, \partial_t u_i + h_i\, u_i^2\, \partial_x u_i + u_i^2\, \partial_x\left(h_i\, u_i\right)\right) \\ &= \frac{1}{2}\int_{\mathbb{T}} \Big[\left(u_i\, \partial_t(h_i\, u_i) + h_i\, u_i\, \partial_t u_i\right) + \left(h_i\, u_i^2\, \partial_x u_i + u_i\, \partial_x\left(h_i\, u_i^2\right)\right) \Big] \\ &+ \frac{1}{2}\int_{\mathbb{T}} u_i^2\, \left(\partial_t h_i + \partial_x\left(h_i\, u_i\right)\right) .\end{aligned}$$

En sommant de 1 à N, cela donne :

$$\begin{aligned}\sum_{i=1}^{N}\int_{\mathbb{T}} A_i\, u_i &= \int_{\mathbb{T}} \frac{1}{2}\Big[\partial_t \sum_{i=1}^{N}\left(h_i\, u_i^2\right) + \partial_x \sum_{i=1}^{N}\left(h_i\, u_i^3\right)\Big] \\ &+ \frac{1}{2}\int_{\mathbb{T}} \sum_{i=1}^{N} u_i^2\, \left(\partial_t h_i + \partial_x\left(h_i\, u_i\right)\right) .\end{aligned}$$

Nous pouvons alors nous servir de la conservation de la masse, qui s'écrit sous la forme :

$$\partial_t h + \partial_x\left(h\, u_N\right) = w_{N-1/2}\,.$$

Par ailleurs, en utilisant (A.1.2), nous pouvons écrire :

$$\partial_x\left(h_i\, u_i\right) = w_{i-1/2} - w_{i+1/2}\,.$$

Ainsi l'équation précédente devient :

$$\begin{aligned}\sum_{i=1}^{N}\int_{\mathbb{T}} A_i\, u_i &= \int_{\mathbb{T}} \frac{1}{2}\,\partial_t \sum_{i=1}^{N}\left(h_i\, u_i^2\right) + \frac{1}{2}\int_{\mathbb{T}} \sum_{i=1}^{N} u_i^2\, \left(w_{i-1/2} - w_{i+1/2}\right) \\ &= \int_{\mathbb{T}} \frac{1}{2}\,\partial_t \sum_{i=1}^{N}\left(h_i\, u_i^2\right) + \frac{1}{2}\int_{\mathbb{T}} \sum_{i=1}^{N-1} w_{i+1/2}\, \left(u_{i+1}^2 - u_i^2\right) .\end{aligned}$$

Le dernier terme du membre de gauche de (A.1.1) est traité simplement :

$$\begin{aligned}\sum_{i=1}^{N}\int_{\mathbb{T}} g\, h_i\, \partial_x H\, u_i &= -g\int_{\mathbb{T}} H \sum_{i=1}^{N} \partial_x\left(h_i\, u_i\right) \\ &= +\tfrac{1}{2}\, g\, \textstyle\int_{\mathbb{T}} \partial_t\left(H^2\right) .\end{aligned}$$

Ensuite, pour le second membre de (A.1.1), les termes de viscosité horizontale se traitent également par intégration par parties :

$$\sum_{i=1}^{N}\int_{\mathbb{T}} \mu\, \partial_x\left(h_i\, \partial_x u_i\right)\, u_i = -\mu\int_{\mathbb{T}} \sum_{i=1}^{N} h_i\, \left(\partial_x u_i\right)^2 .$$

Enfin, pour les derniers termes nous allons à nouveau faire des changements d'indices dans les sommes. Nous notons

$$\left\{ \begin{array}{l} B_i \;=\; u_{i-1/2}\, w_{i-1/2} - u_{i+1/2}\, w_{i+1/2}\,, \\ \\ C_i \;=\; -2\,\mu\, \dfrac{u_i - u_{i-1}}{h_i + h_{i-1}} + 2\,\mu\, \dfrac{u_{i+1} - u_i}{h_i + h_{i+1}}\,. \end{array} \right.$$

Alors

$$\sum_{i=1}^{N} \int_{\mathbb{T}} B_i\, u_i \;=\; \sum_{i=1}^{N-1} \int_{\mathbb{T}} u_{i+1/2}\, w_{i+1/2}\, (u_{i+1} - u_i)\,,$$

$$\sum_{i=1}^{N} \int_{\mathbb{T}} C_i\, u_i \;=\; -2\,\mu \sum_{i=1}^{N-1} \int_{\mathbb{T}} \frac{(u_{i+1} - u_i)^2}{h_i + h_{i+1}} - \kappa \int_{\mathbb{T}} u_1^2\,.$$

Pour conclure, il suffit de remarquer que

$$\frac{1}{2} \int_{\mathbb{T}} \sum_{i=1}^{N-1} w_{i+1/2}\, \left(u_{i+1}^2 - u_i^2\right) \;=\; \sum_{i=1}^{N-1} \int_{\mathbb{T}} u_{i+1/2}\, w_{i+1/2}\, (u_{i+1} - u_i)\,,$$

grâce à la définition des $u_{i+1/2}$.

A.1.2 Estimations supplémentaires ? Existence de solutions faibles ?

Si l'on s'intéresse à l'étude des solutions faibles du modèle multicouche 1D, il est nécessaire de choisir le « bon espace » fonctionnel et d'établir des estimations fournissant assez de compacité sur les solutions. En s'inspirant des solutions faibles pour le système classique de Saint-Venant visqueux (voir [144] par exemple), nous pouvons chercher des solutions avec la régularité suivante pour tout $T > 0$:

$$u_i \;\in\; L^2\left(0, T; L^2(\mathbb{T})\right)\,,$$

$$u_i \;\in\; L^\infty\left(0, T; L^\infty(\mathbb{T})\right) \cap L^\infty\left(0, T; H^1(\mathbb{T})\right) \cap L^2\left(0, T; H^2(\mathbb{T})\right)\,.$$

Une attention particulière doit également être portée sur le choix des conditions initiales, par exemple :

$$\sqrt{h^0} \;\in\; H^1(\mathbb{T})\,,\; h^0 \geqslant 0\,,$$

$$\sqrt{h_i}\, u_i \;\in\; L^2(\mathbb{T})\,.$$

Evidemment, puisque nous ne considérons pas ici les effets de capillarité ni de friction turbulente, notre première estimation d'énergie ne donne pas les mêmes estimations que dans [43, 144].

L'estimation *a priori* d'énergie (A.1.3) établie précédemment permet d'obtenir de la compacité sur plusieurs termes si l'on s'intéresse à l'étude des solutions faibles du système multicouche, à savoir :

$$\begin{cases} H \in L^\infty\left(0,T;L^2(\mathbb{T})\right), \\ \sqrt{h_i}\,u_i \in L^\infty\left(0,T;L^2(\mathbb{T})\right), \\ \sqrt{h_i}\,\partial_x u_i \in L^2\left(0,T;L^2(\mathbb{T})\right), \\ u_1 \in L^2\left(0,T;L^2(\mathbb{T})\right), \\ \dfrac{u_{i+1}-u_i}{\sqrt{h_{i+1}+h_i}} \in L^2\left(0,T;L^2(\mathbb{T})\right). \end{cases} \tag{A.1.4}$$

Nous pouvons également obtenir d'autres estimations, à savoir :

$$\begin{cases} \partial_x h \in L^2\left(0,T;L^2(\mathbb{T})\right), \\ u_i \in L^2\left(0,T;L^2(\mathbb{T})\right). \end{cases} \tag{A.1.5}$$

Néanmoins, le contrôle sur h n'est pas suffisant pour passer à la limite dans les termes non linéaires de (A.1.1) (voir [43] pour la dimension 2 et [144] pour la dimension 1), notamment les termes $h_i\,u_i^2$. L'étape cruciale est d'obtenir de la compacité forte L^2 sur $\sqrt{h_i}\,u_i$. C'est ce qui est délicat dans notre cas. En effet, considérons notre système avec seulement deux couches. La conservation de la masse s'écrit :

$$\partial_t\left(h_1+h_2\right)+\partial_x\left(h_1\,u_1\right)+\partial_x\left(h_2\,u_2\right)=0\,.$$

Si l'on oublie un instant que la hauteur h_1 de la couche inférieure est constante, on peut écrire, en dérivant par rapport à x et en remplaçant $\partial_x h_i$ par $h_i\,\partial_x\left(\log h_i\right)$:

$$\partial_t\left(h_1\,\partial_x\log h_1+h_2\,\partial_x\log h_2\right)+\partial_x\left(h_1\,\partial_x u_1+h_2\,\partial_x u_2\right)+\partial_x\left(h_1\,u_1\,\partial_x\log h_1+h_2\,u_2\,\partial_x\log h_2\right)=0\,.$$

L'idée dans la dérivation de la BD-entropie est de s'affranchir des termes visqueux en multipliant la dérivée de l'équation de la masse par μ et de l'ajouter à l'équation sur la quantité de mouvement. La nouvelle équation décrit l'évolution d'une quantité de mouvement pour une vitesse auxiliaire

$$v_i := u_i+\mu\,\partial_x\log h_i\,.$$

Ensuite, la même technique que pour l'obtention de l'inégalité d'énergie (A.1.3) est employée (multiplication par v_i, intégration sur $\mathbb{T}$ et utilisation de la conservation de la masse et d'intégrations par parties). Nous ne pouvons pas utiliser ici la même méthode, car les deux vitesse u_1 et u_2 sont couplées dans la conservation de la masse.

A.2 D'autres simulations numériques

Nous terminons cette annexe en présentant des simulations numériques supplémentaires de notre modèle multicouche pour des solutions discontinues. Nous considérons des

conditions aux limites périodiques ainsi que les paramètres numériques suivant : viscosité $\mu = 0.001$, friction $\kappa = 0$, gravité $g = 1$, $CFL = 0.95$, domaine spatial $[0, 1]$ ($L = 1$) et 400 points de discrétisation. Nous perturbons le « lac au repos » :

$$\begin{cases} H(x) \equiv 0.1\,, \\ u_i(x) \equiv 0 \ \text{ pour } \ 1 \leqslant i \leqslant N\,, \end{cases}$$

d'abord en vitesse, puis en hauteur.

A.2.1 Perturbation du lac au repos en vitesse, fond plat

Nous choisissons d'abord une vitesse initiale de la couche supérieure discontinue :

$$u_N(0,x) = \begin{cases} -0.2 & \text{si } \ 0.75\,(L/2) \leqslant x \leqslant L/2\,, \\ +0.2 & \text{si } \ L/2 < x \leqslant 1.25\,(L/2)\,, \\ 0 & \text{sinon}\,. \end{cases}$$

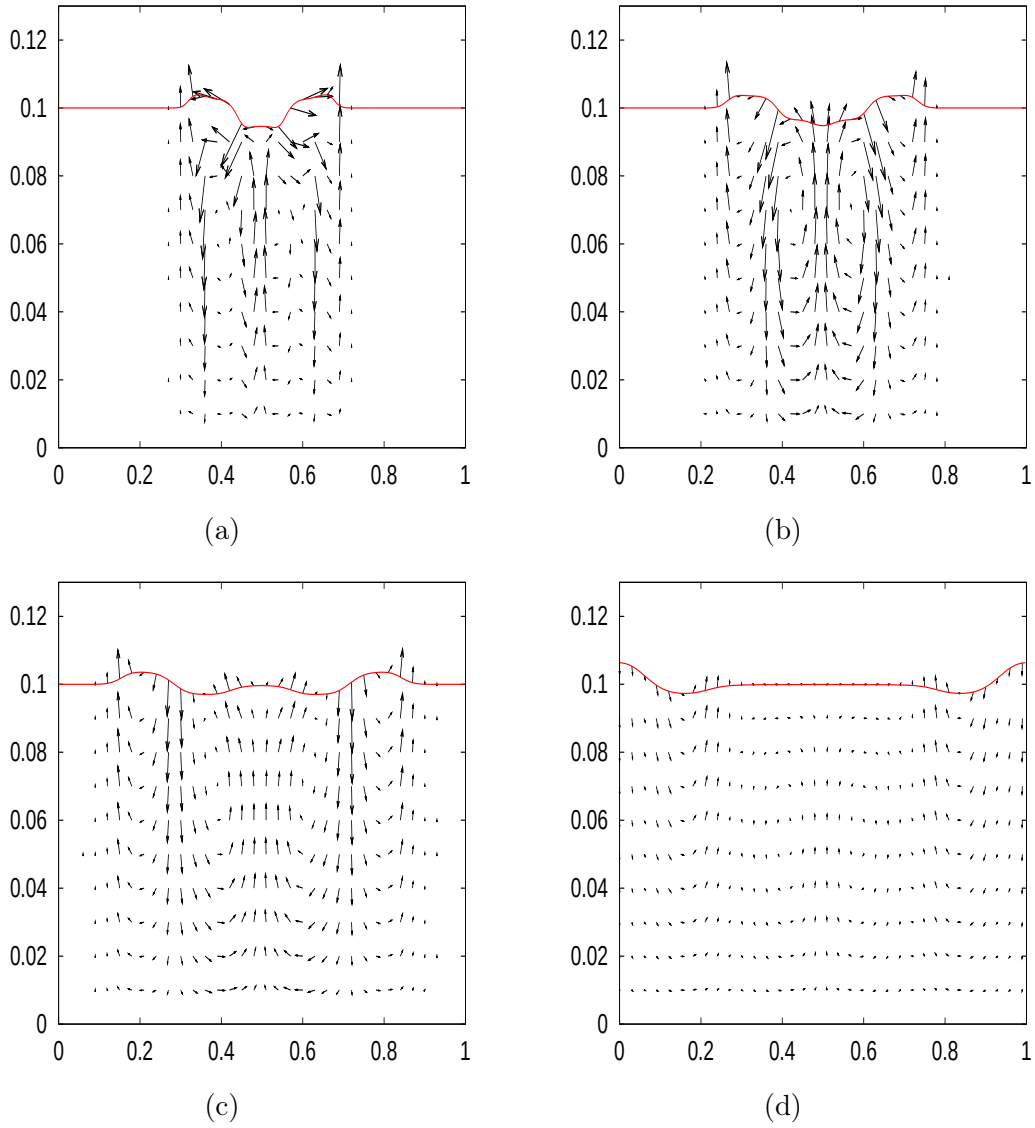


FIGURE A.1 – $N = 10$, *évolution du champs des vitesses :* $t = 2$ *(a),* $t = 4$ *(b),* $t = 8$ *(c),* $t = 16$ *(d)*

Nous observons dans la Figure A.1 des recirculations à l'intérieur de fluide : la perturbation à la surface entraîne des conséquences dans les couches internes au cours du temps, avant que l'équilibre soit rétabli.

A.2.2 Perturbation du lac au repos en hauteur, fond plat

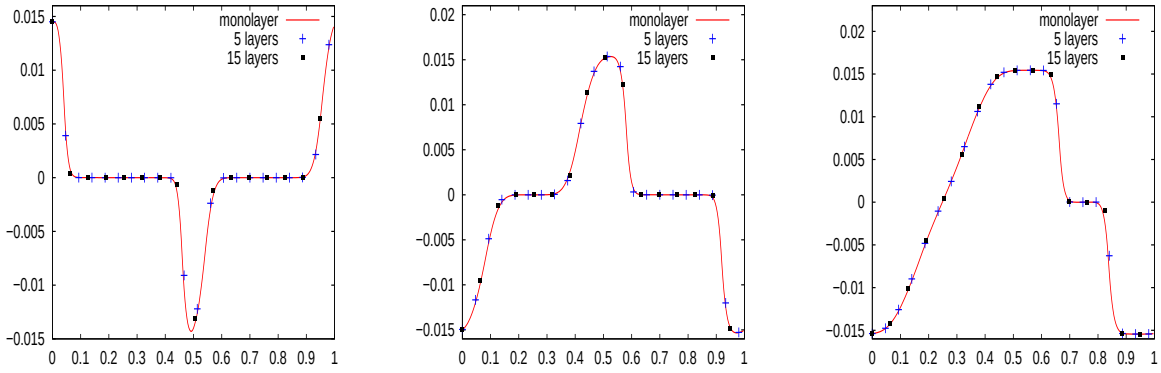
Nous perturbons ensuite le repos en hauteur afin de pouvoir comparer le modèle avec le modèle classique de Saint-Venant, comme c'est fait au Chapitre 3 pour des solutions

régulières. Les conditions initiales sont données par :

$$u_i(0,x) \equiv 0 \ \text{ pour } \ 1 \leqslant i \leqslant N\,,$$

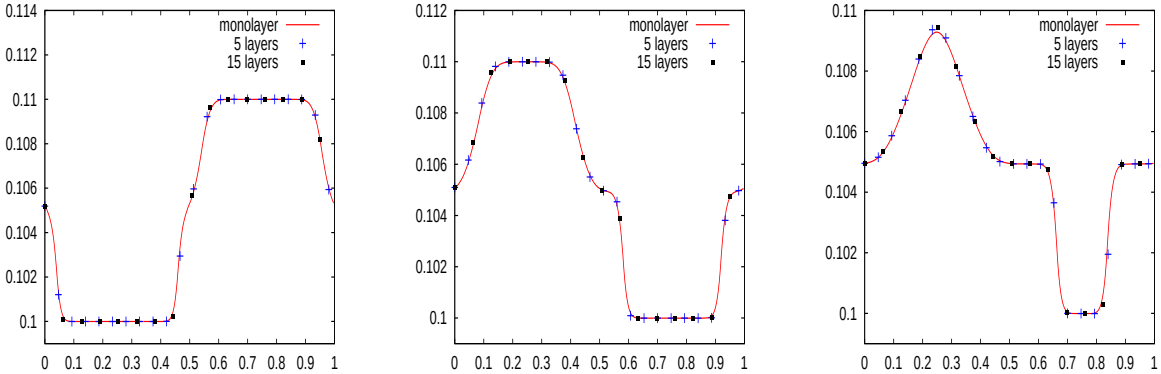
$$h_N(0,x) = \begin{cases} 0.02 & \text{if } \ x \leqslant L/2\,, \\ 0.01 & \text{if } \ x > L/2\,, \end{cases}$$

Nous pouvons alors comparer les vitesses moyennes (Figure A.2) et la surface libre (figure A.3) pour 1, 5 et 15 couches.



(a) (b) (c)

FIGURE A.2 – *Evolution de la vitesse moyenne , $t = 2$ (a), $t = 4$ (b), $t = 8$ (c)*



(a) (b) (c)

FIGURE A.3 – *Evolution de la surface libre , $t = 2$ (a), $t = 4$ (b), $t = 8$ (c)*

Deuxième partie

Construction et analyse d'un schéma préservant l'asymptotique pour les systèmes de relaxation

Chapitre 4

Analysis of an Asymptotic Preserving Scheme for Relaxation Systems

4.1 Introduction

Many physical problems are governed by hyperbolic conservation laws with non vanishing stiff source terms. These problems can describe the effect of relaxation toward an equilibrium, as in the kinetic theory of gases for example. Actually for monatomic gases, when an equilibrium state is perturbed, it gradually relaxes to the equilibrium state with Maxwellian velocity distribution. In the continuum theory of nonmonatomic gases, there are other modes of internal energy besides the translated one, and when the gas is perturbed, the translational energy adjusts to its equilibrium value quickly. Other modes relax to their equilibrium values through collision of gas particles. The time scale for such a relaxation process may not be short and the phenomenon of thermo-nonequilibrium becomes important. In this case, the compressible Euler equations should be supplemented by a rate equation governing the nonequilibrium mode of the internal energy. In the domain of kinetic equations, the relaxation parameter is represented by the dimensionless Knudsen number $\varepsilon > 0$, defined as the ratio of the mean free path of the particles over a typical lenght scale, such as the size of the spatial domain. It measures the rarefiedness of the gas. At the continuous level, it has been shown [21] that, for the Boltzmann equation

$$\partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon} Q(f)\,,$$

when the Knudsen number ε goes to zero, the distribution function f converges to a local Maxwellian $\mathcal{M}$, so the macroscopic model (compressible Euler or Navier-Stokes) becomes more adequate to describe the behavior. For more details about fluid dynamic limits of kinetic equation, we refer to [21, 22, 23, 36].

In this context, developing robust numerical schemes for kinetic equations that also work in the fluid regime becomes challenging. It has been done in the framework of

Asymptotic-Preserving (AP) schemes [92, 119]. With the vocabulary introduced in these papers, a numerical scheme is Asymptotic Preserving if

1. it is a suitable scheme for the relaxation problem (the kinetic equation), and when holding the mesh size and time step fixed, and letting ε go to zero, then the scheme becomes a suitable discretization of the limit (equilibrium) problem;

2. implicit collision terms can be implemented explicitly.

Indeed, from a numerical point of view, the treatment of the stiffness can not be done with explicit schemes, so mathematicians will rather favour the use of semi-implicit of fully implicit schemes.

One solution offered by E. Gabetta, L. Pareschi and G. Toscani [96] to design an Asymptotic Preserving scheme, was to penalize the nonlinear collision operator $Q(f)$ by a linear function λf and then absorb the linearly stiff part into the time variable to remove the stiffness. More recently, F. Filbet & S. Jin [92] proposed to penalize the Boltzman operator by the BGK operator in order to build stable schemes with respect to $\varepsilon > 0$. If such schemes are now numerically validated and extensively used to discretize kinetic equations [81, 92, 93, 96, 119], their mathematical study has only been done in some particular cases [103], because of the complexity of general kinetic equations, where the collision operator is nonlinear. However, hyperbolic conservation laws represents a simplified context where a theoretical study of relaxation schemes can be done [8, 61, 62, 63, 92, 119, 120]. Indeed, there is a strong analogy between the local relaxation approximation of conservation laws, initially proposed in [119], and the study of fluid dynamical limits of kinetic equations [36]. In the domain of hyperbolic conservation laws with stiff source terms, the relaxation parameter plays the role of the Knudsen number in kinetic theory, while the source term is the analogous of the collision operator, which we want to be the most general possible. Few works are devoted to the mathematical analysis of relaxation scheme for the approximation of conservations laws. We refer for instance to the series of papers by D. Aregba-Driollet and R. Natalini [8] and A. Chalabi [61, 62, 63]. But for all of them, the relaxation operator is relatively simple and can be easily treated explicitly without any additional computational cost. Here we want to focus on the approximation of general relaxation system with a nonlinear and stiff source term as in the context of the Boltzmann kinetic equation. Therefore, we will consider throughout the rest of the chapter the following hyperbolic relaxation system. For all (t, x) in $\mathbb{R}^+ \times \mathbb{R}$:

$$\begin{cases} \partial_t u^\varepsilon + \partial_x v^\varepsilon = 0\,, \\ \\ \partial_t v^\varepsilon + a\,\partial_x u^\varepsilon = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u^\varepsilon, v^\varepsilon), \end{cases} \tag{4.1.1}$$

where $a > 0$ is a constant coefficient to be discussed later, ε is the relaxation parameter and $\mathcal{R} : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ is a nonlinear function. The system is completed with the initial conditions:

$$\begin{cases} u^\varepsilon(0, x) = u_0^\varepsilon(x)\,, \\ \\ v^\varepsilon(0, x) = v_0^\varepsilon(x)\,. \end{cases} \tag{4.1.2}$$

The system of equations (4.1.1)-(4.1.2) is often refered to as a two velocity kinetic equation. Here we will assume that the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ possesses a unique local equilibrium, restricted to the manifold $\{v = A(u)\}$ [(1)], that is,

$$\mathcal{R}(u,v) = 0 \;\Leftrightarrow\; v = A(u)\,, \tag{4.1.3}$$

where A is a locally Lipschitz continuous function with $A(0) = 0$. Therefore, under some assumptions on $\mathcal{R}$ and on the initial data, the solution $(u^\varepsilon, v^\varepsilon)$ to (4.1.1)-(4.1.2) converges to (u,v) with $v = A(u)$ and u solution to the conservation laws [65, 142]

$$\begin{cases} \partial_t u + \partial_x A(u) = 0\,, \text{ in } \mathbb{R}^+\times\mathbb{R}, \\ \\ u(t=0) = u_0\,, \end{cases} \tag{4.1.4}$$

where the initial datum u_0 is given by

$$u_0 \;=\; \lim_{\varepsilon\to 0} u_0^\varepsilon\,. \tag{4.1.5}$$

Concerning the mathematical study of the system (4.1.1)-(4.1.2), we refer for instance to [152] [(2)].

Theorem 1.1. *Assume that the initial datum $(u_0^\varepsilon, v_0^\varepsilon)$ is bounded independently of ε in $BV(\mathbb{R})$. Consider $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$, which satisfies (4.1.3) and take the characteristic speed $\sqrt{a} > 0$ large enough. Then there exists a unique globally solution $(u^\varepsilon, v^\varepsilon)$ to the system (4.1.1)-(4.1.2) in $\mathcal{C}\left([0,\infty[, L^1_{loc}(\mathbb{R})^2\right)$ and there exists a constant $C > 0$ which only depends on a and $(u_0^\varepsilon, v_0^\varepsilon)$, such that for any $\varepsilon > 0$:*

$$\begin{cases} \left\| v^\varepsilon(t) \pm \sqrt{a}\, u^\varepsilon(t)\right\|_{L^\infty} \;\le\; C \quad \forall\, t > 0, \\ \\ TV\left(u^\varepsilon(t)\right) \,+\, TV\left(v^\varepsilon(t)\right) \;\le\; C \quad \forall\, t > 0, \\ \\ \|u^\varepsilon(t+\tau) - u^\varepsilon(t)\|_{L^1} \;\le\; C\,\tau, \quad \forall t \in \mathbb{R}^+,\, \tau \in \mathbb{R}^+, \\ \\ \|v^\varepsilon(t+\tau) - v^\varepsilon(t)\|_{L^1} \;\le\; \dfrac{C}{\varepsilon}\,\tau, \quad \forall t \in \mathbb{R}^+,\, \tau \in \mathbb{R}^+, \\ \\ \|v^\varepsilon(t+\tau) - v^\varepsilon(t)\|_{L^1} \;\le\; C_\nu\,\tau, \quad \forall t \;\ge\; \nu,\, \tau \in \mathbb{R}^+, \end{cases} \tag{4.1.6}$$

where $\nu > 0$, and C_ν only depends on a, $(u_0^\varepsilon, v_0^\varepsilon)$ and ν. Moreover, there exists $\beta_0 > 0$ such that,

$$\|v^\varepsilon(t,\cdot) - A\left(u^\varepsilon(t,\cdot)\right)\|_{L^1} \;\le\; e^{-\frac{\beta_0 t}{\varepsilon}}\, \|v_0^\varepsilon - A\left(u_0^\varepsilon\right)\|_{L^1} \,+\, C\,\varepsilon. \tag{4.1.7}$$

[1] This is nothing but the Implicit function Theorem.

[2] The study in [152] is done for the relaxation of Jin-Xin, the arguments are adapted to our case in Annex B, where we prove Theorem 1.1.

Finally, if u_0^ε converges, as ε goes to zero, to u_0 defined by (4.1.5), then the sequence $(u^\varepsilon, v^\varepsilon)$ converges to $(u, A(u))$ when ε goes to 0, such that, for any $\nu > 0$:

$$\left\{ \begin{array}{l} u^\varepsilon \longrightarrow u \quad \mathcal{C}\left([0,\infty); L^1_{loc}(\mathbb{R})\right), \\ \\ v^\varepsilon \longrightarrow A(u) \quad \mathcal{C}\left([\nu,\infty); L^1_{loc}(\mathbb{R})\right), \end{array} \right. \tag{4.1.8}$$

where u is the unique entropic solution to the Cauchy problem (4.1.4).

In this work we propose a rigorous analysis of the Asymptotic Preserving scheme proposed by F. Filbet & S. Jin [92] for a nonlinear relaxation. In other words, denoting by $\mathcal{P}^\varepsilon$ and $\mathcal{P}^0$ respectively the relaxation and the equilibrium Cauchy problems, and $\mathcal{P}_h^\varepsilon$ and $\mathcal{P}_h^0$ the corresponding discrete problems, where h represents the discretization parameter, independent of ε, we will perform a precise analysis of the following Asymptotic Preserving diagram.

$$\begin{array}{ccc} \mathcal{P}_h^\varepsilon & \xrightarrow{\quad\varepsilon\to 0\quad} & \mathcal{P}_h^0 \\ h \downarrow 0 \Big\downarrow & & \Big\downarrow h \downarrow 0 \\ \mathcal{P}^\varepsilon & \xrightarrow[\quad\varepsilon\to 0\quad]{} & \mathcal{P}^0 \end{array}$$

The chapter is organized as follows. We present in Section 4.2 an asymptotic preserving scheme for the relaxation model, and state the both convergence results of the Asymptotic Preserving scheme when the relaxation parameter ε goes to zero (Theorem 2.3) and next when the discretization parameter h goes to zero (Theorem 2.4). Then, we prove different *a priori* estimates in L^∞ and BV on the numerical solution to the Asymptotic Preserving scheme in Section 4.3 in order to prove both the zero relaxation limit (Section 4.4) and the convergence of the scheme (Section 4.5). Finally, we present some numerical results in Section 4.6.

4.2 Numerical schemes and main results

When $\mathcal{R}(u,v) = v - A(u)$, where $A \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ is a given function, the necessary and sufficient stability condition is given by the so called subcharacteristic condition:

$$|A'(u)| < \sqrt{a}\,. \tag{4.2.1}$$

It means that the propagation speed of the equilibrium problem has to be bounded by the speeds of the relaxation system, which has to be dissipative. For more details about this case, we refer to [65, 142, 152]. Moreover, $\nabla\mathcal{R}$ is uniformly bounded with respect to v and locally bounded with respect to u such that, for any (u,v) in $[-U_0, U_0] \times \mathbb{R}$:

$$\left\{ \begin{array}{l} |\partial_u \mathcal{R}(u,v)| \leq g(U_0)\,, \\ \\ 0 < \beta_0(U_0) \leq \partial_v \mathcal{R}(u,v) \leq h(U_0)\,, \end{array} \right. \tag{4.2.2}$$

where β_0, g and h are some constants depending only on U_0.

Hence the subcharacteristic condition reads, in our case:

$$\left|\frac{\partial_u \mathcal{R}\,(u,v)}{\partial_v \mathcal{R}\,(u,v)}\right| \,<\, \sqrt{a}\,.$$

For the sequel, we define for any $N \,>\, 0$ and $\alpha > 0$,

$$\left\{\begin{array}{l} U(N,\alpha) := \left(1+\frac{1}{\sqrt{\alpha}}\right) N\,, \\ \\ F(N,\alpha) := \displaystyle\sup_{|\xi|\le U(N,\alpha)} |A(\xi)|\,, \\ \\ V(N,\alpha) := U(N,\alpha) \;+\; \frac{1}{\sqrt{\alpha}}\, F\,(N,\alpha)\;. \end{array}\right. \tag{4.2.3}$$

We also denote by $I(N,\alpha)$ the compact set

$$I(N,\alpha) \;:=\; \left[-\sqrt{a}\,V(N,\alpha),\, \sqrt{a}\,V(N,\alpha)\right]^2\,. \tag{4.2.4}$$

Moreover, we assume that the initial conditions $u_0^\varepsilon, v_0^\varepsilon$ are bounded independently of ε in $L^\infty(\mathbb{R})$, such that:

$$N_0 := \max\left\{\sup_{\varepsilon>0}\|u_0^\varepsilon\|_{L^\infty}, \sup_{\varepsilon>0}\|v_0^\varepsilon\|_{L^\infty}\right\} \,<\, \infty\,. \tag{4.2.5}$$

Consider any $a_0 > 0$ and assume that the function $\mathcal{R} \in {}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ satisfies (4.1.3) and (4.2.2). We choose the characteristic speed $\sqrt{a} > 0$ and the parameter $\beta > 0$ such that

$$\left\{\begin{array}{l} \sqrt{a} \;>\; \max\left\{\sqrt{a_0}\,,\; \dfrac{g\,(V(N_0,a_0))}{\beta_0\,(V(N_0,a_0))}\right\}\,, \\ \\ \beta \;=\; h\,(V(N_0,a_0))\,, \end{array}\right. \tag{4.2.6}$$

where V is given by (4.2.3).

Remark 2.1. Note that if we differentiate with respect to u the equilibrium equation

$$\mathcal{R}(u, A(u)) = 0\,,$$

we obtain

$$A'(u) = -\frac{\partial_u \mathcal{R}\,(u, A(u))}{\partial_v \mathcal{R}\,(u, A(u))} \tag{4.2.7}$$

and thus recover the well known subcharacteristic condition in the case of semi-linear relaxation, namely:

$$|A'(u)| \,<\, \sqrt{a}\,.$$

We present here the splitting Asymptotic Preserving scheme and its relaxed version. We introduce a space time discretization based on a uniform grid of points $(x_{j+1/2})_{j\in\mathbb{Z}} \subset \mathbb{R}$, with space step Δx, and a grid of points $t^n = n\,\Delta t$, for which the time step Δt is chosen such as the ratio $\Delta t/\Delta x$ is constant and satisfies the CFL condition:

$$0 < \lambda := \frac{\sqrt{a}\,\Delta t}{\Delta x} < 1\,. \tag{4.2.8}$$

We denote by $h = (\Delta t, \Delta x)$ the discretization parameter.

4.2.1 An Asymptotic Preserving scheme for the relaxation system

In this section, we design a numerical scheme for system (4.1.1)-(4.1.2), by introducing a splitting between the linear transport part, and the nonlinear relaxation part, for which we will take advantage of the knowledge of the equilibrium (4.1.3). When ε becomes small, the differential equation (4.1.1) becomes stiff and explicit schemes are subject to severe stability constraints. Of course, implicit schemes allow larger time steps, but new difficulty arises in seeking the numerical solution of a fully nonlinear problem at each time step. Here we want to combine both advantages of implicit and explicit schemes:large time step for stiff problems and low computational complexity of the numerical solution at each time step. This is done, as said in the introduction, in the spirit of Asymptotic Preserving schemes introduced by F. Filbet & S. Jin [92].

Thus we construct a numerical solution $(u_h^\varepsilon, v_h^\varepsilon)$ to (4.1.1)-(4.1.2) in $\mathbb{R}^+ \times \mathbb{R}$ as follows

$$\left\{\begin{array}{ll} u_h^\varepsilon(t,x) & = \displaystyle\sum_{n\in\mathbb{N}} \sum_{j\in\mathbb{Z}} u_j^n\, \mathbf{1}_{C_j}(x)\, \mathbf{1}_{[t^n,t^{n+1}[}(t)\,, \\ \\ v_h^\varepsilon(t,x) & = \displaystyle\sum_{n\in\mathbb{N}} \sum_{j\in\mathbb{Z}} v_j^n\, \mathbf{1}_{C_j}(x)\, \mathbf{1}_{[t^n,t^{n+1}[}(t)\,, \end{array}\right. \tag{4.2.9}$$

where $C_j =]x_{j-1/2}, x_{j+1/2}[$ are the space cells and the sequences $\left(u_j^n\right)_{(n,j)\in\mathbb{N}\times\mathbb{Z}}$ and $\left(v_j^n\right)_{(n,j)\in\mathbb{N}\times\mathbb{Z}}$ depend on ε and are given below.

First, initial data are computed, as usual in the finite volume framework, as the averaged values of (4.1.2) through each space cell: for all j in $\mathbb{Z}$,

$$\left\{\begin{array}{ll} u_j^0 & = \dfrac{1}{\Delta x}\displaystyle\int_{C_j} u_0^\varepsilon(x)\,\mathrm{d}x\,, \\ \\ v_j^0 & = \dfrac{1}{\Delta x}\displaystyle\int_{C_j} v_0^\varepsilon(x)\,\mathrm{d}x\,. \end{array}\right. \tag{4.2.10}$$

Therefore, in order to discretize the system (4.1.1)-(4.1.2), we apply a splitting strategy into a linear transport part and a stiff ordinary differential part as follows. The first part will solve, using an explicit finite volume scheme, the hyperbolic linear system

$$\left\{\begin{array}{l} \partial_t u + \partial_x v = 0\,, \\ \\ \partial_t v + a\,\partial_x u = 0\,, \end{array}\right. \tag{4.2.11}$$

and then the second part deals with the stiff ordinary differential equations

$$\begin{cases} \partial_t u = 0\,, \\ \\ \partial_t v = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u,v)\,. \end{cases} \tag{4.2.12}$$

We first approximate the linear transport part, that is, for a given $(u^n\,,\,v^n)$, we compute an approximate solution $\left(u^{n+1/2},\,v^{n+1/2}\right)$ of (4.2.11) at time t^{n+1} with a standard Finite Volume scheme, that is, for all $j\,\in\,\mathbb{Z}$,

$$\begin{cases} u_j^{n+1/2} \;=\; u_j^n \;-\; \Delta t\, D_h v_j^n\,, \\ \\ v_j^{n+1/2} \;=\; v_j^n \;-\; \Delta t\, a\, D_h u_j^n\,, \end{cases} \tag{4.2.13}$$

where $D_h v_j^n$ and $a\,D_h u_j^n$ are discrete derivatives with respect to x of v and u, given for instance by the Lax-Friedrichs fluxes, namely:

$$\begin{cases} D_h v_j^n \,=\, \dfrac{1}{2\,\Delta x}\,\left[\left(v_{j+1}^n - v_{j-1}^n\right) - \sqrt{a}\,\left(u_{j+1}^n - 2\,u_j^n + u_{j-1}^n\right)\right]\,, \\ \\ a\,D_h u_j^n \,=\, \dfrac{1}{2\,\Delta x}\,\left[a\,\left(u_{j+1}^n - u_{j-1}^n\right) - \sqrt{a}\,\left(v_{j+1}^n - 2\,v_j^n + v_{j-1}^n\right)\right]\,. \end{cases}$$

Remark 2.2. Of course, there is a wide range of possible choices for the numerical fluxes. As we will see below, the main property of the numerical scheme for the linear transport term that we require is the TVD (*Total Variation Diminishing*) property, namely, for all $n\,\in\,\mathbb{N}$,

$$\begin{cases} TV(u^{n+1/2}) \,\leq\, TV(u^n)\,, \\ \\ TV(v^{n+1/2}) \,\leq\, TV(v^n)\,, \end{cases}$$

where $TV(u) \,:=\, \displaystyle\sum_{j\,\in\,\mathbb{Z}} |u_{j+1} - u_j|$.

Hence, the second part of the splitting only consists in approximating the nonlinear ordinary differential equation (4.2.12), for all $j\,\in\,\mathbb{Z}$, starting from $(u_j^{n+1/2}, v_j^{n+1/2})$. We use the decomposition

$$\mathcal{R}(u,v) = \left[\mathcal{R}(u,v) - \beta\,\left(v - A(u)\right)\right] \,+\, \beta\,\left(v - A(u)\right)\,,$$

where $\beta > 0$ is a parameter such that

$$0 \,<\, \sup_{(u,v)} \partial_v \mathcal{R}(u,v) \;<\; \beta.$$

Then, we apply a time exponential scheme on the dissipative part and get the following numerical scheme:

$$\left\{\begin{array}{rcl} u_j^{n+1} = & u_j^{n+1/2}\,, & \\ v_j^{n+1} = & v_j^{n+1/2} - \left(v_j^{n+1/2} - A(u_j^{n+1/2})\right)\left[1 - \left(1 + \dfrac{\beta\,\Delta t}{\varepsilon}\right) e^{-\beta\,\Delta t/\varepsilon}\right] & \\ & -\dfrac{\Delta t}{\varepsilon}\, e^{-\beta\,\Delta t/\varepsilon}\,\mathcal{R}\left(u_j^{n+1/2}, v_j^{n+1/2}\right)\,. & \end{array}\right. \tag{4.2.14}$$

4.2.2 Convergence results

We first establish a convergence result on the asymptotic behavior of the numerical solution to (4.2.13)-(4.2.14) when ε tends to zero.

Theorem 2.3. *Assume that the initial conditions $u_0^\varepsilon, v_0^\varepsilon$ are bounded independently of ε in $BV(\mathbb{R})$ and such that the assumption (4.2.5) is satisfied. Consider $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$, which satisfies (4.1.3)-(4.2.2) and the characteristic speed $\sqrt{a} > 0$ and the parameter $\beta > 0$ are given by (4.2.6). Then, the solution $(u_h^\varepsilon, v_h^\varepsilon)$ given by (4.2.9) to the scheme (4.2.13)-(4.2.14) with the initial data (4.2.10), converges in $L^1(\mathbb{R})$, as $\varepsilon \to 0$, to a numerical solution (u_h, v_h), that is,*

$$\|u_h^\varepsilon(t) - u_h(t)\|_{L^1} + \|v_h^\varepsilon(t) - v_h(t)\|_{L^1} \;\le\; C_t\, e^{-\beta_0 \Delta t/\varepsilon}\,\left[1 + \left\|\delta^0\right\|_{L^1}\right],$$

where (u_h, v_h) is a consistent approximation to the conservation laws (4.1.4) with $v_h = A(u_h)$,

$$u_h(t,x) := \sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} u_j^n\,\mathbf{1}_{C_j}(x)\,\mathbf{1}_{[t^n,t^{n+1}[}(t),$$

and

$$u_j^{n+1} = u_j^n \;+\; \Delta t\, D_h A(u_j^n), \qquad j\in\mathbb{Z},\; n\ge 1,$$

with the initial data

$$u_j^0 \;\; = \frac{1}{\Delta x}\int_{C_j} u_0(x)\,dx\,, \tag{4.2.15}$$

where u_0 is given by (4.1.5).

The convergence of the Asymptotic Preserving scheme is given by the following theorem.

Theorem 2.4. *Assume that the initial conditions $u_0^\varepsilon, v_0^\varepsilon$ are bounded independently of ε in $BV(\mathbb{R})$ and such that the assumption (4.2.5) is satisfied. Consider $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$, which satisfies (4.1.3)-(4.2.2) and the characteristic speed $\sqrt{a} > 0$ and the parameter $\beta > 0$ are given by (4.2.6). Then, the solution $(u_h^\varepsilon, v_h^\varepsilon)$ given by (4.2.9) to the scheme (4.2.13)-(4.2.14) and the initial data (4.2.10), converges in $L^1_{loc}(\mathbb{R}^+\times\mathbb{R})$, as $h \to (0,0)$, to a weak solution $(u^\varepsilon, v^\varepsilon)$ to the relaxation Cauchy problem (4.1.1)-(4.1.2).*

4.3 *A priori* estimates

In this section, we give the precise definition of the parameter β and the assumption on the characteristic speed to ensure the stability of the scheme (4.2.13)-(4.2.14) and prove estimates on the solution to the relaxation problem which are uniform with respect to ε. In following section, we drop the subscripts ε for sake of clarity and investigate the stability property of the Asymptotic Preserving scheme (4.2.13)-(4.2.14).

4.3.1 *A priori* estimate on the relaxation operator

In this section we focus on the second part of the scheme devoted to the approximation of the relaxation source term and give a technical lemma, which establishes a quasi-monotinicity property on the operator $G_{\varepsilon,s}$. In order to do this, we will rather consider the equivalent formulation on the diagonal variables w and z. Let us rewrite the splitting scheme on these variables. For given u and v,

$$\left\{ \begin{array}{l} w \,=\, -v \,-\, \sqrt{a}\, u\,, \\ \\ z \,=\, +v \,-\, \sqrt{a}\, u\,. \end{array} \right. \tag{4.3.1}$$

Therefore, the linear transport scheme (4.2.13) written for (w, z) only becomes the upwind method:for all $j \in \mathbb{Z}$,

$$\left\{ \begin{array}{rcl} w_j^{n+1/2} & = & w_j^n \,-\, \sqrt{a}\, \frac{\Delta t}{\Delta x} \left(w_j^n - w_{j-1}^n \right), \\ \\ z_j^{n+1/2} & = & z_j^n \,+\, \sqrt{a}\, \frac{\Delta t}{\Delta x} \left(z_{j+1}^n - z_j^n \right), \end{array} \right. \tag{4.3.2}$$

whereas the nonlinear stiff part (4.2.14) yields, for all $j \in \mathbb{Z}$

$$\left\{ \begin{array}{rcl} w_j^{n+1} & = & w_j^{n+1/2} \,+\, G_{\varepsilon,\Delta t} \left(w_j^{n+1/2}, z_j^{n+1/2} \right), \\ \\ z_j^{n+1} & = & z_j^{n+1/2} \,-\, G_{\varepsilon,\Delta t} \left(w_j^{n+1/2}, z_j^{n+1/2} \right), \end{array} \right. \tag{4.3.3}$$

with

$$\begin{array}{rcl} G_{\varepsilon,\Delta t}(w,z) & = & \left(\dfrac{z-w}{2} - A\left(-\dfrac{w+z}{2\sqrt{a}} \right) \right) \left[1 - \left(1 + \dfrac{\beta\, \Delta t}{\varepsilon} \right) e^{-\beta\, \Delta t/\varepsilon} \right] \\ \\ & + & \dfrac{\Delta t}{\varepsilon}\, e^{-\beta\, \Delta t/\varepsilon}\, \mathcal{R}\left(-\dfrac{w+z}{2\sqrt{a}}, \dfrac{z-w}{2} \right). \end{array}$$

The main result of this section, from which will be derived L^∞ an BV estimates, follows.

Lemma 3.1. *Assume the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R}, \mathbb{R})$ satisfies (4.1.3)-(4.2.2) and choose $a > 0$, $\beta > 0$ such that (4.2.6) is verified. Then,*

(*i*) *the subcharacteristic condition is satisfied for all* $(w,z) \in I(N_0,a_0)$*, that is,*

$$\left|\frac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}}(u,v)\right| < \sqrt{a}, \tag{4.3.4}$$

where $2\sqrt{a}\,u := -(w+z)$ *and* $2\,v := z-w$.

(*ii*) *for all* $\varepsilon\,,\ s>0$*, the source term operator* $G_{\varepsilon,s}$ *is quasimonotone on the compact set* $I(N_0,a_0)$*, that is,*

$$\begin{cases} -1 \le \partial_w G_{\varepsilon,\,s}(w,z) \le 0, \ \forall\,(w,z) \in I(N_0,a_0), \\ \\ 0 \le \partial_z G_{\varepsilon,\,s}(w,z) \le 1, \ \forall\,(w,z) \in I(N_0,a_0); \end{cases} \tag{4.3.5}$$

(*iii*) *consider for* $i=1,2$*,* (w_i^{n+1}, z_i^{n+1}) *two solutions to (4.3.3) corresponding to two initial data* $(w_i^{n+1/2}, z_i^{n+1/2}) \in I(N_0,a_0)$*. Then there exist* w *and* $z \in \mathbb{R}$ *such that* $|w|,\ |z| \le \sqrt{a}\,V(N_0,a_0)$ *and*

$$\begin{cases} w_1^{n+1} - w_2^{n+1} &= \left(w_1^{n+1/2} - w_2^{n+1/2}\right)\left(1+\partial_w G_{\varepsilon,s}(w, z_1^{n+1/2})\right) \\ &+ \left(z_1^{n+1/2} - z_2^{n+1/2}\right)\partial_z G_{\varepsilon,s}(w_2^{n+1/2}, z)\,, \\ z_1^{n+1} - z_2^{n+1} &= \left(z_1^{n+1/2} - z_2^{n+1/2}\right)\left(1-\partial_z G_{\varepsilon,s}(w_2^{n+1/2}, z)\right) \\ &- \left(w_1^{n+1/2} - w_2^{n+1/2}\right)\partial_w G_{\varepsilon,s}(w, z_1^{n+1/2})\,. \end{cases} \tag{4.3.6}$$

Proof. For any $N_0>0$ and $a_0>0$, we first observe that for $(w,z) \in I(N_0,a_0)$,

$$|u| = \frac{|w+z|}{2\sqrt{a}} \le V(N_0,a_0).$$

Therefore, using the assumption (4.2.2) and the definition (4.2.3), we get that

$$\left|\frac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}}(u,v)\right| \le \frac{g\left(V(N_0,a_0)\right)}{\beta_0\left(V(N_0,a_0)\right)} < \sqrt{a}\,,$$

which proves the first assertion (*i*).

Now we prove (*ii*) the quasi-monotonicity property of $G_{\varepsilon,s}$. Computing the partial derivatives of $G_{\varepsilon,s}$, it yields for all $s>0$,

$$\begin{cases} \partial_w G_{\varepsilon,s} = -\dfrac{1}{2}\left(1-\dfrac{A'(u)}{\sqrt{a}}\right)\left[1-\left(1+\dfrac{\beta\,s}{\varepsilon}\right)e^{-\beta\,s/\varepsilon}\right] - \dfrac{s}{2\,\varepsilon}\,e^{-\beta\,s/\varepsilon}\left(\dfrac{\partial_u \mathcal{R}}{\sqrt{a}} + \partial_v \mathcal{R}\right), \\ \\ \partial_z G_{\varepsilon,s} = +\dfrac{1}{2}\left(1+\dfrac{A'(u)}{\sqrt{a}}\right)\left[1-\left(1+\dfrac{\beta\,s}{\varepsilon}\right)e^{-\beta\,s/\varepsilon}\right] + \dfrac{s}{2\,\varepsilon}\,e^{-\beta\,s/\varepsilon}\left(\dfrac{-\partial_u \mathcal{R}}{\sqrt{a}} + \partial_v \mathcal{R}\right). \end{cases}$$

Hence, from Remark 2.1 and the subcharacteristic condition (4.3.4), we obtain that for all $(u,v) \in I(N_0,a_0)$

$$\partial_w G_{\varepsilon,s}(w,z) \le 0 \quad \text{and} \quad \partial_z G_{a,s}(w,z) \ge 0\,.$$

Moreover, still using condition (4.3.4), we also get for all $(w,z) \in I(N_0,a_0)$

$$\begin{cases} \partial_w G_{\varepsilon,s}(w,z) \ge -\left[1-\dfrac{\beta s}{\varepsilon}\,e^{-\beta s/\varepsilon}\right] - \partial_v \mathcal{R}(u,v)\,\dfrac{s}{\varepsilon}\,e^{-\beta s/\varepsilon}\,, \\ \\ \partial_z G_{\varepsilon,s}(w,z) \le \left[1-\dfrac{\beta s}{\varepsilon}\,e^{-\beta s/\varepsilon}\right] + \partial_v \mathcal{R}(u,v)\,\dfrac{s}{\varepsilon}\,e^{-\beta s/\varepsilon}\,. \end{cases}$$

Now since $|u| \le V(N_0,a_0)$ and from the choice of the parameter β in (4.2.6), it yields $|\partial_v \mathcal{R}(u,v)| \le \beta$. Therefore, we conclude that

$$-1 \le \partial_w G_{\varepsilon,s}(w,z) \quad \text{and} \quad \partial_z G_{\varepsilon,s}(w,z) \le 1, \quad \forall (w,z) \in I(N_0,a_0).$$

Finally (*iii*) follows from a first order Taylor expansion of $G_{\varepsilon,s}$. □

This Lemma allows to obtain the following comparison principle.

Corollary 3.2. *Consider for $i=1,2$, two initial data $(w_i^{n+1/2}, z_i^{n+1/2}) \in I(N_0,a_0)$ satisfying the monotononicity condition*

$$w_1^{n+1/2} \le w_2^{n+1/2} \quad \text{and} \quad z_1^{n+1/2} \le z_2^{n+1/2}.$$

Then, the numerical solution (w_i^{n+1}, z_i^{n+1}), given by (4.3.3) corresponding to the inital data $(w_i^{n+1/2}, z_i^{n+1/2})$ for $i=1,2$, satisfies

$$w_1^{n+1} \le w_2^{n+1} \quad \text{and} \quad z_1^{n+1} \le z_2^{n+1}.$$

Proof. Starting from the equality (4.3.6), it yields to the result applying the estimates (4.3.5). □

4.3.2 L^∞ estimates

In this section, we establish a uniform bound on the numerical solution to the scheme (4.2.13)-(4.2.14), that is, equivalently the scheme (4.3.2)-(4.3.3), with the time-space step $h=(\Delta t, \Delta x)$ such that (4.2.8) is satisfied.

Proposition 3.3. *Consider any $a_0 > 0$ and*

$$N_0 = \max\left\{\sup_{\varepsilon>0}\|u_0\|_{L^\infty}, \sup_{\varepsilon>0}\|v_0\|_{L^\infty}\right\}.$$

We assume that the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ satisfies (4.1.3)-(4.2.2) and choose $a>0$, $\beta>0$ such that (4.2.6) is verified. Then, for all $n \in \mathbb{N}$

$$\|u^n\|_{L^\infty} \le V(N_0,a_0), \quad \|v^n\|_{L^\infty} \le \sqrt{a}\,V(N_0,a_0).$$

Proof. We will proceed in two steps :

- find a particular solution $(\overline{w}^n, \overline{z}^n) \in \mathbb{R}^2$ to the scheme (4.3.2)-(4.3.3) which is uniformly bounded,
- apply the comparison principle on the compact set $I(N_0, a_0)$ to prove an L^∞ bound on (u^n, v^n).

For $R_0 = (1 + \sqrt{a})\, N_0$, we consider the numerical solution $(\overline{w}^n, \overline{z}^n)$ to (4.3.2)-(4.3.3) with the particular initial data $(\overline{w}^0, \overline{z}^0) = (R_0, R_0)$, which does not depend on the space variable so that the transport step (4.3.2) is invariant. Then we apply the relaxation scheme (4.3.3), which yields

$$\overline{w}^n \,=\, -\overline{v}^n \,-\, \sqrt{a}\,\overline{u}^n, \quad \overline{z}^n \,=\, +\overline{v}^n \,-\, \sqrt{a}\,\overline{u}^n$$

where $(\overline{u}^n, \overline{v}^n)$ are only given by

$$\left\{ \begin{array}{rcl} \overline{u}^n & = & \overline{u}^0 = -\dfrac{R_0}{\sqrt{a}} = \left(1 + \dfrac{1}{\sqrt{a}}\right) N_0, \\ \\ \overline{v}^n & = & \left(1 + \dfrac{\beta\,\Delta t}{\varepsilon}\right) e^{-\beta \Delta t/\varepsilon}\, \overline{v}^{n-1} \,+\, \left(1 - \left(1 + \dfrac{\beta\,\Delta t}{\varepsilon}\right) e^{-\beta \Delta t/\varepsilon}\right) A\left(\overline{u}^0\right) \\ \\ & - & \dfrac{\Delta t}{\varepsilon}\, e^{-\beta \Delta t/\varepsilon}\, \mathcal{R}\left(\overline{u}^0, \overline{v}^{n-1}\right). \end{array} \right.$$

Then, we proceed by induction to show that:

$$\forall\, n \,\in\, \{0, \ldots, N\}, \quad (\overline{w}^n, \overline{z}^n) \in I(N_0, a_0)\,.$$

We assume that $(\overline{w}^{n-1}, \overline{z}^{n-1}) \in I(N_0, a_0)$, for some $n \geq 1$. Let us prove that $(\overline{w}^n, \overline{z}^n) \in I(N_0, a_0)$. On the one hand since $\overline{u}^n = \overline{u}^0$, it yields

$$\|\overline{u}^n\|_{L^\infty} \,=\, \|\overline{u}^0\|_{L^\infty} \,\leq\, \left[1 \,+\, \frac{1}{\sqrt{a}}\right] N_0 \,\leq\, U(N_0, a_0).$$

On the other hand, using a first order Taylor expansion of the source term $\mathcal{R}(\overline{u}^0, .)$, we get that there exists $\tilde{v}^{n-1} \in \mathbb{R}$ such that

$$\begin{array}{rcl} \overline{v}^n & = & \left(1 + \dfrac{\beta - \partial_v \mathcal{R}(\overline{u}^0, \tilde{v}^k)}{\varepsilon}\, \Delta t\right) e^{-\beta \Delta t/\varepsilon}\, \overline{v}^{n-1} \\ & + & \left(1 - \left(1 + \dfrac{\beta - \partial_v \mathcal{R}(\overline{u}^0, \tilde{v}^k)}{\varepsilon}\, \Delta t\right) e^{-\beta \Delta t/\varepsilon}\right) A\left(\overline{u}^0\right). \end{array}$$

Therefore, denoting by $\lambda_k \in \mathbb{R}$, the real number such that

$$\lambda_k := \left(1 + \frac{\beta - \partial_v \mathcal{R}(\overline{u}^0, \tilde{v}^k)}{\varepsilon}\, \Delta t\right) e^{-\beta \Delta t/\varepsilon}, \qquad \forall k \in \mathbb{N},$$

with $|\tilde{v}^k| \leq F(N_0, a_0)$, hence we get

$$\overline{v}^n = \lambda_{n-1}\, \overline{v}^{n-1} \ + \ (1 - \lambda_{n-1})\, A(\overline{u}^0)$$

and since $\overline{v}^0 = 0$

$$\overline{v}^n \ = \ \left(1 - \prod_{k=0}^{n-1} \lambda_k\right) A(\overline{u}^0).$$

Moreover, since $|\overline{u}^0| \leq U(N_0, a_0)$ and $\tilde{v}^k \leq \sqrt{a}\, V(N_0, a_0)$, for all $k \in \mathbb{N}$, we get from (4.2.2) and (4.2.6),

$$0 \ < \ \left(1 + \frac{\beta - \beta_0}{\varepsilon}\, \Delta t\right) e^{-\beta \Delta t/\varepsilon} \ \leq \ \lambda_k \ \leq \ \left(1 + \frac{\beta\, \Delta t}{\varepsilon}\right) e^{-\beta \Delta t/\varepsilon} \ < \ 1, \quad \forall k \in \mathbb{N}.$$

Therefore, $\|\overline{v}^n\|_{L^\infty} \leq F(N_0, a_0)$ and

$$\|\overline{w}^n\|_{L^\infty}\,, \ \|\overline{z}^n\|_{L^\infty} \ \leq \ F(N_0, a_0) \ + \ (1 + \sqrt{a})\, N_0 \leq \sqrt{a}\, V(N_0, a_0),$$

that is, $(\overline{w}^n, \overline{z}^n) \in I(N_0, a_0)$.

Moreover, starting from the following initial datum $(\underline{w}^0, \underline{z}^0) = (-R_0, -R_0)$, we construct another particular solution $(\underline{w}^n, \underline{z}^n) \in I(N_0, a_0)$ for all $n \in \{0, \ldots, N\}$.

Now, we apply the comparison principle of Corollary 3.2 to prove an L^∞ estimate for any initial data u^0, $v^0 \in L^\infty(\mathbb{R})$ given by (4.2.10). From the definition of N_0, we have:

$$\|u^0\|_{L^\infty}, \ \|v^0\|_{L^\infty} \ \leq \ N_0.$$

Then, we have for the initial data (w^0, z^0)

$$\|w^0\|_{L^\infty}\,, \ \|z^0\|_{L^\infty} \ \leq \ (1 + \sqrt{a})\, N_0 = R_0 \ \leq \sqrt{a}\, V(N_0, a_0).$$

In other words, we have initially:

$$\underline{w}^0 \ \leq \ w^0 \ \leq \ \overline{w}^0, \quad \underline{z}^0 \ \leq \ z^0 \ \leq \ \overline{z}^0\,.$$

Thus, we proceed by induction and assume that

$$\underline{w}^n \ \leq \ w^n \ \leq \ \overline{w}^n, \quad \underline{z}^n \ \leq \ z^n \ \leq \ \overline{z}^n\,.$$

We first apply the linear transport step (4.3.2) to (w^n, z^n) and get that

$$\underline{w}^n \ \leq \ w^{n+1/2} \ \leq \ \overline{w}^n, \quad \underline{z}^n \ \leq \ z^{n+1/2} \ \leq \ \overline{z}^n\,.$$

Then, by applying Corollary 3.2 to the two solutions to (4.3.3) associated to the initial conditions $(w_1^{n+1/2}, z_1^{n+1/2}) = (w^{n+1/2}, z^{n+1/2})$ and $(w_2^{n+1/2}, z_2^{n+1/2}) = (\underline{w}^n, \underline{z}^n)$ (resp. $(\overline{w}^n, \overline{z}^n)$), we have

$$\underline{w}^{n+1} \ \leq \ w^{n+1} \ \leq \ \overline{w}^{n+1}, \quad \underline{z}^{n+1} \ \leq \ z^{n+1} \ \leq \ \overline{z}^{n+1},$$

which finally gives for all $n \in \mathbb{N}$, that $(w^n, z^n) \in I(N_0, a_0)$. By construction of (u^n, v^n) we have proven that

$$\|u^n\|_{L^\infty} \ \leq \ V(N_0, a_0), \quad \|v^n\|_{L^\infty} \ \leq \ \sqrt{a}\, V(N_0, a_0).$$

□

4.3.3 BV estimates

In this section, we obtain a BV estimate on the numerical solution to the scheme (4.2.13)-(4.2.14), that is, equivalently the scheme (4.3.2)-(4.3.3), with the time-space step $h = (\Delta t, \Delta x)$ such that (4.2.8) is satisfied.

Proposition 3.4. *Assume that u_0, v_0 are uniformly bounded with respect to ε in $BV(\mathbb{R})$. For any $a_0 > 0$ and $N_0 = \max\left\{\sup_{\varepsilon>0}\|u_0\|_{L^\infty}, \sup_{\varepsilon>0}\|v_0\|_{L^\infty}\right\}$, we assume that the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ satisfies (4.1.3)-(4.2.2) and choose $a > 0$, $\beta > 0$ such that (4.2.6) is verified. Then, for all $n \in \mathbb{N}$, we have:*

$$TV(w^{n+1}) + TV(z^{n+1}) \leq TV(w^n) + TV(z^n).$$

Proof. First we note that u^0, $v^0 \in BV(\mathbb{R})$, then by construction, w^0, $z^0 \in BV(\mathbb{R})$ also. To prove the BV estimate, we proceed in two steps. On the one hand, using the TVD property of the upwind scheme, we get that

$$TV(w^{n+1/2}) \leq TV(w^n) \quad \text{and} \quad TV(z^{n+1/2}) \leq TV(z^n).$$

On the other hand, we apply the nonlinear relaxation step (4.3.3) and from Lemma 3.1 (*iii*) with $w_1^{n+1/2} = w^{n+1/2}(.)$, $z_1^{n+1/2} = z^{n+1/2}(.)$ and $w_2^{n+1/2} = w^{n+1/2}(.+\Delta x)$, $z_2^{n+1/2} = z^{n+1/2}(.+\Delta x)$, it yields for any $j \in \mathbb{Z}$,

$$|w_{j+1}^{n+1} - w_j^{n+1}| + |z_{j+1}^{n+1} - z_j^{n+1}| \leq |w_{j+1}^{n+1/2} - w_j^{n+1/2}| + |z_{j+1}^{n+1/2} - z_j^{n+1/2}|.$$

Summing over $j \in \mathbb{Z}$, we get that

$$\begin{aligned} TV(w^{n+1}) + TV(z^{n+1}) &\leq TV(w^{n+1/2}) + TV(z^{n+1/2}) \\ &\leq TV(w^n) + TV(z^n). \end{aligned}$$

□

4.4 Trend to equilibrium

In this section we first focus on the asymptotic behavior of the numerical solution to (4.2.13)-(4.2.14) when ε goes to zero or when times goes to infinity. Then, we prove that the numerical solution to (4.2.13)-(4.2.14) converges to a consistent approximation of the conservation laws (4.1.4) when ε goes to zero. It corresponds to the limit $\mathcal{P}_h^\varepsilon \to \mathcal{P}_h^0$, when $\varepsilon \to 0$.

4.4.1 Asymptotic behavior

In this subsection, we drop the subscripts ε for sake of clarity.

Proposition 4.1. *Assume that u_0, v_0 are uniformly bounded with respect to ε in $BV(\mathbb{R})$. For any $a_0 > 0$ and N_0 as before, we assume that the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ satisfies (4.1.3)-(4.2.2) and choose $a > 0$, $\beta > 0$ such that (4.2.6) is verified. Then the (discrete) deviation from the equilibrium, $\delta = v - A(u)$ is controled as follows. For all $n \in \mathbb{N}$ and all $\varepsilon > 0$*

$$\left\{ \begin{array}{l} \left\|\delta^{n+1/2}\right\|_{L^1} \leq \left\|\delta^n\right\|_{L^1} + C\,\Delta t, \\ \\ \left\|\delta^n\right\|_{L^1} \leq e^{-\beta_0\, t^n/\varepsilon}\, \left\|\delta^0\right\|_{L^1} + C\,\varepsilon. \end{array} \right. \tag{4.4.1}$$

where $C > 0$ is a constant only depending on the parameters a, β_0 and the BV norm of the initial data.

Moreover, if $\varepsilon < \Delta t$ then we get

$$\left\|\delta^n\right\|_{L^1} \leq e^{-\beta_0\, t^n/\varepsilon}\, \left\|\delta^0\right\|_{L^1} + C_a\,\Delta t\, e^{-\beta_0 \Delta t/\varepsilon}. \tag{4.4.2}$$

Proof. We set, for $j \in \mathbb{Z}$, $n \in \mathbb{N}$ the sequence of the deviations from the equilibrium:

$$\delta_j^n = v_j^n - A\left(u_j^n\right).$$

We first consider the transport step (4.2.13) of the numerical scheme: for all $j \in \mathbb{Z}$, there exists ξ_j^n such that $\left|\xi_j^n\right| \leq V(N_0, a_0)$ and

$$\begin{array}{rcl} \delta_j^{n+1/2} & = & \delta_j^n - \dfrac{\Delta t}{2\,\Delta x}\left[a\left(u_{j+1}^n - u_{j-1}^n\right) - \sqrt{a}\left(v_{j+1}^n - 2\,v_j^n + v_{j-1}^n\right)\right] \\ & - & \dfrac{\Delta t}{2\,\Delta x}\, A'(\xi_j^n)\left[\left(v_{j+1}^n - v_{j-1}^n\right) - \sqrt{a}\left(u_{j+1}^n - 2\,u_j^n + u_{j-1}^n\right)\right]. \end{array}$$

Thanks to the uniform BV estimate, proven in Proposition 3.4, the subcharacteristic condition

$$\left|A'(\xi_j^n)\right| < \sqrt{a},$$

and the TVD property of the numerical fluxes we get the first estimate (4.4.1), by multiplying by Δx and summing over $j \in \mathbb{Z}$:

$$\|\delta^{n+1/2}\|_{L^1} \leq \|\delta^n\|_{L^1} + \Delta t\, C_a\left[TV(v^0) + \sqrt{a}\, TV(u^0)\right], \tag{4.4.3}$$

where $C_a > 0$ is a constant only depending on a.

Then, we consider the second step of the numerical scheme (4.2.14). On the one hand, since $u^{n+1} = u^{n+1/2}$, it yields

$$\delta_j^{n+1} = \delta_j^{n+1/2}\left[1 + \beta\,\frac{\Delta t}{\varepsilon}\right] e^{-\beta\,\Delta t/\varepsilon} - \frac{\Delta t}{\varepsilon}\, e^{-\beta\,\Delta t/\varepsilon}\,\mathcal{R}(u_j^{n+1/2}, v_j^{n+1/2}).$$

On the other hand, applying a Taylor expansion, we get that there exists η such that $|\eta| \leq \sqrt{a}\, V(N_0, a_0)$ and:

$$\mathcal{R}(u_j^{n+1/2}, v_j^{n+1/2}) = \partial_v \mathcal{R}(u_j^{n+1/2}, \eta)\,\delta_j^{n+1/2}.$$

Hence, we have

$$\delta_j^{n+1} \,=\, \delta_j^{n+1/2} \left[1 \,+\, \left(1 - \frac{\partial_v \mathcal{R}(u_j^{n+1/2}, \eta)}{\beta}\right) \frac{\beta \Delta t}{\varepsilon}\right] e^{-\beta\, \Delta t/\varepsilon}.$$

Therefore under the assumption (4.2.2), we set for all $s \geq 0$

$$g(s) \,=\, \left[1 + \left(1 - \frac{\beta_0}{\beta}\right) s\right] e^{-s},$$

for which we easily show that for all $s \in \mathbb{R}^+$, we have that $e^{-s} \,\leq\, g\,(s) \,\leq\, e^{-\beta_0\, s/\beta}$. Hence, for $s = \Delta t/\varepsilon$

$$\left\|\delta^{n+1}\right\|_{L^1} \,\leq\, e^{-\beta_0\, \Delta t/\varepsilon} \left\|\delta^{n+1/2}\right\|_{L^1}. \tag{4.4.4}$$

Finally, gathering (4.4.3) and (4.4.4), we obtain that there exists a constant $C_1 > 0$ depending only on a, $TV(u^0)$ and $TV(v^0)$ such that

$$\left\|\delta^{n+1}\right\|_{L^1} \,\leq\, e^{-\beta_0\, \Delta t/\varepsilon} \left[\, \|\delta^n\|_{L^1} \,+\, C_1\, \Delta t\,\right].$$

By induction, we easily get

$$\|\delta^n\|_{L^1} \,\leq\, e^{-\beta_0\, t^n/\varepsilon} \left\|\delta^0\right\|_{L^1} \,+\, C_a\, \Delta t\, \frac{e^{-\beta_0\, \Delta t/\varepsilon}}{1 - e^{-\beta_0\, \Delta t/\varepsilon}}. \tag{4.4.5}$$

To conclude we only observe that $x\, e^{-x} \,\leq\, 1 - e^{-x}$, for any $x \,\geq\, 0$. This plugged into (4.4.5), it gives the second estimate of (4.4.1), that is, there exists a constant $C > 0$, only depending on a, β_0, $TV(u^0)$ and $TV(v^0)$ such that

$$\|\delta^n\|_{L^1} \,\leq\, e^{-\beta_0\, t^n/\varepsilon} \left\|\delta^0\right\|_{L^1} \,+\, C\, \varepsilon.$$

Moreover, when $\varepsilon < \Delta t$, we again start from the estimate (4.4.5) and note that $1/(1 - e^{-\beta_0 \Delta t/\varepsilon}) \leq 1/(1 - e^{-\beta_0})$. Thus, there exists another constant $C > 0$, only depending on a, β_0, $TV(u^0)$ and $TV(v^0)$ such that

$$\|\delta^n\|_{L^1} \,\leq\, e^{-\beta_0\, t^n/\varepsilon} \left\|\delta^0\right\|_{L^1} \,+\, C\, \Delta t\, e^{-\beta_0 \Delta t/\varepsilon},$$

which gives (4.4.2). □

4.4.2 Proof of Theorem 2.3

We are now ready to perform a rigorous asymptotic analysis of the numerical scheme (4.2.13)-(4.2.14) when ε goes to zero.

Let us consider the numerical solution $(u_h^\varepsilon, v_h^\varepsilon)$ to the scheme (4.2.13)-(4.2.14) written in the form (4.3.2)-(4.3.3) with

$$\left\{\begin{array}{l} w_h^\varepsilon \,=\, -v_h^\varepsilon \,-\, \sqrt{a}\, u_h^\varepsilon\,, \\[2ex] z_h^\varepsilon \,=\, +v_h^\varepsilon \,-\, \sqrt{a}\, u_h^\varepsilon\,, \end{array}\right.$$

such that

$$\begin{cases} w_h^\varepsilon(t,x) = \displaystyle\sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} w_j^{\varepsilon,n}\,\mathbf{1}_{C_j}(x)\,\mathbf{1}_{[t^n,t^{n+1}[}(t), \\ \\ z_h^\varepsilon(t,x) = \displaystyle\sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} z_j^{\varepsilon,n}\,\mathbf{1}_{C_j}(x)\,\mathbf{1}_{[t^n,t^{n+1}[}(t). \end{cases}$$

Let us also define (w_h, z_h) the numerical solution to the scheme (4.3.2)-(4.3.3) in the asymptotic limit $\varepsilon = 0$. w_h and z_h are given by

$$\begin{cases} w_h(t,x) = \displaystyle\sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} w_j^{n}\,\mathbf{1}_{C_j}(x)\,\mathbf{1}_{[t^n,t^{n+1}[}(t), \\ \\ z_h(t,x) = \displaystyle\sum_{n\in\mathbb{N}}\sum_{j\in\mathbb{Z}} z_j^{n}\,\mathbf{1}_{C_j}(x)\,\mathbf{1}_{[t^n,t^{n+1}[}(t). \end{cases}$$

Therefore, the values $(w_j^n, z_j^n)_{(n,j)\in\mathbb{N}\times\mathbb{Z}}$ are given by

$$\begin{cases} w_j^{n+1/2} \;=\; w_j^n \;-\sqrt{a}\,\dfrac{\Delta t}{\Delta x}\,\left(w_j^n - w_{j-1}^n\right), \\ \\ z_j^{n+1/2} \;=\; z_j^n \;+\sqrt{a}\,\dfrac{\Delta t}{\Delta x}\,\left(z_{j+1}^n - z_j^n\right), \end{cases}$$

and then

$$\begin{cases} w_j^{n+1} &=& w_j^{n+1/2} \,+\, G_{\varepsilon,\Delta t}\left(w_j^{n+1/2}, z_j^{n+1/2}\right) \,-\, \Delta t\,\mathcal{E}_j^n(\varepsilon), \\ \\ z_j^{n+1} &=& z_j^{n+1/2} \,-\, G_{\varepsilon,\Delta t}\left(w_j^{n+1/2}, z_j^{n+1/2}\right) \,+\, \Delta t\,\mathcal{E}_j^n(\varepsilon). \end{cases}$$

where $\Delta t\,\mathcal{E}_j^n(\varepsilon)$ represents the consistency error of the operator $G_{\varepsilon,\Delta t}$ with respect to ε, that is,

$$\Delta t\,\mathcal{E}_j^n(\varepsilon) := G_{\varepsilon,\Delta t}\left(w_j^{n+1/2}, z_j^{n+1/2}\right) - G_{0,\Delta t}\left(w_j^{n+1/2}, z_j^{n+1/2}\right).$$

Therefore, we apply Lemma 3.1 (ii) and (iii), with $(w_1, z_1) = (w_j^\varepsilon, z_j^\varepsilon)$ and $(w_2, z_2) = (w_j, z_j)$, it yields

$$\begin{aligned} |w_j^{\varepsilon,n+1} - w_j^{n+1}| \,+\, |z_j^{\varepsilon,n+1} - z_j^{n+1}| \;&\leq\; |w_j^{\varepsilon,n+1/2} - w_j^{n+1/2}| \,+\, |z_j^{\varepsilon,n+1/2} - z_j^{n+1/2}| \\ &+\; 2\,\left|\Delta t\,\mathcal{E}_j^n(\varepsilon)\right|, \end{aligned}$$

and by linearity of the transport scheme (4.3.2), we have for all $n \geq 0$

$$|w_j^{\varepsilon,n+1} - w_j^{n+1}| \,+\, |z_j^{\varepsilon,n+1} - z_j^{n+1}| \leq |w_j^{\varepsilon,n} - w_j^{n}| \,+\, |z_j^{\varepsilon,n} - z_j^{n}| \,+\, 2\,\left|\Delta t\,\mathcal{E}_j^n(\varepsilon)\right|.$$

Thus, multiplying by Δx, summing over $j \in \mathbb{Z}$ and applying a straighforward induction, we get the stability result

$$\begin{aligned} \sum_{j\in\mathbb{Z}} \Delta x\,\left(|w_j^{\varepsilon,n} - w_j^n| \,+\, |z_j^{\varepsilon,n} - z_j^n|\right) \;&\leq\; \sum_{j\in\mathbb{Z}} \Delta x\,\left(|w_j^{\varepsilon,0} - w_j^0| \,+\, |z_j^{\varepsilon,0} - z_j^0|\right) \\ &+\; 2\sum_{k=0}^{n-1}\sum_{j\in\mathbb{Z}} \Delta t\,\Delta x\,\left|\mathcal{E}_j^k(\varepsilon)\right|. \end{aligned}$$

It now remains to evaluate the error $\mathcal{E}_j^n(\varepsilon)$. Using that for any $(u,v) \in I(N_0,a_0)$, the function $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}^+,\mathbb{R})$ such that $\beta_0 \leq \partial_v \mathcal{R}(u,v) \leq \beta$, then we have

$$\begin{aligned}\Delta t \left|\mathcal{E}_j^n(\varepsilon)\right| &= e^{-\beta\,\Delta t/\varepsilon}\left|-\left(v_j^{n+1/2}-A(u_j^{n+1/2})\right)\left(1+\frac{\beta\,\Delta t}{\varepsilon}\right)+\frac{\Delta t}{\varepsilon}\mathcal{R}\left(u_j^{n+1/2},v_j^{n+1/2}\right)\right|,\\ &\leq e^{-\beta_0\Delta t/\varepsilon}\left|v_j^{n+1/2}-A(u_j^{n+1/2})\right|.\end{aligned}$$

Thanks to the estimates (4.4.1) and (4.4.2) in Proposition 4.1 on the deviation applied to $v^{n+1/2}-A(u^{n+1/2})$ which is also valid in the asymptotic $\varepsilon \to 0$, it yields

$$\|v^{n+1/2}-A(u^{n+1/2})\|_{L^1} \leq \begin{cases} \left\|\delta^0\right\|_{L^1} + C\,\Delta t & \text{if } n=0,\\ C\,\Delta t & \text{if } n>0.\end{cases}$$

Then, we get for $k \geq 0$ and $\varepsilon \leq \Delta t$,

$$\sum_{j\in\mathbb{Z}} \Delta x\,\Delta t\left|\mathcal{E}_j^k(\varepsilon)\right| \leq \begin{cases} e^{-\beta_0\Delta t/\varepsilon}\left(\left\|\delta^0\right\|_{L^1} + C\,\Delta t\right) & \text{if } k=0,\\ C\,e^{-\beta_0\Delta t/\varepsilon}\,\Delta t & \text{if } k>0.\end{cases}$$

Hence summing over $0 \leq k \leq n$, it gives

$$\sum_{k=0}^{n}\sum_{j\in\mathbb{Z}} \Delta x\,\Delta t\left|\mathcal{E}_{\varepsilon,\Delta t}^k\right| \leq e^{-\beta_0\Delta t/\varepsilon}\left[\left\|\delta^0\right\|_{L^1} + C\,t^{n+1}\right].$$

Finally, we get the estimate

$$\begin{aligned}\|w_h^\varepsilon(t^n)-w_h(t^n)\|_{L^1}+\|z_h^\varepsilon(t^n)-z_h(t^n)\|_{L^1} &\leq \|w_h^\varepsilon(0)-w_h(0)\|_{L^1}+\|z_h^\varepsilon(0)-z_h(0)\|_{L^1}\\ &+ 2\,e^{-\beta_0\Delta t/\varepsilon}\left[\left\|\delta^0\right\|_{L^1} + C\,t^n\right]\end{aligned}$$

and the result follows $(u_h^\varepsilon, v_h^\varepsilon) \to (u_h, v_h)$, when ε goes to zero.

4.5 Proof of Theorem 2.4

In this section, we prove the convergence of the relaxation Asymptotic Preserving scheme stated in Theorem 2.4. More precisely, we will obtain the following error estimate between the solution $(u_h^\varepsilon, v_h^\varepsilon)$ to the scheme and the solution $(u^\varepsilon, v^\varepsilon)$ to the continuous problem.

Proposition 5.1. *Consider a discretization parameter $h=(\Delta t,\Delta x)$ satisfying the CFL condition (4.2.8). Assume that the initial conditions $u_0^\varepsilon, v_0^\varepsilon$ are bounded independently of ε in $BV(\mathbb{R})$ and such that the assumption (4.2.5) is satisfied. Consider $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$, which satisfies (4.1.3)-(4.2.2) and the characteristic speed $\sqrt{a}>0$ and the parameter $\beta>0$ are given by (4.2.6). Denote by $(u^\varepsilon, v^\varepsilon)$ the weak solution to the relaxation Cauchy problem*

(4.1.1)-(4.1.2); while $(u_h^\varepsilon, v_h^\varepsilon)$, given by (4.2.9), is the solution to the scheme (4.2.13)-(4.2.14) with initial data (4.2.10). Then it holds, for all $t, \varepsilon > 0$:

$$\int_{\mathbb{R}} |u_h^\varepsilon(t,x) - u^\varepsilon(t,x)| + |v_h^\varepsilon(t,x) - v^\varepsilon(t,x)| \, dx \leq \frac{C}{\varepsilon} \left(\Delta t \left(\frac{\left\| \delta^0 \right\|_{L^1}}{\varepsilon} + 1 \right) + \Delta x^{1/2} \right). \tag{4.5.1}$$

As in the stability analysis of the relaxation scheme, we will rather consider the diagonal variables w and z and drop the subscripts ε for sake of clarity when it is not necessary. In the first subsection we study the consistency error, while in the second we compute the error and prove estimate (4.5.1).

4.5.1 Consistency error

Consider (w, z) the exact solution to (4.1.1)-(4.1.2) with (4.3.1). Unfortunately, this solution is not smooth enough to study the consistency error, then we introduce a regularization (w_δ, z_δ) given by

$$\begin{cases} w_\delta(t,x) = w \star \rho_\delta(t,x), \\ \\ z_\delta(t,x) = z \star \rho_\delta(t,x), \end{cases}$$

where $\star$ denotes the convolution product with respect to $x \in \mathbb{R}$ and

$$\rho_\delta(x) = \frac{1}{\delta} \rho\left(\frac{x}{\delta}\right) \quad \text{and} \quad \rho \in \mathcal{C}_c^\infty(\mathbb{R}), \quad \rho \geq 0, \quad \int_{\mathbb{R}} \rho(z) \, \mathrm{d}z = 1.$$

Thus, the couple (w_δ, z_δ) is solution to

$$\begin{cases} \partial_t w_\delta + \sqrt{a} \, \partial_x w_\delta = +\dfrac{1}{\varepsilon} \mathcal{R}_\delta(u,v), \\ \\ \partial_t z_\delta - \sqrt{a} \, \partial_x z_\delta = -\dfrac{1}{\varepsilon} \mathcal{R}_\delta(u,v), \end{cases} \tag{4.5.2}$$

with $\mathcal{R}_\delta = \mathcal{R} \star \rho_\delta$ and (u, v) solution to (4.1.1)-(4.1.2). Therefore, the solution can be written as

$$\begin{cases} w_\delta(t^{n+1}, x) = w_\delta(t^n, x - \sqrt{a}\Delta t) + \dfrac{1}{\varepsilon} \displaystyle\int_0^{\Delta t} \mathcal{R}_\delta(u,v)(t^n + s, x - \sqrt{a}(\Delta t - s)) \, \mathrm{d}t, \\ \\ z_\delta(t^{n+1}, x) = z_\delta(t^n, x + \sqrt{a}\Delta t) - \dfrac{1}{\varepsilon} \displaystyle\int_0^{\Delta t} \mathcal{R}_\delta(u,v)(t^n + s, x + \sqrt{a}(\Delta t - s)) \, \mathrm{d}t. \end{cases} \tag{4.5.3}$$

Then we set

$$\tilde{w}_j^n = \frac{1}{\Delta x} \int_{C_j} w_\delta(t^n, x) \, \mathrm{d}x, \quad \tilde{z}_j^n = \frac{1}{\Delta x} \int_{C_j} z_\delta(t^n, x) \, \mathrm{d}x. \tag{4.5.4}$$

Integrating over $x \in C_j$ (4.5.3) and dividing by Δx, it yields

$$\left\{ \begin{array}{l} \tilde{w}_j^{n+1} = \tilde{w}_j^{n+1/2} \, + \, G_{\varepsilon,\Delta t}\left(\tilde{w}_j^{n+1/2}, \tilde{z}_j^{n+1/2}\right) \, + \, \Delta t\, \mathcal{E}_{1,j}^n \, + \, \Delta t\, \mathcal{E}_{2,j}^n, \\ \\ \tilde{z}_j^{n+1} = \tilde{z}_j^{n+1/2} \, - \, G_{\varepsilon,\Delta t}\left(\tilde{w}_j^{n+1/2}, \tilde{z}_j^{n+1/2}\right) \, + \, \Delta t\, \mathcal{E}_{3,j}^n \, + \, \Delta t\, \mathcal{E}_{4,j}^n, \end{array} \right. \tag{4.5.5}$$

with

$$\left\{ \begin{array}{l} \tilde{w}_j^{n+1/2} = \tilde{w}_j^n \, - \sqrt{a}\, \dfrac{\Delta t}{\Delta x}\, \left(\tilde{w}_j^n - \tilde{w}_{j-1}^n\right), \\ \\ \tilde{z}_j^{n+1/2} = \tilde{z}_j^n \, + \sqrt{a}\, \dfrac{\Delta t}{\Delta x}\, \left(\tilde{z}_{j+1}^n - \tilde{z}_j^n\right). \end{array} \right.$$

The consistency errors related to the transport operator $\mathcal{E}_{1,j}^n$, $\mathcal{E}_{3,j}^n$ are respectively defined by

$$\Delta t\, \mathcal{E}_{1,j}^n = \frac{\varepsilon_{1,j+1/2}^n - \varepsilon_{1,j-1/2}^n}{\Delta x}, \quad \Delta t\, \mathcal{E}_{3,j}^n = \frac{\varepsilon_{3,j+1/2}^n - \varepsilon_{3,j-1/2}^n}{\Delta x},$$

where $\varepsilon_{1,j+1/2}^n$ and $\varepsilon_{3,j+1/2}^n$ are the consistency errors of the numerical flux and are given by

$$\left\{ \begin{array}{l} \varepsilon_{1,j+1/2}^n = -\displaystyle\int_0^{\sqrt{a}\Delta t} w_\delta(t^n, x_{j+1/2} - s)\mathrm{d}s \, + \, \sqrt{a}\, \Delta t\, \tilde{w}_j^n, \\ \\ \varepsilon_{3,j+1/2}^n = +\displaystyle\int_0^{\sqrt{a}\Delta t} z_\delta(t^n, x_{j+1/2} + s)\mathrm{d}s \, - \, \sqrt{a}\, \Delta t\, \tilde{z}_{j+1}^n, \end{array} \right.$$

whereas the consistency errors $\Delta t \mathcal{E}_{2,j}^n$ and $\Delta t \mathcal{E}_{4,j}^n$ correspond to the stiff source term and are given by

$$\left\{ \begin{array}{l} \Delta t\, \mathcal{E}_{2,j}^n \, = \, +\dfrac{1}{\Delta x} \displaystyle\int_{C_j} \int_0^{\Delta t} \frac{1}{\varepsilon} \mathcal{R}_\delta\left(u, v\right)\left(t^n + s, x - \sqrt{a}(\Delta t - s)\right) \mathrm{d}s \, - \, G_{\varepsilon,\Delta t}\left(\tilde{w}_j^{n+1/2}, \tilde{z}_j^{n+1/2}\right) \mathrm{d}x, \\ \\ \Delta t\, \mathcal{E}_{4,j}^n \, = \, -\dfrac{1}{\Delta x} \displaystyle\int_{C_j} \int_0^{\Delta t} \frac{1}{\varepsilon} \mathcal{R}_\delta\left(u, v\right)\left(t^n + s, x + \sqrt{a}(\Delta t - s)\right) \mathrm{d}s \, - \, G_{\varepsilon,\Delta t}\left(\tilde{w}_j^{n+1/2}, \tilde{z}_j^{n+1/2}\right) \mathrm{d}x. \end{array} \right.$$

We then evaluate successively each consistency error term. On the one hand, we prove the following consistency error for smooth solutions, which is related to the transport approximation.

Proposition 5.2. *Let (w, z) be given by (4.3.1), where (u, v) is the exact solution to (4.1.1)-(4.1.2) and such that w, $z \in L^\infty(\mathbb{R}^+, BV(\mathbb{R}))$. Then the consistency error related to the transport part satisfies*

$$\sum_{j \in \mathbb{Z}} \Delta x \, \left[\, |\mathcal{E}_{1,j}^n| \, + \, |\mathcal{E}_{3,j}^n| \,\right] \, \leq \, C\, \frac{\Delta x}{\delta}\, \left(\, TV(w(t^n)) \, + \, TV(z(t^n))\,\right).$$

Proof. We first study the consistency error for $w \in L^\infty(\mathbb{R}^+, BV(\mathbb{R}))$. We perform a simple change of variable, which yields since $\sqrt{a}\Delta t = \lambda\,\Delta x$,

$$\begin{aligned}\varepsilon^n_{1,j+1/2} &= -\lambda \int_0^{\Delta x} w_\delta(t^n, x_{j+1/2} - \lambda s)\mathrm{d}s \,+\, \lambda \int_0^{\Delta x} w_\delta(t^n, x_{j+1/2} - s)\mathrm{d}s,\\ &= \lambda \int_0^{\Delta x}\int_{\lambda s}^{s} \partial_x w_\delta(t^n, x_{j+1/2} - r)\mathrm{d}r\,\mathrm{d}s.\end{aligned}$$

Therefore, since w_δ is smooth we have

$$\begin{aligned}\left|\mathcal{E}^n_{1,j}\right| &= \frac{\sqrt{a}}{\Delta x^2}\left|\int_0^{\Delta x}\int_{\lambda s}^{s} \partial_x w_\delta(t^n, x_{j+1/2} - r) \,-\, \partial_x w_\delta(t^n, x_{j-1/2} - r)\mathrm{d}r\,\mathrm{d}s\right|,\\ &\le \sqrt{a}\int_{x_{i-3/2}}^{x_{i+1/2}} \left|\partial^2_{xx} w_\delta(t^n, x)\right|\,\mathrm{d}x.\end{aligned}$$

By multiplying by Δx and summing over $j \in \mathbb{Z}$, we get an estimate for a smooth solution $w_\delta(t^n) \in W^{2,1}(\mathbb{R})$,

$$\sum_{j\in\mathbb{Z}} \Delta x\,\left|\mathcal{E}^n_{1,j}\right| \,\le\, 2\sqrt{a}\,\Delta x\,\|\partial^2_{xx} w_\delta(t^n)\|_{L^1}.$$

To achieve the proof, we need to estimate $\|\partial^2_{xx} w_\delta(t^n)\|_{L^1}$ with respect to w and ρ_δ. Using the convolution properties, we easily get

$$\|\partial^2_{xx} w_\delta(t^n)\|_{L^1} \,\le\, \frac{C}{\delta}\,\|\partial_x w_\delta(t^n)\|_{L^1} \,\le\, \frac{C}{\delta}\,TV(w(t^n)),$$

which allows to conclude that

$$\sum_{j\in\mathbb{Z}} \Delta x\,\left|\mathcal{E}^n_{1,j}\right| \,\le\, C\,\frac{\Delta x}{\delta}\,TV(w(t^n)).$$

Using a similar technique, we also get for a smooth solution $z \in L^\infty(\mathbb{R}^+, BV(\mathbb{R}))$,

$$\sum_{j\in\mathbb{Z}} \Delta x\,\left|\mathcal{E}^n_{3,j}\right| \,\le\, C\,\frac{\Delta x}{\delta}\,TV(z(t^n)).$$

□

On the other hand, we treat the consistency errors $\mathcal{E}^n_{2,j}$ and $\mathcal{E}^n_{4,j}$, which are related to the stiff source term.

Proposition 5.3. *Let (w, z) be given by (4.3.1), where (u, v) is the exact solution to (4.1.1)-(4.1.2). Assume that w, $z \in L^\infty(\mathbb{R}^+, BV(\mathbb{R}))$. Then there exists a constant $C > 0$, only depending on u and v such that the consistency error related to the stiff source part satisfies*

$$\sum_{j\in\mathbb{Z}} \Delta x\,\left|\mathcal{E}^n_{2,j}\right| \,\le\, C\,\left[\frac{\Delta t}{\varepsilon}\left(e^{-\beta_0\, t^n/\varepsilon}\,\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon} \,+\, 1\right) + \frac{\Delta x}{\varepsilon} \,+\, \frac{\delta}{\varepsilon}\right]$$

and

$$\sum_{j\in\mathbb{Z}} \Delta x \, \left|\mathcal{E}^n_{4,j}\right| \;\leq\; C\left[\frac{\Delta t}{\varepsilon}\left(e^{-\beta_0\, t^n/\varepsilon}\,\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon} \,+\, 1\right) \,+\, \frac{\Delta x}{\varepsilon} \,+\, \frac{\delta}{\varepsilon}\right].$$

Proof. We first define $(\tilde{u}^n_j, \tilde{v}^n_j)$ such that

$$\tilde{u}^n_j = -\frac{\tilde{w}^n_j + \tilde{z}^n_j}{2\sqrt{a}} \quad \text{and} \quad \tilde{v}^n_j = \frac{\tilde{z}^n_j - \tilde{w}^n_j}{2}.$$

Therefore, we split the consistency error $\mathcal{E}^n_{2,j}$ as

$$\mathcal{E}^n_{2,j} \;=\; \mathcal{E}^n_{21,j} \,+\, \mathcal{E}^n_{22,j} \,+\, \mathcal{E}^n_{23,j} \,+\, \mathcal{E}^n_{24,j} \,+\, \mathcal{E}^n_{25,j},$$

with

$$\left\{\begin{array}{l}
\Delta t\, \mathcal{E}^n_{21,j} \;=\; -\left[1-\left(1+\dfrac{\beta\,\Delta t}{\varepsilon}\right)\, e^{-\beta\,\Delta t/\varepsilon}\right]\,\left(\tilde{v}^{n+1/2}_j - A\left(\tilde{u}^{n+1/2}_j\right)\right),\\[2ex]
\Delta t\, \mathcal{E}^n_{22,j} \;=\; \left(1-e^{-\beta\,\Delta t/\varepsilon}\right)\,\dfrac{\Delta t}{\varepsilon}\,\mathcal{R}(\tilde{u}^{n+1/2}_j, \tilde{v}^{n+1/2}_j),\\[2ex]
\Delta t\, \mathcal{E}^n_{23,j} \;=\; \dfrac{1}{\varepsilon\,\Delta x}\displaystyle\int_{C_j}\int_0^{\Delta t} \mathcal{R}_\delta(u,v)(t^n+s, x-\sqrt{a}(\Delta t - s)) - \mathcal{R}_\delta(u,v)(t^n, x-\sqrt{a}(\Delta t))\mathrm{d}s\mathrm{d}x,\\[2ex]
\Delta t\, \mathcal{E}^n_{24,j} = \dfrac{\Delta t}{\varepsilon\,\Delta x}\displaystyle\int_{C_j} \mathcal{R}_\delta(u,v)(t^n, x-\sqrt{a}(\Delta t)) - \mathcal{R}(u,v)(t^n, x-\sqrt{a}(\Delta t))\mathrm{d}x,\\[2ex]
\Delta t\, \mathcal{E}^n_{25,j} = \dfrac{\Delta t}{\varepsilon\,\Delta x}\displaystyle\int_{C_j} \mathcal{R}(u,v)(t^n, x-\sqrt{a}(\Delta t)) - \mathcal{R}\left(\tilde{u}^{n+1/2}_j, \tilde{v}^{n+1/2}_j\right)\,\mathrm{d}x.
\end{array}\right.$$

On the one hand, the two terms $\mathcal{E}^n_{21,j}$ and $\mathcal{E}^n_{22,j}$ can be easily evaluated using a Taylor expansion of $s \mapsto e^{-\beta s/\varepsilon}$:it yields

$$\Delta t\,\left|\mathcal{E}^n_{21,j}\right| \;\leq\; \frac{1}{2}\left(\frac{\beta\Delta t}{\varepsilon}\right)^2 \left|\tilde{v}^{n+1/2}_j - A\left(\tilde{u}^{n+1/2}_j\right)\right|.$$

Using that $\mathcal{R}(u, A(u)) = 0$ and $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ with $\partial_v \mathcal{R}(u,v) \leq \beta$, we also obtain that

$$\Delta t\,\left|\mathcal{E}^n_{22,j}\right| \;\leq\; \left(\frac{\beta\Delta t}{\varepsilon}\right)^2 \left|\tilde{v}^{n+1/2}_j - A\left(\tilde{u}^{n+1/2}_j\right)\right|$$

Therefore, from (4.4.1) in Proposition 4.1, we have

$$\sum_{j\in\mathbb{Z}} \Delta x \,\left[\,\left|\mathcal{E}^n_{21,j}\right| \,+\, \left|\mathcal{E}^n_{22,j}\right|\,\right] \;\leq\; C\,\frac{\Delta t}{\varepsilon}\left(e^{-\beta_0\, t^n/\varepsilon}\,\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon} \,+\, 1\right). \tag{4.5.6}$$

On the other hand, we proceed to the evaluation of the terms $\mathcal{E}^n_{23,j}$, $\mathcal{E}^n_{24,j}$ and $\mathcal{E}^n_{25,j}$. First, for $s \in [0, \Delta t]$, we set

$$\varphi_{\delta,x}(s) = [\mathcal{R}(u,v) \star \rho_\delta]\,(t^n + s, x - \sqrt{a}(\Delta t - s)).$$

Then, from (4.2.2) and (4.2.6), we know that $|\partial_u \mathcal{R}(u,v)| \le \sqrt{a}\,\beta$ and $|\partial_v \mathcal{R}(u,v)| \le \beta$, for any $(u,v) \in I(N_0, a_0)$, we obtain

$$\begin{aligned}
\sum_{j\in\mathbb{Z}} \Delta x \Delta t \left|\mathcal{E}^n_{23,j}\right| &\le \frac{1}{\varepsilon}\int_{\mathbb{R}} \left| \int_0^{\Delta t}\int_0^s \varphi'_{\delta,x}(\eta)\,\mathrm{d}\eta\,\mathrm{d}s \right| \mathrm{d}x, \\
&\le C\,\frac{\Delta t}{\varepsilon}\int_{\mathbb{R}}\int_{t^n}^{t^{n+1}} \left(|\partial_t u_\delta| + |\partial_x u_\delta|\right)(t,x)\,\mathrm{d}t\,\mathrm{d}x \\
&+ C\,\frac{\Delta t}{\varepsilon}\int_{\mathbb{R}}\int_{t^n}^{t^{n+1}} \left(|\partial_t v_\delta| + |\partial_x v_\delta|\right)(t,x)\,\mathrm{d}t\,\mathrm{d}x.
\end{aligned}$$

Thus we can use the estimates on the continuous relaxation system listed in Theorem 1.1. Indeed, since

$$\begin{cases} \partial_t u_\delta = -\partial_x v_\delta, \\ \\ \partial_t v_\delta = -a\,\partial_x u_\delta - \dfrac{1}{\varepsilon}\mathcal{R}_\delta(u,v), \end{cases}$$

we obtain, by applying a first order Taylor expansion of $\mathcal{R}$, the inequalities

$$\begin{aligned}
\int_{\mathbb{R}} \left(|\partial_t u_\delta| + |\partial_x u_\delta|\right)(t,x)\,\mathrm{d}x &\le TV(u(t)) + TV(v(t)), \\
\int_{\mathbb{R}} \left(|\partial_t v_\delta| + |\partial_x v_\delta|\right)(t,x)\,\mathrm{d}x &\le C\left(TV(u(t)) + \frac{1}{\varepsilon}\,\|(v - A(u))(t)\|_{L^1}\right).
\end{aligned}$$

Hence, integrating over $t \in (t^n, t^{n+1})$ and using (4.1.6) and (4.1.7), we get:

$$\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{23,j}\right| \le C\,\frac{\Delta t}{\varepsilon}\left(TV(u(t^n)) + TV(v(t^n)) + \frac{e^{-\beta_0 t^n/\varepsilon}}{\varepsilon}\,\|\delta^0\|_{L^1} + 1\right), \quad (4.5.7)$$

where $C > 0$ only depends on $\sqrt{a}$ and β.

Now we treat the term $\mathcal{E}^n_{24,j}$ using the smoothness properties of $\mathcal{R}$ (4.2.2) and (4.2.6), it gives

$$\begin{aligned}
\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{24,j}\right| &= \frac{1}{\varepsilon}\int_{\mathbb{R}} \left| \int_{\mathbb{R}} \left[\mathcal{R}(u,v)(t^n, x - y - \sqrt{a}\Delta t) - \mathcal{R}(u,v)(t^n, x - \sqrt{a}\Delta t)\right]\,\rho_\delta(y)\,\mathrm{d}y \right| \mathrm{d}x, \\
&\le \frac{C}{\varepsilon}\int_{\mathbb{R}^2} \left[|u(t^n,x) - u(t^n,x-y)| + |v(t^n,x) - v(t^n,x-y)|\,\right]\rho_\delta(y)\,\mathrm{d}y\,\mathrm{d}x.
\end{aligned}$$

Thus, applying Fubini's theorem the BV estimate on the exact solution (4.1.6) and the value of the integral of ρ_δ, we get

$$\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{24,j}\right| \;\leq\; C\,\frac{\delta}{\varepsilon}\left[\,TV(u(t^n)) \,+\, TV(v(t^n))\,\right]. \tag{4.5.8}$$

Finally, to deal with the last term $\mathcal{E}^n_{25,j}$, we split it in two parts

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \Delta x\,|\mathcal{E}^n_{25,j}| \;&\leq\; \frac{1}{\varepsilon}\int_{\mathbb{R}} \left|\mathcal{R}\left(u,v\right)(t^n, x-\sqrt{a}\Delta t) - \mathcal{R}\left(u_\delta,v_\delta\right)(t^n, x-\sqrt{a}\Delta t)\right|\,\mathrm{d}x\\ &+\; \frac{1}{\varepsilon}\sum_{j\in\mathbb{Z}}\int_{C_j} \left|\mathcal{R}\left(u_\delta,v_\delta\right)(t^n, x-\sqrt{a}\Delta t) - \mathcal{R}\left(\tilde{u}_j^{n+1/2},\tilde{v}_j^{n+1/2}\right)\right|\,\mathrm{d}x\end{aligned}$$

and treat the different terms as for $\mathcal{E}^n_{24,j}$, we get for the first one

$$\int_{\mathbb{R}} \left|\mathcal{R}\left(u,v\right)(t^n,x) - \mathcal{R}\left(u_\delta,v_\delta\right)(t^n,x)\right|\,\mathrm{d}x \;\leq\; C\,\delta\,\left[\,TV(u(t^n)) \,+\, TV(v(t^n))\,\right].$$

and for the latter one using the BV estimate on the exact solution (4.1.6),

$$\sum_{j\in\mathbb{Z}}\int_{C_j} \left|\mathcal{R}\left(u_\delta,v_\delta\right)(t^n, x-\sqrt{a}\Delta t) - \mathcal{R}\left(\tilde{u}_j^{n+1/2},\tilde{v}_j^{n+1/2}\right)\right|\,\mathrm{d}x \leq C\,\Delta x\,\left[\,\|\partial_x u_\delta(t^n)\|_{L^1} \,+\, \|\partial_x v_\delta(t^n)\|_{L^1},\right].$$

Thus, we have

$$\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{25,j}\right| \;\leq\; C\left(\frac{\delta}{\varepsilon}+\frac{\Delta x}{\varepsilon}\right)\left[\,TV(u(t^n)) \,+\, TV(v(t^n))\,\right]. \tag{4.5.9}$$

Gathering (4.5.6), (4.5.7) , (4.5.8) and (4.5.9), and finally using the uniform in time bound on the BV norms of (u,v), it yields

$$\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{2,j}\right| \;\leq\; C\left[\frac{\Delta t}{\varepsilon}\left(e^{-\beta_0\,t^n/\varepsilon}\,\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon}\,+\,1\right)+\frac{\Delta x}{\varepsilon}\,+\,\frac{\delta}{\varepsilon}\right].$$

Using the same arguments we also prove that

$$\sum_{j\in\mathbb{Z}} \Delta x \left|\mathcal{E}^n_{4,j}\right| \;\leq\; C\left[\frac{\Delta t}{\varepsilon}\left(e^{-\beta_0\,t^n/\varepsilon}\,\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon}\,+\,1\right)+\frac{\Delta x}{\varepsilon}\,+\,\frac{\delta}{\varepsilon}\right].$$

□

4.5.2 Convergence proof.

Now we perform a rigorous analysis of the numerical scheme (4.2.13)-(4.2.14) when $h = (\Delta t, \Delta x)$ goes to zero and prove Proposition 5.1. We consider the numerical solution $(u_h^\varepsilon, v_h^\varepsilon)$ to the scheme (4.2.13)-(4.2.14) and $(u^\varepsilon, v^\varepsilon)$ the exact solution to (4.1.1)-(4.1.2) and define $(w^\varepsilon, z^\varepsilon)$ using (4.3.1). Then we denote by

$$\bar{w}_j^n = \frac{1}{\Delta x}\int_{C_j} w^\varepsilon(t^n, x)\,\mathrm{d}x, \quad \bar{z}_j^n = \frac{1}{\Delta x}\int_{C_j} z^\varepsilon(t^n, x)\,\mathrm{d}x$$

and $(w_j^n, z_j^n)_{(j,n)\in\mathbb{Z}\times\mathbb{N}}$ the numerical solution given by (4.3.2)-(4.3.3). Thus,

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \Delta x \left[\, |w_j^n - \bar{w}_j^n| + |z_j^n - \bar{z}_j^n| \,\right] \;&\le\; \sum_{j\in\mathbb{Z}} \Delta x \left[\, |w_j^n - \tilde{w}_j^n| + |z_j^n - \tilde{z}_j^n| \,\right] \\ &+ \sum_{j\in\mathbb{Z}} \Delta x \left[\, |\tilde{w}_j^n - \bar{w}_j^n| + |\tilde{z}_j^n - \bar{z}_j^n| \,\right],\end{aligned}$$

where $(\tilde{w}_j^n, \tilde{z}_j^n)_{(j,n)\in\mathbb{Z}\times\mathbb{N}}$ is given by (4.5.4). On the one hand, we estimate the second terms of the right hand side using the convolution properties and have

$$\sum_{j\in\mathbb{Z}} \Delta x \left[\, |\tilde{w}_j^n - \bar{w}_j^n| + |\tilde{z}_j^n - \bar{z}_j^n| \,\right] \le C\,\delta\,[\,TV(u) + TV(v)\,]. \tag{4.5.10}$$

On the other hand, we apply the consistency error analysis to estimate the first term of the right hand side. Applying (4.3.5)- (4.3.6) in Lemma 3.1 with $(\tilde{w}_j, \tilde{z}_j)$ and (w_j, z_j), it yields

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \Delta x\, |\tilde{w}_j^{n+1} - w_j^{n+1}| \;&\le\; \sum_{j\in\mathbb{Z}} \Delta x\, |\tilde{w}_j^{n+1/2} - w_j^{n+1/2}| \left(1 + \partial_w G_{\varepsilon,\Delta t}(w_j, z_j^{n+1/2})\right) \\ &+ \sum_{j\in\mathbb{Z}} \Delta x |\tilde{z}_j^{n+1/2} - z_j^{n+1/2}|\, \partial_z G_{\varepsilon,\Delta t}(\tilde{w}_j^{n+1/2}, z_j) \\ &+ \sum_{j\in\mathbb{Z}} \Delta x\, \Delta t \left[\, |\mathcal{E}_{1,j}^n| + |\mathcal{E}_{2,j}^n| \,\right]\end{aligned}$$

and

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \Delta x\, |\tilde{z}_j^{n+1} - z_j^{n+1}| \;&\le\; \sum_{j\in\mathbb{Z}} \Delta x\, |\tilde{z}_j^{n+1/2} - z_j^{n+1/2}| \left(1 - \partial_z G_{\varepsilon,\Delta t}(\tilde{w}_j^{n+1/2}, z_j)\right) \\ &- \sum_{j\in\mathbb{Z}} \Delta x |\tilde{w}_j^{n+1/2} - w_j^{n+1/2}|\, \partial_w G_{\varepsilon,\Delta t}(w_j, z_j^{n+1/2}). \\ &+ \sum_{j\in\mathbb{Z}} \Delta x\, \Delta t \left[\, |\mathcal{E}_{3,j}^n| + |\mathcal{E}_{4,j}^n| \,\right].\end{aligned}$$

Summing the two inequalities and using that the scheme (4.3.3) is TVD, we get the following inequality

$$\begin{aligned}\sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{n+1}-z_j^{n+1}|+|\tilde w_j^{n+1}-w_j^{n+1}|\,\right] &\le \sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{n}-z_j^{n}|+|\tilde w_j^{n}-w_j^{n}|\,\right]\\ &+ \sum_{j\in\mathbb{Z}}\Delta x\,\Delta t\left[\,|\mathcal{E}_{1,j}^n|+|\mathcal{E}_{2,j}^n|+|\mathcal{E}_{3,j}^n|+|\mathcal{E}_{4,j}^n|\,\right].\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{n+1}-z_j^{n+1}|+|\tilde w_j^{n+1}-w_j^{n+1}|\,\right] &\le \sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{0}-z_j^{0}|+|\tilde w_j^{0}-w_j^{0}|\,\right]\\ &+\sum_{k=0}^{n}\sum_{j\in\mathbb{Z}}\Delta x\,\Delta t\left[\,|\mathcal{E}_{1,j}^k|+|\mathcal{E}_{2,j}^k|+|\mathcal{E}_{3,j}^k|+|\mathcal{E}_{4,j}^k|\,\right].\end{aligned}$$

Finally the consistency error analysis performed in Propositions 5.2 and 5.3, we have taking $\delta=\Delta x^{1/2}$

$$\begin{aligned}\sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{n+1}-z_j^{n+1}|+|\tilde w_j^{n+1}-w_j^{n+1}|\,\right] &\le \sum_{j\in\mathbb{Z}}\Delta x\,\left[\,|\tilde z_j^{0}-z_j^{0}|+|\tilde w_j^{0}-w_j^{0}|\,\right]\\ &+\frac{C}{\varepsilon}\left(\Delta t\,(\Delta t+\varepsilon)\left(\frac{\left\|\delta^0\right\|_{L^1}}{\varepsilon}+1\right)+t^n\left[\Delta x+\varepsilon\Delta x^{1/2}+\Delta x^{1/2}\right]\right).\end{aligned}\tag{4.5.11}$$

Gathering (4.5.10) and (4.5.11), we get (4.5.1), in which the right hand side goes to zero as h goes to 0. This ends the proof of Theorem 2.4.

4.6 Numerical simulations for the Broadwell system

In this section, we apply our scheme to the Broadwell model, which can be seen as a simple one-dimensional lattice Boltzmann equation, with only four discrete velocities [7, 48, 135]. The gas is defined by a density function in phase space satisfying the equation

$$\left\{\begin{array}{rcl}\partial_t f_+ + \partial_x f_+ &=& -\frac{1}{\varepsilon}\left(f_+\,f_- - f_0^2\right),\\ \partial_t f_0 &=& \frac{1}{\varepsilon}\left(f_+\,f_- - f_0^2\right),\\ \partial_t f_- - \partial_x f_- &=& -\frac{1}{\varepsilon}\left(f_+\,f_- - f_0^2\right).\end{array}\right.$$

Here f_+, f_- and f_0 denote the particle density distribution at time t, position x with velocity 1, -1 and 0 respectively; ε is the mean free path. We can rewrite the system with fluid dynamical variables. We define the density ρ, the momentum m and z as:

$$\left\{\begin{array}{l}\rho = f_+ + 2\,f_0 + f_-,\\ m = f_+ - f_-,\\ z = f_+ + f_-.\end{array}\right.\tag{4.6.1}$$

The system can then be written as follows.

$$\left\{\begin{array}{rcl} \partial_t \rho + \partial_x m &=& 0\,, \\ \partial_t m + \partial_x z &=& 0\,, \\ \partial_t z + \partial_x m &=& -\frac{1}{\varepsilon}\,\left(\rho\, z - \frac{1}{2}\,\left(\rho^2 + m^2\right)\right)\,. \end{array}\right. \tag{4.6.2}$$

Hence, denoting

$$\mathbf{u} = (u_1, u_2) := (\rho, m) \text{ and } v := z\,,$$

we can rewrite the system under the form (4.1.1) as follows.

$$\left\{\begin{array}{rcl} \partial_t u_1 + \partial_x u_2 &=& 0\,, \\ \partial_t u_2 + \partial_x v &=& 0\,, \\ \partial_t v + \partial_x u_2 &=& -\frac{1}{\varepsilon}\,\mathcal{R}\,(\mathbf{u}, v)\,, \end{array}\right. \tag{4.6.3}$$

where

$$\mathcal{R}\,(\mathbf{u}, v) = u_1\, v - \frac{1}{2}\,\left(u_1^2 + u_2^2\right)\,.$$

The local equilibrium is defined by

$$v = A\,(\mathbf{u})\,,\ \text{where}\ A\,(\mathbf{u}) = \frac{1}{2}\,\left(u_1 + \frac{u_2^2}{u_1}\right)\,.$$

Hence, when ε goes to zero, we obtain the following "Euler" system:

$$\partial_t \mathbf{u} + \partial_x F(\mathbf{u}) \;=\; \mathbf{0}\,, \tag{4.6.4}$$

with

$$F(\mathbf{u}) \;=\; \left(\begin{array}{cc} u_2 & A(\mathbf{u}) \end{array}\right)^T\,.$$

Here, we have to examine the generalization of the subcharateristic condition (4.3.4). Indeed, the new stability criterion is expressed as follows. The eigenvalues of the limit problem (4.6.4) are required to be enterlaced between the ones of the relaxation problem (4.6.3) [29, 145, 129], that is:

$$-1 \;<\; \lambda_- \;<\; 0 \;<\; \lambda_+ \;<\; 1\,,$$

where $\lambda_\pm = \dfrac{1}{2}\,\left(\dfrac{u_2}{u_1} \pm \sqrt{2 - \dfrac{u_2^2}{u_1^2}}\right)$.

4.6.1 The Riemann problem

We present first several simulations for the problem (4.6.3) with different relaxation parameters ε, from the rarefied regime to the fluid regime. The initial data is given by the local equilibrium:

$$(\rho_0, m_0, z_0) = \left\{\begin{array}{ll} (1, 0, 1), & \text{if } -1 \le x \le 0, \\ & \\ (0.25, 0, 0.125), & \text{if } 0 < x \le 1. \end{array}\right. \tag{4.6.5}$$

We integrate the Broadwell system over the space domain $[-1, 1]$, with reflecting boundary condition and a Courant number $\lambda = 0.9$ and output the solution at different times. We used only 100 grid points, so that the time step is fixed equal to 0.002, and compare the results with a fully explicit solver (for which the time step has to be of the order of ε) for different values of the relaxation parameter.

In Figure 4.1, we take $\varepsilon = 0.5$ and 0.1. For such values of ε, the problem is not stiff and this test is performed to compare the accuracy of our AP scheme with a fully explicit scheme (global Lax-Friedrichs method with slope limiters and explicit Euler discretization in time). The density ($u_1 = \rho$), the momentum ($u_2 = m$), $v = z$, and the deviation to the equilibrium $v - A(\mathbf{u})$ are plotted at different times $t = 0.05$, 0.2, 0.35 and 0.5. At the kinetic regime, we can observe that our method gives the same accuracy as a standard fully explicit scheme.

Next we investigate the cases of small values of ε. The same time step for the AP scheme is used, whereas the fully explicit scheme requires it to be of order $O(\varepsilon)$. We report the numerical results for $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-3}$ in Figure 4.6.1. In this case, we add in the comparison the numerical solution to the *Euler* system (4.6.4), obtained with a standard first order finite volume scheme.

We observe that the AP scheme and the fully explicit scheme still agree, aven if the time step is at least ten times larger with our method. Moreover, now that we are closer to the fluid regime, we see that the macroscopic quantities are in good agreement with the ones obtained with the limit problem. Yet some differences between the AP and the explicit schemes can be observed for $\varepsilon = 10^{-3}$, this comes from the fact that we used a very small number of points for both discretizations.

Finally, an approximation of the L^1 error is plotted in logscale in Figure 4.2, that is, we compute

$$E = \|(\mathbf{u}_h^\varepsilon - \mathbf{u}_{2h}^\varepsilon, v_h^\varepsilon - v_{2h}^\varepsilon)\|,$$

where h is the discretization parameter and $\|\cdot\|$ is the discrete L^1 norm.

We observe that for discontinuous initial data, we get an order of convergence which is not better than $\sqrt{h}$, as in the estimate (4.5.11).

4.6.2 Approximation of smooth solutions

For this test, we considered the smooth initial data [158]:

$$\begin{aligned} \rho_0(x) &= 1 + 0.2\,\sin(\pi\, x), \\ m_0(x) &= 0, \\ z_0(x) &= \tfrac{1}{2}\,\left(\rho_0 + m_0^2/\rho_0\right), \end{aligned} \tag{4.6.6}$$

which is again in the local equilibrium. We used periodic boundary condition and $\lambda = 0.9$.

We first plot the L^1 , L^2 and L^∞ errors at time $t = 1$ (Figures 4.3, 4.4 and 4.5), for different relaxation regimes. We observe that the order of accuracy of our method is not deteriorated, from big values to very small values of the relaxation parameter ε.

Next our goal is to investigate numerically the long-time behavior of the solution to the Broadwell system. If $f = (f_+, f_0, f_-)^T$ is a reasonable smooth solution to (4.6.1), then it is expected to converge to the (unique) global equilibrium $\mathcal{M}_g$ as t goes to $+\infty$. (Desvillettes-Villani, Filbet-Jin, Filbet-Mouhot-Pareschi, Guo-Strain for the Boltzmann equation). In order to observe this damping phenomenon for the simpler Broadwell model, as well as the time oscillations conjectured by Desvillettes and Villani, we investigate the behavior of the quantities

$$\mathcal{E}_1 = \|\rho(t) - \rho_g\|_{L^1}, \quad \mathcal{E}_2 = \|m(t) - m_g\|_{L^1},$$

where, in our case, the global equilibrium is:

$$\rho_g = 1, \quad m_g = 0.$$

Therefore the initial local equilibrium data (4.6.6) is a perturbation of the global equilibrium.

In Figure 4.6, one can then observe oscillations of $\mathcal{E}_1$, $\mathcal{E}_2$ for the relaxation model and the Euler system. The frequency of oscillations does not depend on ε, contrary to the slope of the envelop curve is smaller when ε decreases, that is when we get closer to the hydrodynamic regime. In other words, it appears that the equilibration is much more rapid in the rarefied regime (ε large), and the convergence seems exponential. While there is no equilibration in the hydrodynamic regime, where the quantities $\mathcal{E}_1$, $\mathcal{E}_2$ are simply transported.

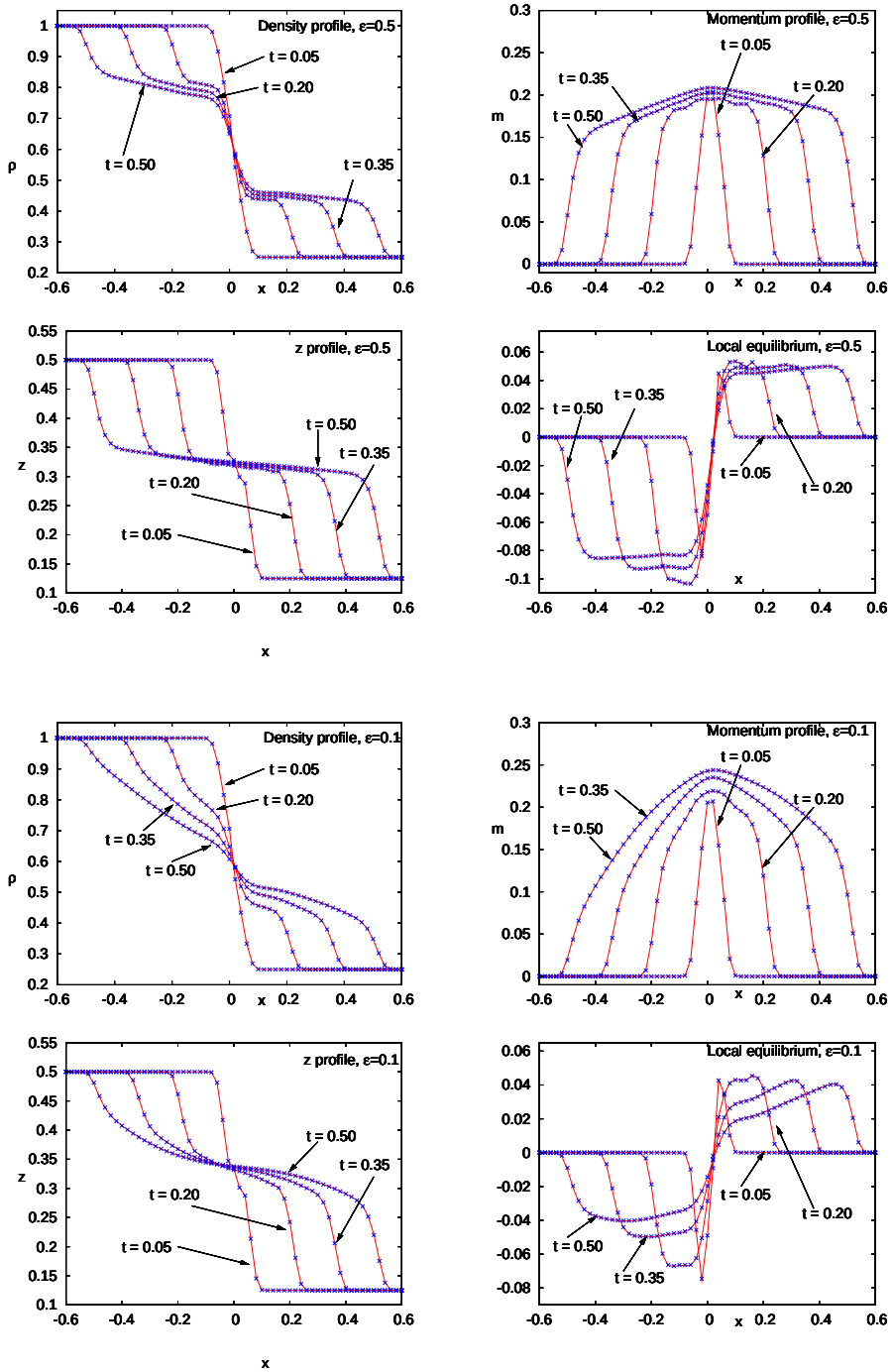
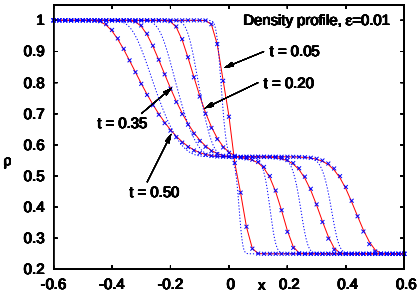
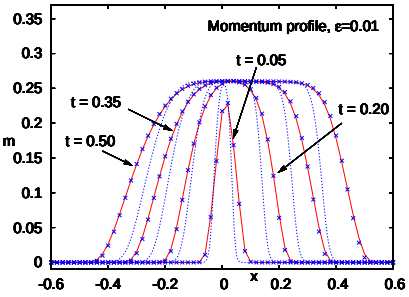
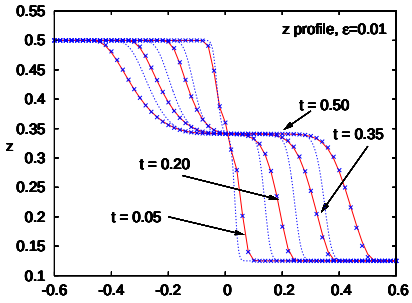
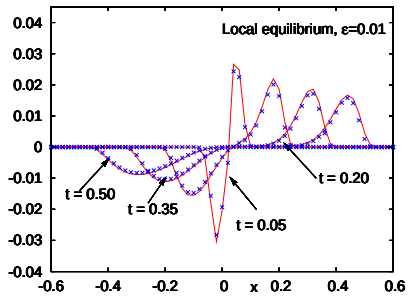


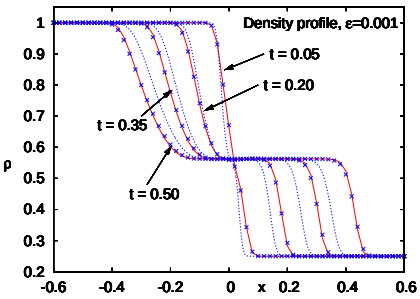
Figure 4.1: *Solution to the AP scheme (crosses) and solution the explicit solver (lines) with initial data (4.6.5) in the kinetic regime:* $\varepsilon = 0.5$ *(top) and* $\varepsilon = 0.1$ *(bottom). Evolution of the density* u_1*, the momentum* u_2*,* v *and the deviation to the local equilibrium* $v - A(\mathbf{u})$*.*

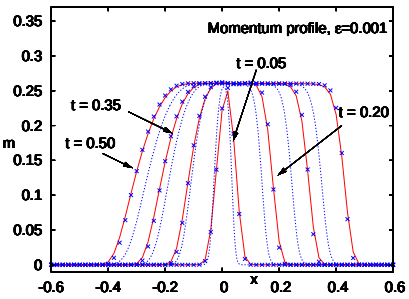

Density profile, ε=0.01
t = 0.05
t = 0.20
t = 0.35
t = 0.50
ρ
x

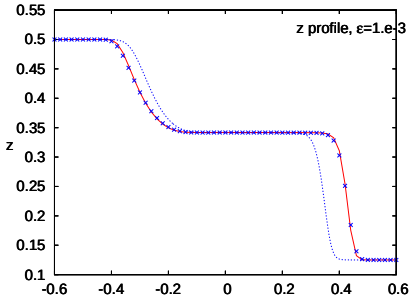

Momentum profile, ε=0.01
t = 0.05
t = 0.35
t = 0.20
t = 0.50
m
x

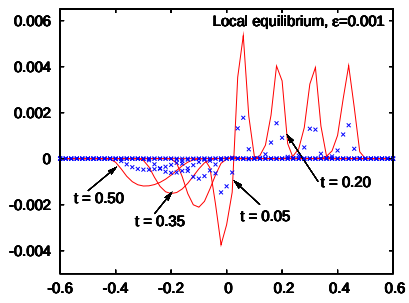

z profile, ε=0.01
t = 0.50
t = 0.35
t = 0.20
t = 0.05
z
x


Local equilibrium, ε=0.01
t = 0.50
t = 0.35
t = 0.05
t = 0.20
x


Density profile, ε=0.001
t = 0.05
t = 0.20
t = 0.35
t = 0.50
ρ
x


Momentum profile, ε=0.001
t = 0.05
t = 0.35
t = 0.20
t = 0.50
m
x


z profile, ε=1.e-3
z
x


Local equilibrium, ε=0.001
t = 0.50
t = 0.35
t = 0.05
t = 0.20

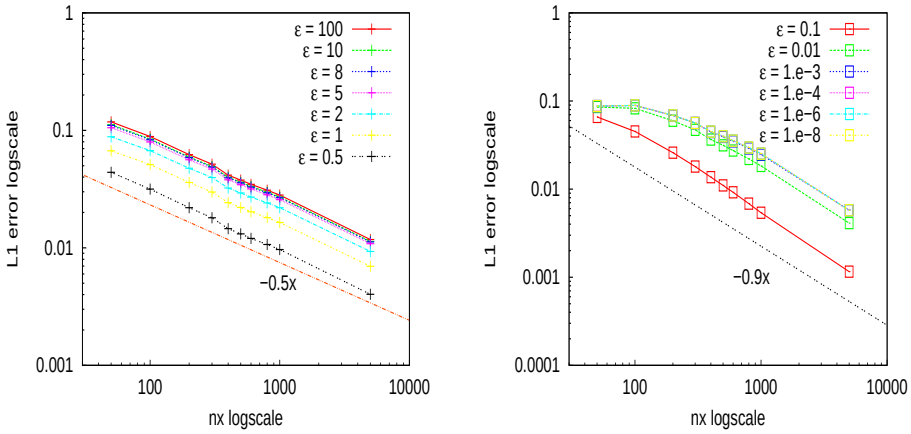


Figure 4.2: L^1 *error at time* $t = 1$ *for different regimes, with initial data (4.6.5).*

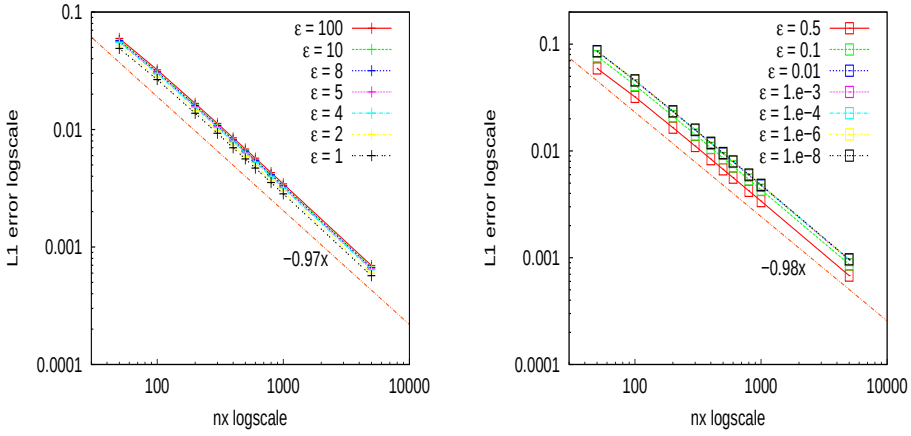


Figure 4.3: L^1 *error at time* $t = 1$ *for different regimes, with initial data (4.6.6).*

Figure 4.4: L^2 *error at time* $t = 1$ *for different regimes with initial data (4.6.6).*

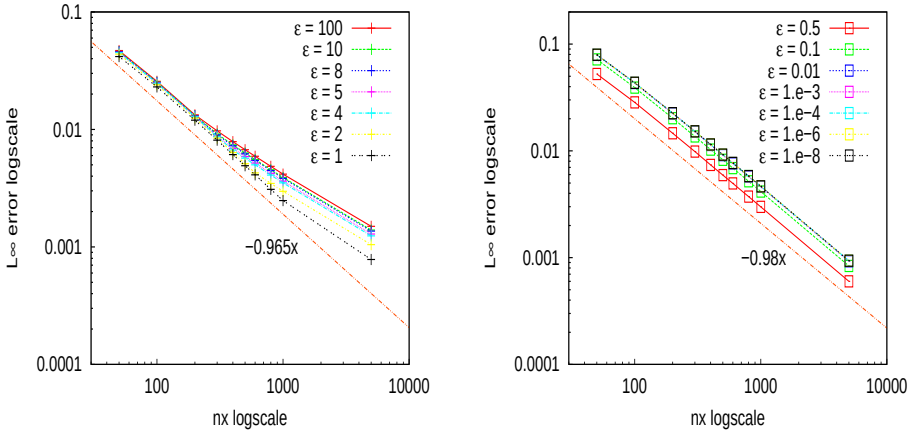


Figure 4.5: L^∞ *error at time* $t = 1$ *for different regimes with initial data (4.6.6).*

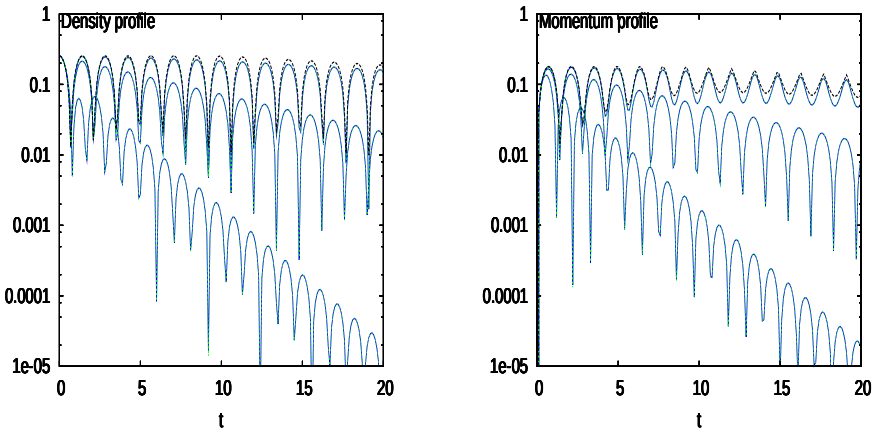


Figure 4.6: *Influence of the relaxation parameter ε: evolution of $\mathcal{E}_1$ (left) and $\mathcal{E}_2$ (right) for $\varepsilon = 0.5$ (biggest slope of the envelop curve), 0.05 and 0.005 (smallest slope), and comparison with the Euler system (dashed black line).*

Annexe B

Compléments sur le système continu : preuve de la convergence vers l'équilibre local

Dans cette annexe, nous établissons rigoureusement la convergence vers l'équilibre local au niveau continu pour le problème étudié au Chapitre 4, autrement dit la limite

$$\mathcal{P}^\varepsilon \underset{\varepsilon\to 0}{\longrightarrow} \mathcal{P}^0$$

du diagramme AP. Précisément nous montrons le théorème 1.1 énoncé au Chapitre précédent. Il s'agit ainsi d'adapter à notre cas (terme de relaxation non linéaire) les arguments de R. Natalini [152] pour la relaxation semi-linéaire de Jin et Xin. La preuve est composée des mêmes étapes que pour le problème discret : nous obtenons d'abord des estimations L^∞ et BV uniformes en ε, avant de passer à la limite.

B.1 Introduction

B.1.1 Rappel des notations et hypothèses

Rappelons le système hyperbolique que nous étudions :

$$\begin{cases} \partial_t u^\varepsilon + \partial_x v^\varepsilon = 0\,, \\ \\ \partial_t v^\varepsilon + a\,\partial_x u^\varepsilon = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u^\varepsilon, v^\varepsilon), \end{cases} \tag{B.1.1}$$

complété par les conditions initiales

$$\begin{cases} u^\varepsilon(0,x) = u_0^\varepsilon(x)\,, \\ \\ v^\varepsilon(0,x) = v_0^\varepsilon(x)\,. \end{cases} \tag{B.1.2}$$

Nous rappelons ensuite les hypothèses sur le terme source. La fonction $\mathcal{R} \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R}, \mathbb{R})$ possède un unique équilibre local défini par $\{v = A(u)\}$, c'est-à-dire :

$$\mathcal{R}(u, v) = 0 \iff v = A(u)\,, \tag{B.1.3}$$

où A est une fonction continue localement Lipschitz telle que $A(0) = 0$. En d'autres termes, la fonction $\mathcal{R}$ satisfait les hypothèses du théorème des fonctions implicites, et en réalité un peu plus, puisque l'on demande non seulement que sa dérivée partielle par rapport à la seconde variable ne change pas de signe, mais en outre qu'elle soit strictement positive pour assurer le caractère dissipatif du problème (voir la condition (B.1.6) ci-après). Le problème limite s'écrit donc :

$$\begin{cases} \partial_t u + \partial_x A(u) = 0\,, \text{ sur } \mathbb{R}^+ \times \mathbb{R}, \\ \\ u(t = 0) = u_0\,, \end{cases} \tag{B.1.4}$$

où la condition initiale u_0 est donnée par :

$$u_0 \;=\; \lim_{\varepsilon \to 0} u_0^\varepsilon \;\text{ dans }\; L^1_{loc}(\mathbb{R})\,. \tag{B.1.5}$$

Enfin, les hypothèses et notations que nous utiliserons dans la suite (déjà énoncées au chapitre précédent) sont réécrites ci-dessous.
On se donne un $U_0 > 0$ tel que, pour tout (u, v) dans $[-U_0, U_0] \times \mathbb{R}$:

$$\begin{cases} |\partial_u \mathcal{R}(u, v)| \;\leq\; g(U_0)\,, \\ \\ 0 \;<\; \beta_0(U_0) \;\leq\; \partial_v \mathcal{R}(u, v) \;\leq\; h(U_0)\,, \end{cases} \tag{B.1.6}$$

où β_0, g et h ne dépendent que de U_0.

La condition sous-caractéristique s'écrit :

$$\left| \frac{\partial_u \mathcal{R}\,(u, v)}{\partial_v \mathcal{R}\,(u, v)} \right| \;<\; \sqrt{a}\,. \tag{B.1.7}$$

On définit ensuite, pour tous $N > 0$ et $\alpha > 0$,

$$\begin{cases} U(N, \alpha) := \left(1 + \frac{1}{\sqrt{\alpha}}\right) N\,, \\ \\ F(N, \alpha) := \sup\limits_{|\xi| \leq U(N,\alpha)} |A(\xi)|\,, \\ \\ V(N, \alpha) := U(N, \alpha) \;+\; \frac{1}{\sqrt{\alpha}} \frac{h(U(N,\alpha))}{\beta_0(U(N,\alpha))} F\,(N, \alpha)\;. \end{cases} \tag{B.1.8}$$

Le sous-ensemble convexe compact $I(N, \alpha)$ de $\mathbb{R}^2$ dans lequel sera confinée la solution est défini par :

$$I(N, \alpha) \;:=\; \left[-\sqrt{a}\,V(N, \alpha),\, \sqrt{a}\,V(N, \alpha)\right]^2\;. \tag{B.1.9}$$

Enfin, nous introduisons

$$N_0 := \max\left\{\sup_{\varepsilon>0}\|u_0^\varepsilon\|_{L^\infty}, \sup_{\varepsilon>0}\|v_0^\varepsilon\|_{L^\infty}\right\} < \infty\,. \tag{B.1.10}$$

Ainsi, pour un certain $a_0 > 0$, nous choisissons la fonction $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ satisfaisant (B.1.3) et (B.1.6), ainsi que la vitesse caractéristique $\sqrt{a} > 0$ et le paramètre $\beta > 0$ tel que :

$$\begin{cases} \sqrt{a} > \max\left\{1\,,\,\sqrt{a_0}\,,\,\dfrac{g\left(V(N_0,a_0)\right)}{\beta_0\left(V(N_0,a_0)\right)}\right\}, \\ \\ \beta = h\left(V(N_0,a_0)\right), \end{cases} \tag{B.1.11}$$

où V est donné par (B.1.8).

Rappelons maintenant quelques résultats classiques sur les systèmes semi-linéaires. Ces résultats seront utiles pour l'étude de notre problème sous sa formulation diagonale. Pour plus de détails ainsi que pour les preuves, nous renvoyons le lecteur à l'article [108], dans lequel B. Hanouzet et R. Natalini traitent même le cas plus général des systèmes *quasi-linéaires*, mais le cadre semi-linéaire suffit à notre étude.

B.1.2 Rappels sur les systèmes semi-linéaires

Considérons un problème de Cauchy générique semi-linéaire et hyperbolique, de la forme :

$$\begin{cases} \partial_t u_i + \lambda_i\,\partial_x u_i = h_i(U)\,,\ i = 1\ldots n\,, \\ \\ U(0,x) = U_0(x)\,, \end{cases} \tag{B.1.12}$$

où $U = (u_1,\ldots,u_n) \in \mathbb{R}^n$, les λ_i ($i \in \{1,\ldots,n\}$) sont les valeurs propres réelles, et le terme source $H = (h_1,\ldots,h_n)$ est une fonction donnée continue et localement lipschitzienne sur $\mathbb{R}^n$. La donnée initiale est choisie avec la régularité suivante,

$$U_0 = (u_{1,0},\ldots,u_{n,0}) \in L^\infty(\mathbb{R})^n\,.$$

La notion de solution que nous considèrerons est celle de solution faible au sens suivant.

Définition 13. On dit qu'une fonction $U \in L^\infty((0,T)\times\mathbb{R})^n$ ($T > 0$) est une solution faible du problème de Cauchy (B.1.12) si pour toute fonction test $\varphi \in \mathcal{C}_0^\infty((0,T)\times\mathbb{R})$ et pour tout $i = 1,\ldots,n$ on a :

$$\int\int \Big[u_i\left(\partial_t\varphi + \lambda_i\,\partial_x\varphi\right) + h_i(U)\,\varphi\Big]\,\mathrm{d}x\,\mathrm{d}t = 0\,,$$

De plus, pour tout intervalle $J \subset \mathbb{R}$:

$$\lim_{T\to 0}\frac{1}{T}\int_0^T\int_J |u_i(t,x) - u_{i,0}(x)|\,\mathrm{d}x\,\mathrm{d}t = 0\,.$$

Le théorème suivant assure l'existence de telles solutions.

Théorème 14. *(Existence et unicité de la solution faible)*

Pour toute donnée initiale $U_0 \in L^\infty(\mathbb{R})$, il existe un $T^* > 0$ (dépendant seulement de la norme infinie de la condition initiale) tel que dans $(0, T^*) \times \mathbb{R}$, il existe une unique solution faible U du problème de Cauchy (B.1.12) vérifiant :

$$U \in \mathcal{C}\left([0, T^*]; L^1_{loc}(\mathbb{R})^n\right) .$$

En outre, il n'y a que deux possibilités :

- *soit $T^* = +\infty$ et $U \in L^\infty\left((0,T) \times \mathbb{R}\right)$ pour tout $T > 0$,*
- *soit $T^* < \infty$ et $\lim_{T \to T^*} \|U\|_{L^\infty((0,T)\times\mathbb{R})} = +\infty$.*

Un principe de comparaison des solutions pour de tels problèmes semi-linéaires est donné dans le théorème suivant.

Théorème 15. *(Principe de comparaison pour les systèmes semi-linéaires)*

On note $L = \max\{|\lambda_i|\}$. Soient U et $\tilde{U}$ deux solutions faibles sur $(0,T)\times\mathbb{R}$ de (B.1.12), associées respectivement aux conditions initiales U_0 et $\tilde{U}_0$ et aux termes sources H et $\tilde{H}$. Alors pour tout intervalle $[c,d]$ de $\mathbb{R}$, pour presque tout $t \in (0, \min(T, (d-c)/2L))$ on a :

$$\sum_{i=1}^{n} \int_c^d |u_i(t,x) - \tilde{u}_i(t,x)|\, dx \le \sum_{i=1}^{n} \int_{c-Lt}^{d+Lt} |u_{i,0}(x) - \tilde{u}_{i,0}(x)|\, dx$$
$$+ \sum_{i=1}^{n} \int_0^t \int_{c-L(t-s)}^{d+L(t-s)} \operatorname{sgn}\Big(u_i(s,x) - \tilde{u}_i(s,x)\Big)\Big(h_i(U(s,x)) - \tilde{h}_i(\tilde{U}(s,x))\Big)\, dx\, ds\,. \quad \text{(B.1.13)}$$

Nous verrons que le problème que nous étudions appartient à une famille restreinte des systèmes semi-linéaires hyperboliques, celle des systèmes *quasi-monotones*. De fait, le principe de comparaison précédent aura une formulation plus simple dans notre cas (voir le Théorème 17 ci-après).

Définition 16. Soient Ω un ouvert convexe de $\mathbb{R}^n$ et H une fonction de Ω dans $\mathbb{R}^n$. H est dite *quasi-monotone* non décroissante si chacune de ses composantes h_i est non décroissante par rapport aux variables x_j pour $j \neq i$. De plus, le système semi-linéaire (B.1.12) est dit *quasi-monotone* si le terme source $H = (h_1, \ldots, h_n)$ est quasi-monotone.

Les systèmes quasi-monotones font l'objet d'un principe de comparaison plus fort que le Théorème précédent et qui nous sera utile pour l'obtention d'estimations *a priori* uniformes en ε.

Théorème 17. *(Principe de comparaison pour les systèmes quasi-monotones)*

Soient U et $\tilde{U}$ deux solutions faibles sur $(0,T)\times\mathbb{R}$ de (B.1.12), associées respectivement aux conditions initiales U_0 et $\tilde{U}_0$. Soit Ω un ouvert convexe de $\mathbb{R}^n$ tel que :

- *H est quasi-monotone sur Ω,*
- *Pour presque tout $(t,x) \in (0,T)\times\mathbb{R}$, U et $\tilde{U}$ appartiennent à Ω.*

Si $U_0 \leqslant \tilde{U}_0$, pour presque tout $x \in \mathbb{R}$, alors $U \leqslant \tilde{U}$ pour presque tout $(t,x) \in (0,T)\times\mathbb{R}$.

B.2 Estimations *a priori*

Dans cette section, nous établissons des estimations sur $(u^\varepsilon, v^\varepsilon)$ qui sont *uniformes* par rapport au paramètre de relaxation ε. Nous omettrons la plupart du temps les exposants ε par souci de clarté. L'obtention des estimations repose essentiellement sur la structure hyperbolique et quasi-linéaire de la formulation diagonale du système (B.1.1), à savoir :

$$\begin{cases} \partial_t w + \sqrt{a}\,\partial_x w = \dfrac{1}{\varepsilon}\,G(w,z)\,, \\[2ex] \partial_t z - \sqrt{a}\,\partial_x z = -\dfrac{1}{\varepsilon}\,G(w,z)\,, \end{cases} \tag{B.2.1}$$

où w et z sont données par :

$$w = -v - \sqrt{a}\,u\,, \quad z = v - \sqrt{a}\,u\,, \tag{B.2.2}$$

tandis que le terme source s'écrit :

$$G(w,z) = \mathcal{R}\left(\frac{-w-z}{2\sqrt{a}}, \frac{-w+z}{2}\right). \tag{B.2.3}$$

Enfin, les conditions initiales du problème diagonal sont :

$$\begin{cases} w(0,x) = w_0(x) = -v_0 - \sqrt{a}\,u_0\,, \\[2ex] z(0,x) = z_0(x) = v_0 - \sqrt{a}\,u_0\,. \end{cases} \tag{B.2.4}$$

Nous verrons en particulier que ce système est quasi-monotone au sens de la Définition 16 sous réserve que la condition sous-caractéristique (B.1.7) soit satisfaite.

B.2.1 Estimations L^∞

Proposition 18. *(Estimations L^∞)*
Soient N_0 défini par (B.1.10) et $a_0 > 0$. Soit la fonction $\mathcal{R} \in \mathcal{C}^1(\mathbb{R}\times\mathbb{R},\mathbb{R})$ vérifiant (B.1.3) et (B.1.6). On choisit $a > 0$ et $\beta > 0$ tels que les conditions (B.1.11) soient vérifiées, à savoir :

$$\begin{cases} \sqrt{a} > \max\left\{1\,,\ \sqrt{a_0}\,,\ \dfrac{g\left(V(N_0,a_0)\right)}{\beta_0\left(V(N_0,a_0)\right)}\right\}, \\[2ex] \beta = h\left(V(N_0,a_0)\right), \end{cases}$$

où V est donnée par (B.1.8), et h et g sont données par (B.1.6). Alors il existe une unique solution faible globale $(u^\varepsilon, v^\varepsilon)$ du problème de Cauchy (B.1.1)-(B.1.2) dans $\mathcal{C}\left([0,\infty[, L^1_{loc}(\mathbb{R})^2\right)$. De plus il existe une constante $C>0$ dépendant uniquement de a, a_0 et N_0, telle que, pour tout $\varepsilon>0$:

$$\left\|v^\varepsilon(t) \pm \sqrt{a}\, u^\varepsilon(t)\right\|_{L^\infty} \leq C \quad \forall\, t>0. \tag{B.2.5}$$

Enfin, la condition sous-caractéristique est vérifiée : pour tout $\varepsilon>0$, pour presque tout (t,x) dans $(0,\infty)\times\mathbb{R}$:

$$\left|\frac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}}\Big(u^\varepsilon(t,x); v^\varepsilon(t,x)\Big)\right| < \sqrt{a}\,. \tag{B.2.6}$$

Démonstration. Omettons les exposants ε. Nous considérons la formulation diagonale (B.2.1). Commençons par supposer les conditions initiales plus régulières : w_0 et z_0 dans $\mathcal{C}^1_0(\mathbb{R})$. Comme la fonction $\mathcal{R}$ est de classe $\mathcal{C}^1$, le Théorème 14 assure l'existence d'un $T^*>0$ tel qu'il existe une unique solution faible du problème de Cauchy (B.2.1)-(B.2.4) :

$$(w,z) \in \mathcal{C}\left([0,T^*]; L^1_{loc}(\mathbb{R})^2\right).$$

En réalité, la régularité de la condition initiale assure même que :

$$(w,z) \in \mathcal{C}^1\left([0,T^*]\times\mathbb{R}\right)^2.$$

Pour l'instant, T^* peut dépendre de ε, mais nous allons établir que $T^*=+\infty$.
Le choix de a assure que la condition sous-caractéristique est satisfaite initialement. En effet, pour tout x de $\mathbb{R}$:

$$|u_0(x)| \leqslant N_0 \leqslant V(N_0,a_0)\,,$$

donc d'après (B.1.6) et (B.1.11) :

$$\left|\frac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}}(u_0(x),v_0(x))\right| \leqslant \frac{g(V(N_0,a_0))}{h(V(N_0,a_0))} < \sqrt{a}\,.$$

De plus la dernière inégalité étant stricte, il existe un $\eta>0$ tel que

$$\sqrt{a} \geqslant \frac{g(V(N_0,a_0))}{h(V(N_0,a_0))} + \eta\,,$$

donc la continuité en temps de la solution et des dérivées de $\mathcal{R}$ assurent que le temps T_η défini par

$$T_\eta := \sup\left\{T \leqslant T^* : \quad \left\|\frac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}}(u,v)\right\|_{L^\infty((0,T)\times\mathbb{R})} + \frac{\eta}{2} \leqslant \sqrt{a}\right\},$$

est strictement positif. La condition sous-caractéristique est donc assurée jusqu'au temps $T_\eta>0$.

Nous pouvons maintenant montrer que le système est quasi-monotone sur $(0,T_\eta)\times\mathbb{R}$, grâce à la condition sous-caractéristique. En effet, le système (B.2.1) est quasi-monotone

au sens de la Définition 16 sur un ouvert Ω non vide convexe de $\mathbb{R}^2$ si pour tout (w,z) de Ω :

$$\partial_w G(w,z) \leqslant 0\,, \quad \partial_z G(w,z) \geqslant 0\,.$$

Or les dérivées partielles de G s'écrivent sous la forme :

$$\begin{cases} \partial_w G(w,z) = \dfrac{-\partial_v \mathcal{R}}{2}\left(1+\dfrac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}\,\sqrt{a}}\right)\left(\dfrac{-w-z}{2\sqrt{a}}, \dfrac{-w+z}{2}\right), \\ \\ \partial_z G(w,z) = \dfrac{\partial_v \mathcal{R}}{2}\left(1-\dfrac{\partial_u \mathcal{R}}{\partial_v \mathcal{R}\,\sqrt{a}}\right)\left(\dfrac{-w-z}{2\sqrt{a}}, \dfrac{-w+z}{2}\right). \end{cases}$$

Ainsi, pour tout (t,x) de $(0,T_\eta)\times\mathbb{R}$, la solution (w,z) de (B.2.1)-(B.2.4) reste dans un convexe Ω de $\mathbb{R}^2$ sur leque la condition sous-caractéristique est satisfaite, ce qui entraîne la quasi-monotonie du système.

Nous disposons donc du principe de comparaison donné par le Théorème 17, au moins sur $(0,T_\eta)\times\mathbb{R}$. C'est avec un choix particulier d'une sur-solution $(\overline{w},\overline{z})$ et d'une sous-solution $(\underline{w},\underline{z})$ du problème (B.2.1) que nous obtiendrons à la fois la borne L^∞ indépendante de ε pour u et v, et le fait que l'on peut prendre $T_\eta = T^* = +\infty$. Précisément, nous allons construire deux solutions du système (B.2.1) sur $(0,T_\eta)\times\mathbb{R}$, avec des conditions initiales telles que :

$$\begin{cases} \underline{w}_0 \leq w_0 \leq \overline{w}_0\,, \\ \\ \underline{z}_0 \leq z_0 \leq \overline{z}_0\,, \end{cases}$$

ce qui entraînera, en appliquant le Théorème 17 que pour tout (t,x) de $(0,T_\eta)\times\mathbb{R}$ les relations d'ordre sont conservées :

$$\begin{cases} \underline{w}(t,x) \leq w(t,x) \leq \overline{w}(t,x)\,, \\ \\ \underline{z}(t,x) \leq z(t,x) \leq \overline{z}(t,x)\,. \end{cases}$$

Afin d'exhiber $(\underline{w},\underline{z})$ et $(\overline{w},\overline{z})$, nous résolvons, comme dans le cas de la relaxation semi-linéaire [152], le système différentiel ordinaire :

$$\begin{cases} w' = \dfrac{1}{\varepsilon}\,G(w,z)\,, \\ \\ z' = -\dfrac{1}{\varepsilon}\,G(w,z)\,, \end{cases} \tag{B.2.7}$$

avec les conditions initiales

$$\begin{cases} w(0) = R_0\,, \\ \\ z(0) = R_0\,, \end{cases} \tag{B.2.8}$$

où R_0 sera choisi ultérieurement, soit plus grand, soit plus petit que w_0 et z_0. A la différence du problème de Jin et Xin, ce système ne s'intègre pas exactement, mais nous pouvons

cependant obtenir des bornes L^∞ sur les solutions, ce qui suffira. Pour cela, nous préférons la formulation en u et v de ce système, à savoir :

$$\left\{ \begin{array}{l} u' = 0\,, \\ \\ v' = -\dfrac{1}{\varepsilon}\,\mathcal{R}(u,v)\,, \end{array} \right. \tag{B.2.9}$$

avec les conditions initiales

$$\left\{ \begin{array}{l} u(0) = -\dfrac{1}{\sqrt{a}}\,R_0 := U_0\,, \\ \\ v(0) = 0\,. \end{array} \right. \tag{B.2.10}$$

Ce système différentiel ordinaire possède une unique solution *globale* que nous allons « expliciter » grâce au développement de Taylor de $\mathcal{R}$ et à la formule de Duhamel. Tout d'abord la fonction u est identiquement égale à U_0 pour tout temps. Ensuite, la formule de Taylor appliquée à $\mathcal{R}$ entre les points $(U_0, A(U_0))$ et $(U_0, v(t))$ donne, pour tout $t > 0$:

$$\mathcal{R}\,(U_0, v(t)) \;=\; \alpha(t)\;(v(t) - A(U_0))\;,$$

où la fonction α s'écrit :

$$\alpha(t) \;:=\; \partial_v \mathcal{R}\big(U_0;\, \sigma\, v(t) + (1-\sigma)\, A(U_0)\big)\,,$$

avec le paramètre σ compris entre 0 et 1. Nous pouvons maintenant écrire la formule de Duhamel pour le système (B.2.9) -(B.2.10) ; elle donne, pour tout $t > 0$: pour tout $t > 0$:

$$\left\{ \begin{array}{l} u(t) \;=\; U_0\,, \\ \\ v(t) \;=\; \dfrac{1}{\varepsilon}\,A(U_0) \displaystyle\int_0^t e^{-\int_\tau^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s}\,\alpha(\tau)\,\mathrm{d}\tau\,. \end{array} \right. \tag{B.2.11}$$

Ainsi, grâce à l'hypothèse (B.1.6) sur le gradient de $\mathcal{R}$, la fonction α reste bornée pour tout $t > 0$ (quel que soit le paramètre σ) comme suit :

$$\beta_0(U_0) \;\leq\; \alpha(t) \;\leq\; h(U_0)\,. \tag{B.2.12}$$

Choisissons donc la valeur de R_0, donc celle de $U_0 = \frac{1}{\sqrt{a}}\,R_0$. Comme les conditions initiales (B.1.2) de notre problème de Cauchy vérifient

$$|u_0(x)|\,, \quad |v_0(x)| \;\leqslant\; N_0\,,$$

il vient pour les variables diagonales :

$$|w_0(x)|\,, \quad |z_0(x)| \;\leqslant\; (1+\sqrt{a})\,N_0\,.$$

Nous prenons donc :

$$R_0^{\pm} := \pm(1+\sqrt{a})\,N_0\,,$$

de sorte qu'en variable u cela donne, grâce aux notations introduites en (B.1.8) :

$$U_0^{\pm} := \pm U(N_0,a)\,,$$

avec le signe + pour la sur-solution $(\overline{w},\overline{z})$ et le signe − pour la sous-solution $(\underline{w},\underline{z})$. Reprenons maintenant (B.2.11) et majorons la solution du système différentiel (B.2.9)-(B.2.10) avec $R_0 = R_0^{\pm}$ indifféremment. Le choix de R_0 rend valide l'encadrement de la fonction α (B.2.12) ; il vient donc, pour tout $t \geqslant 0$:

$$\begin{aligned} |\overline{u}(t)|\,,\ |\underline{u}(t)| &\leqslant U(N_0,a) \leqslant U(N_0,a_0)\,,\\ |\overline{v}(t)|\,,\ |\underline{v}(t)| &\leqslant \frac{h\left(U(N_0,a_0)\right)}{\beta_0\left(U(N_0,a_0)\right)}\,F(N_0,a_0)\,. \end{aligned} \tag{B.2.13}$$

En repassant aux variables diagonales, nous obtenons pour la sur-solution :

$$|\overline{w}(t)|\,,\ |\overline{z}(t)| \leqslant \sqrt{a}\left(|\overline{u}| + \frac{1}{\sqrt{a}}\,|\overline{v}|\right) \leqslant \sqrt{a}\,V(N_0,a_0)\,, \tag{B.2.14}$$

où V est donné par (B.1.8) ; avec la même majoration pour la sous-solution. Néanmoins, si ces estimations de la solution de (B.2.7)-(B.2.8) sont vraies pour tout temps, la quasi-monotonie n'est pour l'instant assurée que jusqu'à T_η. Précisément, pour tout $t \leqslant T_\eta$, pour tout x dans $\mathbb{R}$, on a :

$$|w(t,x)|\,,\ |z(t,x)| \leqslant \sqrt{a}\,V(N_0,a_0)\,. \tag{B.2.15}$$

Supposons par l'absurde que $T_\eta < +\infty$. Alors, d'après le Théorème 14, la norme L^∞ de la solution de (B.2.1)-(B.2.4) doit donc exploser en temps fini, ce qui n'est pas possible à cause de (B.2.15). Donc $T_\eta = T^* = +\infty$, la condition sous-caractéristique (B.2.6) est vérifiée pour tout temps, et la majoration (B.2.5) également :

$$\left\|v^\varepsilon(t) \pm \sqrt{a}\,u^\varepsilon(t)\right\|_{L^\infty} \leqslant \sqrt{a}\,V(N_0,a_0) \quad \forall\, t>0. \tag{B.2.16}$$

Pour terminer la preuve, nous étendons le résultat précédent pour des données initiales plus générales en approchant dans $L^1_{loc}(\mathbb{R})$ les fonctions u_0 et v_0 par des fonctions régulières u_0^δ et v_0^δ à supports compacts, et en appliquant le Théorème 15. □

B.2.2 Estimations BV

Comme dans le cas discret, nous établissons maintenant des estimations uniformes en ε dans $BV(\mathbb{R})$ de la solution faible du problème de relaxation, l'objectif étant d'obtenir de la compacité en espace dans L^1_{loc} grâce au Théorème de Helly.

Proposition 19. *Sous les hypothèses et notations de la Proposition 18, Soient $\varepsilon > 0$ et $(u^\varepsilon, v^\varepsilon)$ et $(\tilde{u}^\varepsilon, \tilde{v}^\varepsilon)$ les deux solutions faibles globales de (B.1.1) associées respectivement aux*

conditions initiales $(u_0^\varepsilon, v_0^\varepsilon)$ *et* $(\tilde{u}_0^\varepsilon, \tilde{v}_0^\varepsilon)$*. Alors, pour tout intervalle* $[c,d]$ *de* $\mathbb{R}$*, pour presque tout* $t \geq 0$*, on a l'estimation :*

$$\int_c^d \Big(|u^\varepsilon(t,x) - \tilde{u}^\varepsilon(t,x)| + |v^\varepsilon(t,x) - \tilde{v}^\varepsilon(t,x)| \Big)\, dx$$

$$\leqslant \frac{(1+\sqrt{a})^2}{\sqrt{a}} \int_{c-\sqrt{a}\,t}^{d+\sqrt{a}\,t} \Big(|u_0^\varepsilon(x) - \tilde{u}_0^\varepsilon(x)| + |v_0^\varepsilon(x) - \tilde{v}_0^\varepsilon(x)| \Big)\, dx\,. \quad \text{(B.2.17)}$$

Démonstration. Nous omettons les exposants ε pour plus de clarté. En considérant le système sous sa formulation diagonale (B.2.1), nous pouvons appliquer le Théorème 15. Alors, pour tout intervalle $[c,d]$ de $\mathbb{R}$, pour presque tout $t \geq 0$:

$$\int_c^d \Big(|w(t) - \tilde{w}(t)| + |z(t) - \tilde{z}(t)| \Big)\, \mathrm{d}x \;\leq\; \int_{c-\sqrt{a}\,t}^{d+\sqrt{a}\,t} \Big(|w_0 - \tilde{w}_0| + |z_0 - \tilde{z}_0| \Big)\, \mathrm{d}x$$

$$+\; \frac{1}{\varepsilon} \int_0^t \int_{c-\sqrt{a}\,(t-s)}^{d+\sqrt{a}\,(t-s)} \Big[G(w,z) - G(\tilde{w},\tilde{z}) \Big] \Big(\operatorname{sgn}(w-\tilde{w}) - \operatorname{sgn}(z-\tilde{z}) \Big)\, \mathrm{d}x\, \mathrm{d}s\,.$$

Or le deuxième terme du membre de droite de cette inégalité est négatif en raison des propriétés de quasimonotonie du système. En effet, notons

$$E := \Big[G(w,z) - G(\tilde{w},\tilde{z}) \Big] \Big(\operatorname{sgn}(w-\tilde{w}) - \operatorname{sgn}(z-\tilde{z}) \Big)\,.$$

Alors, E peut s'écrire sous la forme :

$$\begin{aligned} E \;&=\; \Big(\operatorname{sgn}(w-\tilde{w}) - \operatorname{sgn}(z-\tilde{z}) \Big) \Bigg[(w-\tilde{w}) \int_0^1 \partial_w G\left(\theta w + (1-\theta)\tilde{w}; z\right)\, \mathrm{d}\theta \\ &+\; (z-\tilde{z}) \int_0^1 \partial_z G\left(\tilde{w}; \theta z + (1-\theta)\tilde{z}\right)\, \mathrm{d}\theta \Bigg]\,. \end{aligned}$$

Ainsi, si $w - \tilde{w}$ et $z - \tilde{z}$ sont du même signe, alors $E = 0$. Tandis que s'ils sont de signes opposés, on a :

$$E = 2\,|w-\tilde{w}| \int_0^1 \partial_w G\left(\theta w + (1-\theta)\tilde{w}; z\right)\, \mathrm{d}\theta \;- 2\,|z-\tilde{z}| \int_0^1 \partial_z G\left(\tilde{w}; \theta z + (1-\theta)\tilde{z}\right)\, \mathrm{d}\theta\,.$$

Ainsi, puisque $\partial_w G \leq 0$ et que $\partial_z G \geq 0$, on a bien toujours $E \leq 0$, et donc :

$$\int_c^d \Big(|w(t) - \tilde{w}(t)| + |z(t) - \tilde{z}(t)| \Big)\, \mathrm{d}x \;\leq\; \int_{c-\sqrt{a}\,t}^{d+\sqrt{a}\,t} \Big(|w_0 - \tilde{w}_0| + |z_0 - \tilde{z}_0| \Big)\, \mathrm{d}x\,. \quad \text{(B.2.18)}$$

On peut maintenant revenir aux variables u et v. Comme $u = -(w+z)/2\sqrt{a}$ et $v = (-w+z)/2$, on a d'une part :

$$|u-\tilde{u}| + |v-\tilde{v}| \;\leq\; \frac{1+\sqrt{a}}{2\sqrt{a}} \Big(|w-\tilde{w}| + |z-\tilde{z}| \Big) .$$

Et d'autre part :

$$|w_0-\tilde{w}_0| + |z_0-\tilde{z}_0| \;\leq\; 2\,(1+\sqrt{a}) \Big(|u_0-\tilde{u}_0| + |v_0-\tilde{v}_0| \Big) .$$

En injectant ces inégalités dans (B.2.18), on obtient l'estimation annoncée, ce qui termine la preuve. □

Ce résultat permet d'obtenir une estimation BV de la solution uniforme par rapport à ε, donnée dans le corollaire ci-après.

Corollaire 20. *(Estimations BV)*
Sous les hypothèses et notations de la Proposition 18, Soient $\varepsilon > 0$ et $(u^\varepsilon, v^\varepsilon)$ la solution faible du problème de Cauchy (B.1.1)-(B.1.2). Alors il existe une constante $C > 0$, indépendante de ε, telle que, pour tout intervalle $[c,d]$ de $\mathbb{R}$, pour tout $t \geq 0$:

$$\|(u^\varepsilon(\cdot,t); v^\varepsilon(\cdot,t))\|_{BV(c,d)} \;\leq\; C\; \|(u_0^\varepsilon; v_0^\varepsilon)\|_{BV(c-\sqrt{a}\,t, d+\sqrt{a}\,t)} \;. \tag{B.2.19}$$

Démonstration. Nous omettons encore les exposants ε dans la preuve. Nous appliquons la Proposition 19 à des cas particulier de solutions faibles u, $\tilde{u}$, v, $\tilde{v}$ de (B.1.1) comme suit. Si $u_0\, v_0 \in BV$, alors on peut choisir, pour h assez petit :

$$\begin{cases} \tilde{u}_0(x) := u_0(x+h)\,, \qquad \tilde{v}_0(x) := v_0(x+h)\,, \\ \\ \tilde{u}(t,x) := u(t,x+h)\,, \quad \tilde{v}(t,x) := v(t,x+h)\,. \end{cases}$$

On obtient ainsi l'estimation annoncée, avec $C = \dfrac{(1+\sqrt{a})^2}{\sqrt{a}}$. □

Ainsi, comme la solution est bornée dans $BV(\mathbb{R})$ pour tout temps $t \geq 0$, le Théorème de Helly (voir par exemple [45]) assure que la suite $(u^\varepsilon(t,\cdot), v^\varepsilon(t,\cdot))_{\varepsilon>0}$ reste dans un compact de $L^1_{loc}(\mathbb{R})^2$ pour tout $t \geq 0$. En vue d'appliquer le Théorème d'Ascoli pour passer à la limite, nous aurons besoin de l'équicontinuité en temps, c'est l'objet du paragraphe suivant.

B.2.3 Equicontinuité en temps

Dans cette section, pour établir les résultats d'équicontinuité en temps, uniformément relativement à ε, on utilise essentiellement le résultat suivant, originellement dû à Kružkov [132].

Lemme 21. *(Condition suffisante d'équicontinuité)*
Soit g un fonction mesurable bornée, définie sur $(-R-h_0; R+h_0)\times[0,T]$, avec T, R, $h_0>0$. On suppose qu'il existe une fonction $\omega_R\in\mathcal{C}([0,h_0])$, non décroissante, avec $\omega_R(0)=0$, vérifiant, pour tout $t\in(0,T)$, pour tout $|h|<h_0$:

$$\int_{-R-h_0}^{R+h_0}|g(t,x+h)-g(t,x)|\,dx\;\leq\;\omega_R(|h|)\,. \tag{B.2.20}$$

On suppose également que, pour tous t, $t+\tau\in(0,T)\,(\tau>0)$, pour toute fonction $\phi\in\mathcal{C}^2([-R,R])$, on a

$$\left|\int_{-R}^{R}(g(t+\tau,x)-g(t,x))\,\phi(x)\,dx\right|\;\leq\;C_R\,\tau\,\|\phi\|_{\mathcal{C}^2}\,. \tag{B.2.21}$$

Alors, pour tous t, $t+\tau\in(0,T)\,(\tau>0)$, on a :

$$\int_{-R}^{R}|g(t+\tau,x)-g(t,x)|\;dx\;\leq\;\tilde{\omega}_R(\tau)\,, \tag{B.2.22}$$

où

$$\tilde{\omega}_R(\tau)=C_R\,\min\left\{|h|+\omega_R(|h|)+\frac{\tau}{h^2}\,;\;|h|\leq h_0\right\}.$$

Ce lemme permet d'établir d'abord l'équicontinuité en temps de u^ε.

Proposition 22. *(Equicontinuité de u)*
Sous les hypothèses et notations de la Proposition 18, alors, pour tout intervalle (c,d) de $\mathbb{R}$, pour tout $T>0$, il existe une fonction continue non décroissante $\omega\in\mathcal{C}([0,T])$, indépendante de ε, avec $\omega(0)=0$, telle que, pour tout $0\leq t\leq t+\tau\leq T$:

$$\int_c^d|u^\varepsilon(t+\tau,x)-u^\varepsilon(t,x)|\,dx\;\leq\;\omega(\tau)\,.$$

Démonstration. On omet l'indice ε. Afin d'appliquer le Lemme 21 à la fonction u, nous nous assurons qu'elle en vérifie bien les hypothèses. L'inégalité (B.2.20) découle du Corollaire 20. Il suffit donc de montrer l'inégalité (B.2.21). Cette inégalité se démontre exactement comme dans le cas de la relaxation semi-linéaire [152]. On se donne une fonction test ϕ à support compact (pour supprimer les termes de bords), et on calcule :

$$\begin{aligned}\left|\int_c^d(u(t+\tau,x)-u(t,x))\,\phi(x)\,\mathrm{d}x\right| &= \left|\int_c^d\left(\int_t^{t+\tau}\partial_t u(s,x)\,\mathrm{d}s\right)\,\phi(x)\,\mathrm{d}x\right|\\ &= \left|\int_c^d\left(\int_t^{t+\tau}v(s,x)\,\mathrm{d}s\right)\,\partial_x\phi(x)\,\mathrm{d}x\right|\\ &\leq \tau\,(d-c)\,\sqrt{a}\,V(N_0,a_0)\,\|\phi\|_{\mathcal{C}^1}\,,\end{aligned}$$

où la dernière inégalité vient de l'estimation L^∞ de v donnée par la Proposition 18. Ainsi, les hypothèses du Lemme 21 sont satisfaites et la famille (u^ε) est bien équicontinue en temps. □

L'équicontinuité en temps de v^ε est légèrement plus délicate.

Proposition 23. *(Equicontinuité de v)*
Sous les mêmes hypothèses que précédemment, alors pour tout $(c,d) \subset \mathbb{R}$, pour tous $0 < \nu < T$, il existe une fonction continue non décroissante $\omega^\nu \in \mathcal{C}([0, T-\nu))$, indépendante de ε, avec $\omega^\nu(0) = 0$, telle que, pour tous $\nu \leq t \leq t+\tau \leq T$:

$$\int_c^d |v^\varepsilon(t+\tau,x) - v^\varepsilon(t,x)|\, dx \;\leq\; \omega^\nu(\tau)\,.$$

Démonstration. On omet l'indice ε. Comme pour u, il suffit de montrer l'inégalité (B.2.21) afin d'utiliser le Lemme 21. De la même manière que précédemment, nous écrivons la formule de Duhamel pour v, en utilisant le développement de Taylor de $\mathcal{R}$:

$$v(t,x) \;=\; v_0(x)\, e^{-\int_0^t \frac{\alpha(\lambda)}{\varepsilon}\,\mathrm{d}\lambda} + \int_0^t e^{-\int_\lambda^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s} \Big(\frac{1}{\varepsilon}\,\alpha(\lambda)\, A(u(\lambda)) - a\,\partial_x u(\lambda)\Big)\,\mathrm{d}\lambda\,,$$

où la fonction α s'écrit :

$$\alpha(t) \;:=\; \partial_v \mathcal{R}\big(u(t);\, \sigma\, v(t) + (1-\sigma)\, A(u(t))\big)\,,$$

avec le paramètre σ compris entre 0 et 1. Décomposons maintenant la différence $(v(t+\tau,x) - v(t,x))$ comme suit :

$$\begin{aligned}
v(t+\tau,x) - v(t,x) \;&=\; v_0(x)\left(e^{-\int_0^{t+\tau} \frac{\alpha(\lambda)}{\varepsilon}\,\mathrm{d}\lambda} - e^{-\int_0^t \frac{\alpha(\lambda)}{\varepsilon}\,\mathrm{d}\lambda}\right) \\
&+\; \frac{1}{\varepsilon}\int_0^t e^{-\int_\lambda^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s} \Big(\alpha(\lambda+\tau)\, A\big(u(\lambda+\tau)\big) - \alpha(\lambda)\, A\big(u(\lambda)\big)\Big)\,\mathrm{d}\lambda \\
&-\; a\int_0^t e^{-\int_\lambda^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s} \Big(\partial_x u(\lambda+\tau) - \partial_x u(\lambda)\Big)\,\mathrm{d}\lambda \\
&\quad + \int_{-\tau}^0 \Big(\frac{1}{\varepsilon}\,\alpha(\lambda+\tau)\, A\big(u(\lambda+\tau)\big) - a\,\partial_x u(\lambda+\tau)\Big)\, e^{-\int_\lambda^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s}\,\mathrm{d}\lambda\,.
\end{aligned}$$

Dans la suite, nous noterons $\mathcal{E}$ la fonction définie par :

$$\mathcal{E}(\lambda,t) \;:=\; e^{-\int_\lambda^t \frac{\alpha(s)}{\varepsilon}\,\mathrm{d}s}\,.$$

Pour montrer l'inégalité (B.2.21), il nous faut multiplier par une fonction test positive ϕ et intégrer en espace : nous introduisons donc la quantité à estimer

$$W(t) := \left|\int_c^d (v(t+\tau,x) - v(t,x))\,\phi(x)\,\mathrm{d}x\right|,$$

et la majorons comme suit :

$$W(t) \;\leq J_1 + J_2 + J_3 + J_4\,,$$

avec

$$\left\{\begin{array}{rcl} J_1 &=& \displaystyle\int_c^d |v_0(x)| \Big|\mathcal{E}(0,t+\tau)-\mathcal{E}(0,t)\Big|\,\phi(x)\,\mathrm{d}x\,, \\ J_2 &=& \displaystyle\frac{1}{\varepsilon}\int_c^d \left|\int_0^t \mathcal{E}(\lambda,t)\Big(\alpha(\lambda+\tau)\,A\big(u(\lambda+\tau)\big)-\alpha(\lambda)\,A\big(u(\lambda)\big)\Big)\,\mathrm{d}\lambda\right|\phi(x)\,\mathrm{d}x\,, \\ J_3 &=& \displaystyle a\left|\int_c^d\int_0^t \mathcal{E}(\lambda,t)\Big(u(\lambda+\tau)-u(\lambda)\Big)\,\partial_x\phi(x)\,\mathrm{d}\lambda\,\mathrm{d}x\right|, \\ J_4 &=& \displaystyle\int_c^d \left|\int_{-\tau}^0 \Big(\frac{1}{\varepsilon}\,\alpha(\lambda+\tau)\,A\big(u(\lambda+\tau)\big)-a\,\partial_x u(\lambda+\tau)\Big)\,\mathcal{E}(\lambda,t)\,\mathrm{d}\lambda\right|\phi(x)\,\mathrm{d}x\,. \end{array}\right.$$

Remarquons maintenant que les bornes du gradient de $\mathcal{R}$ (B.1.6) nous donnent, pour tous $0<\nu\le\lambda\le t$, l'encadrement suivant :

$$e^{-\frac{\beta}{\varepsilon}(t-\lambda)} \;\le\; \mathcal{E}(\lambda,t) \;\le\; e^{-\frac{\beta_0}{\varepsilon}(t-\lambda)}\,.$$

Par ailleurs, pour tous $0<\nu\le\lambda\le t_i$ ($i=1$ ou 2), nous avons :

$$\begin{aligned}\mathcal{E}(\lambda,t_1)-\mathcal{E}(\lambda,t_2) &= \partial_t\mathcal{E}(\lambda,t^*)\,(t_1-t_2) \\ &= -\frac{\alpha(t^*)}{\varepsilon}\,\mathcal{E}(\lambda,t^*)\,(t_1-t_2)\,.\end{aligned}$$

Nous pouvons alors majorer les J_i, pour $i=1,\dots,4$. D'abord, pour tous $0<\nu\le t$:

$$\begin{aligned}J_1 &\le (d-c)\,\|v_0\|_\infty\,\frac{\beta\,\tau}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,\|\phi\|_{\mathcal{C}^0} \\ &\le C_1\,\frac{\tau}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,.\end{aligned}$$

Ensuite, nous majorons le deuxième en utilisant l'équicontinuité de u ainsi que la propriété Lipschitz de A :

$$\begin{aligned}J_2 &\le (d-c)\,\frac{\beta\,\tau}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,\|\phi\|_{\mathcal{C}^0}\,\omega(\tau) \\ &\le C_2\,\frac{\omega(\tau)}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,.\end{aligned}$$

Il en est de même pour J_3 :

$$\begin{aligned}J_3 &\le a\,\frac{\varepsilon}{\beta_0}\,\|\phi\|_{\mathcal{C}^1}\,\omega(\tau) \\ &\le C_3\,\varepsilon\,\omega(\tau)\,.\end{aligned}$$

Enfin, la majoration de J_4 utilise simplement les bornes L^∞ et BV de la solution, puisque la longueur de l'intervalle d'intégration est égale à τ :

$$\begin{aligned} J_4 &\leq (d-c)\,\frac{\beta\,\tau}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,\|\phi\|_{\mathcal{C}^1}\;\left(F(V(N_0,a_0))+a\,\|u\|_{BV}\right) \\ &\leq C_4\,\frac{\tau}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}\,. \end{aligned}$$

Pour conclure, il suffit de remarquer que la fonction

$$\varepsilon \longmapsto \varepsilon + \frac{1}{\varepsilon}\,e^{-\frac{\beta_0}{\varepsilon}\,\nu}$$

est bornée sur $\mathbb{R}^+$. □

B.2.4 Déviation par rapport à l'équilibre

Nous aurons également besoin, comme dans le cas discret, d'évaluer la déviation par rapport à l'équilibre en norme L^1.

Proposition 24. *(Déviation par rapport à l'équilibre)*

Sous les mêmes hypothèses que précédemment, alors pour tout $(c,d) \subset \mathbb{R}$, pour tout $t>0$, on a :

$$\begin{aligned} \int_c^d |v^\varepsilon(t,x) - A\,(u^\varepsilon(t,x))|\;dx &\leq C_1\,e^{-\frac{\beta t}{\varepsilon}}\;\int_c^d |v_0^\varepsilon - A\,(u_0^\varepsilon)|\;dx \\ &+ \varepsilon\,C_2\;\|(u_0^\varepsilon;v_0^\varepsilon)\|_{BV(c-\sqrt{a}\,t,d+\sqrt{a}\,t)}\;, \end{aligned} \tag{B.2.23}$$

où les constantes C_1 et C_2 sont indépendantes de ε.

Démonstration. On omet les ε. Considérons d'abord des conditions initiales régulières, et définissons la fonction déviation $\delta = v - A(u)$. Elle est solution de l'équation suivante :

$$\partial_t\delta + \frac{\beta}{\varepsilon}\,\delta \;=\; \frac{1}{\varepsilon}\;[\beta\,\delta - \mathcal{R}(u,v)] + A'(u)\,\partial_x v - a\,\partial_x u\,.$$

Ainsi, en appliquant la formule de Duhamel et en développant la fonction $\mathcal{R}$ à l'ordre 1 entre (u,v) et $(u,A(u))$, il vient, pour tout $t>0$:

$$\begin{aligned} \delta(t) &= \delta_0\,e^{-\frac{\beta\,t}{\varepsilon}} + \frac{1}{\varepsilon}\int_0^t (\beta - \partial_v\mathcal{R})\;\delta(s)\,e^{-\frac{\beta\,(t-s)}{\varepsilon}}\,\mathrm{d}s \\ &- a\int_0^t \partial_x u(s)\,e^{-\frac{\beta\,(t-s)}{\varepsilon}}\,\mathrm{d}s \;+\; \int_0^t A'\,(u(s))\;\partial_x v(s)\,e^{-\frac{\beta\,(t-s)}{\varepsilon}}\,\mathrm{d}s\,, \end{aligned}$$

où la dérivée partielle de $\mathcal{R}$ est calculée en un point $(u(s), \xi)$, avec $\xi \in (v(s);\, A(u(s)))$. Ensuite, en passant au module et en intégrant en x sur l'intervalle $[c, d]$, nous obtenons la majoration :

$$\begin{aligned}\int_c^d |\delta(t,x)|\,\mathrm{d}x \;&\leq\; e^{-\frac{\beta t}{\varepsilon}} \int_c^d |\delta_0(x)|\,\mathrm{d}x + \int_0^t \left(\int_c^d |\delta(s,x)|\,\mathrm{d}x\right) \frac{(\beta-\beta_0)}{\varepsilon}\, e^{-\frac{\beta(t-s)}{\varepsilon}}\,\mathrm{d}s \\ &+\; \sqrt{a}\int_0^t \left(\int_c^d \left(\sqrt{a}\,|\partial_x u(s,x)| + |\partial_x v(s,x)|\right)\mathrm{d}x\right) e^{-\frac{\beta(t-s)}{\varepsilon}}\,\mathrm{d}s\,.\end{aligned}$$

Enfin, nous concluons en appliquant le Lemme de Gronwall et en utilisant l'estimation BV obtenue au Corollaire 20. Cela entraîne le résultat pour des données initiales plus générales à variations bornées et termine la preuve. □

B.3 Convergence forte

Théorème 25. *(Convergence vers l'équilibre local)*

Sous les hypothèses précédentes, considérons $(u^\varepsilon, v^\varepsilon)$ la solution faible globale du problème de Cauchy (B.1.1)-(B.1.2) donnée par la Proposition 18. Alors il existe une solution faible $\overline{u}$ du problème de l'équilibre (B.1.4) avec la condition initiale donnée par (B.1.5), ainsi qu'une sous-suite encore notée $(u^\varepsilon, v^\varepsilon)$, telles que, pour tout $\nu > 0$:

$$u^\varepsilon \;\to\; \overline{u} \;\; \textit{dans} \;\; \mathcal{C}\left([0,\infty[;\, L^1_{loc}(\mathbb{R})\right)\,, \tag{B.3.1}$$

$$v^\varepsilon \;\to\; A(\overline{u}) \;\; \textit{dans} \;\; \mathcal{C}\left([\nu,\infty[;\, L^1_{loc}(\mathbb{R})\right)\,, \tag{B.3.2}$$

lorsque $\varepsilon \to 0_+$.

Démonstration. La solution $(u^\varepsilon, v^\varepsilon)$ considérée appartient à l'espace $\mathcal{C}\left(0, T; L^1_{loc}(\mathbb{R})\right)$ pour tout $T > 0$. D'après les estimations obtenues à la section précédentes, uniformes en ε, les familles $(u^\varepsilon)_{\varepsilon>0}$ et $(v^\varepsilon)_{\varepsilon>0}$ sont bornées dans $L^\infty \cap BV(\mathbb{R})$ pour presque tout $t > 0$. Donc par le théorème de Helly elles sont relativement compactes dans $L^1_{loc}(\mathbb{R})$. En outre, les résultats de la section précédente assurent qu'elles sont uniformément équicontinues en temps, sur $(0,\infty)$ pour u^ε, et sur (ν,∞) pour v^ε $(\forall\, \nu > 0)$.
Ainsi, pour tout $T > 0$, par le Théorème d'Ascoli, on peut en extraire des sous-suites convergentes : dans $\mathcal{C}\left(0, T; L^1_{loc}\right)$ pour u^ε et dans $\mathcal{C}\left(\nu, T; L^1_{loc}\right)$ pour v^ε. Précisément, il existe $\overline{u}$ et $\overline{v}$ telles que :

$$u^\varepsilon \;\to\; \overline{u}\,, \quad v^\varepsilon \;\to\; \overline{v}\,.$$

Enfin, en utilisant l'estimation uniforme en ε (B.2.23) (de la Proposition 24) de la déviation par rapport à l'équilibre, ainsi que la propriété de la fonction A (continue, localement lipschitzienne) nous obtenons, par unicité de la limite :

$$\overline{v} \;=\; A\left(\overline{u}\right)\,.$$

□

Troisième partie

Un modèle d'écoulement sanguin dans des artères avec stents

Chapitre 5

Asymptotic analysis of blood flow in stented arteries : time dependency

This work aims to extend results recently obtained in [148]. We will focus on the possible extension of our results to the time dependent case. So we consider the time dependent rough problem for a simplified heat equation in a straight channel that mimics the axial velocity under an oscillating pressure gradient. We derive first order approximations with respect to ε, the size of the roughness. In order to understand the problem and set up correct boundary layer approximations, we perform a time periodic Fourier analysis and check that no frequency can interact with the roughness. We show rigorously on this toy problem that the boundary layers remain stationary in time (independent on the frequency number). Finally we perform numerical tests validating our theoretical approach.

5.1 Introduction

Rupture of aneurysm are common lethal pathologies in western countries. It is mainly due to a loss of elastic properties of tissues that constitutes the arterial walls on some branching. Recently emerged a new kind of stent: a metallic wired mutli-layered prosthesis (see fig. 5.1 right) that unlike the classical endograft stent need not to be sutured to the arterial wall. Their form-memory metallic structure allow self-expansion recovering the original form without a need of a balloon.

A recent work of the first author establishes, thanks to asymptotic analysis tools, several advantages of this new device [148]. These can be summarized as follows :

- the presence of a stent at the inlet of a collateral artery (see fig. 5.1 left) gives rise to a secondary flow explicitly computable : it depends on the pressure jump occurring at zero order (when the stent totally closes the inlet of the collateral artery) and on some periodic microscopic resistivity (computed independently of any kind of macroscopic flow).
- the presence of a stent above a closed aneurysmal sac (see fig. 5.1 middle) imposes a

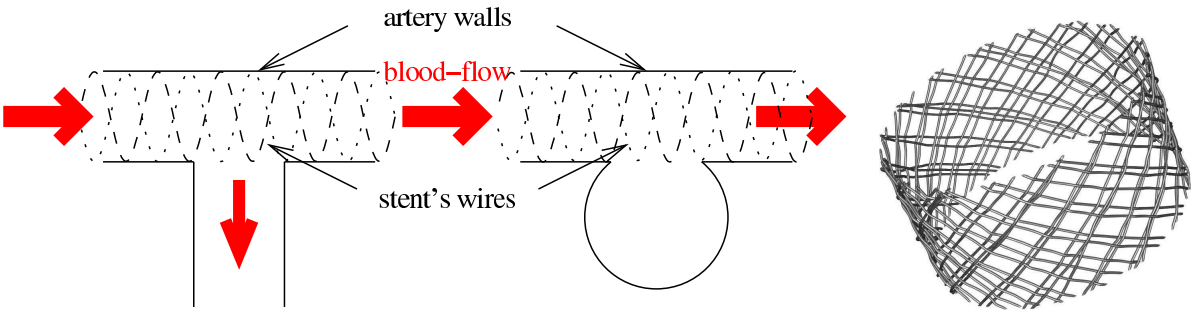


Figure 5.1: A sketch of stented arteries: with a collateral artery (left), an aneurysmal sac (middle) and a 3D example of a real metallic multi-wired stent (right)

constant averaged pressure inside the sac, it also inverts the direction of rotation of the vortex running inside the sac: without a stent, the cavity is driven by the mean flow in the artery, the vortex is tangential to the mean flow whereas the pressure jump across the interface imposes an entering velocity profile upward the sac and an outgoing profile downward the middle of the sac.

These results were established theoretically and numerically for the steady Stokes system of equations. In this work, we set up a preliminary toy framework in order to extend those results to the unsteady case.

Although this is not a first attempt to consider the unsteady regime within the boundary layer framework (let's mention [113, 67]), we set up here very basic model for the time periodic case. In the context of blood flow this regime is quite well-suited since the heart delivers a periodic pressure flow impulse to the cardio-vascular system. Another advantage of this work is that it is self-consistent: extending tools presented in [40], we provide self-contained proofs for every step of our approximation process. We give, for instance, a direct proof for time periodic very weak solutions.

The chapter is organized as follows: in Section 5.2 we give the basic notations and hypotheses of this work, in the next Section we perform a time Fourier expansion and we construct a boundary layer approximation. Then we show that an averaged approximation, cheaper from the computational point of view, is possible. At each step we provide theoretical error estimates wrt the direct rough solution. An interesting feature of the wall-law is exhibited: we show that although we recover the standard $\varepsilon^{3/2}$ convergence rate in $L^2(\Omega_0)$ norm, the *a priori* estimates provide only $\varepsilon^{1/2}$ rate performing a similar error as the zero order estimate itself. Section 5.4 is devoted to the derivation of an implicit macroscopic wall-law in the smooth domain. Finally, Section 5.5 validates numerically theoretical claims stated and proved in previous sections. The poor $H^1(\Omega_0)$ error is observed also on the numerical side.

5.2 Notations and problem setting

In this work, Ω^ε denotes the rough domain in $\mathbb{R}^2$ depicted in fig. 5.2, Ω_0 denotes the smooth one, and $\Omega_\varepsilon \setminus \Omega_0$ the complementary rough sub-domain. Γ_ε is the rough boundary and Γ_0 (resp. Γ_1) the lower (resp. upper) smooth one (see fig. 5.2).

Hypotheses 5.2.1. The rough boundary Γ_ε is described as a periodic repetition at the microscopic scale of a single boundary cell P_0. The latter can be parametrized as the graph of a Lipschitz function $f : [0, 2\pi[\to [-1 : 0[$ such that

$$P_0 = \{y \in [0, 2\pi] \times [-1 : 0[\text{ s.t. } y_2 = f(y_1)\}. \tag{5.2.1}$$

Moreover, we suppose that f is negative definite, i.e. there exists a positive constant δ such that $f(y_1) < \delta$ for all $y_1 \in [0, 2\pi]$. Then the macroscopic boundary Γ_ε is parametrized as

$$\Gamma_\varepsilon = \left\{ x \in \mathbb{R}^2 \text{ s.t. } x_2 = \varepsilon f\left(\frac{x}{\varepsilon}\right) \right\}.$$

We assume that the ratio between L (the width of Ω_0) and $2\pi\varepsilon$ (the width of the periodic cell) is always an integer called N. We consider a simplified setting that avoids the

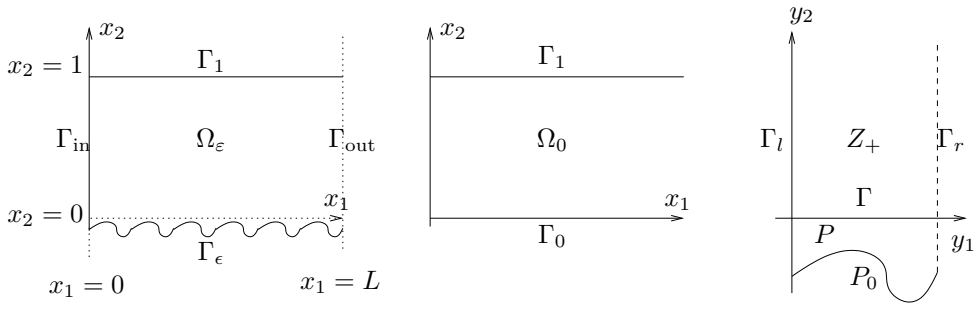


Figure 5.2: *Rough, smooth and cell domains*

theoretical difficulties and the non-linear complications of the full time-dependent Navier-Stokes equations. Starting from the time-dependent Stokes system, we consider a heat-like simplified problem for u_ε, the axial component of the velocity. The oscillating pressure gradient is assumed to reduce to a time-periodic space-constant right hand side $C(t)$. For sake of conciseness, we consider only periodic inflow and outflow boundary conditions on u_ε. The simplified problem reads : find u_ε such that

$$\begin{cases} \partial_t u_\varepsilon - \Delta u_\varepsilon = C(t), & \text{for } x \in \Omega_\varepsilon, \\ u_\varepsilon = 0, & \text{for } x \in \Gamma_\varepsilon \cup \Gamma_1, \\ u_\varepsilon \text{ is } x_1 \text{ periodic.} \end{cases} \tag{5.2.2}$$

We underline that the results below can be directly extended to rough domains with smooth holes and to the Stokes system in the case of a simple sheared flow.

In what follows, functions that do depend on $y = x/\varepsilon$ should be indexed by an ε (e.g. $\hat{\mathcal{U}}_{\epsilon,k} = \hat{\mathcal{U}}_{\epsilon,k}(x, x/\varepsilon)$).

5.3 Time Fourier analysis and boundary layer approximations

Applying the time-Fourier transform on (5.2.2) one obtains for each frequency-mode $k \in \mathbb{Z}^*$ the problem: find $\hat{u}_{\epsilon,k}$ s.t.

$$\begin{cases} (ik - \Delta)\,\hat{u}_{\epsilon,k} = \hat{C}_k & \text{in } \Omega_\varepsilon\,, \\ \hat{u}_{\epsilon,k} = 0 & \text{on } \Gamma_\varepsilon \cup \Gamma_1\,, \\ \hat{u}_{\epsilon,k}\ x_1 - \text{periodic on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{5.3.1}$$

where $\hat{C}_k$ is the Fourier mode associated to the frequency $k \in \mathbb{Z}$:

$$\hat{C}_k := \frac{1}{2\pi}\int_0^{2\pi} C(t)e^{ikt}dt, \quad C(t) = \sum_k \hat{C}_k e^{-ikt}.$$

For the rest of the chapter, one denotes $\mathcal{L}_k := (ik - \Delta)$. When $k \equiv 0$ one returns to the steady case already extensively studied in [40], so we only consider $k \in \mathbb{Z}^*$ for the rest of this chapter.

5.3.1 The zero order approximation

Passing to the limit formally wrt ε in (5.3.1), one shows rigorously below that actually $\hat{u}_{\epsilon,k}$ converges to $\hat{u}_{0,k}$ solving

$$\begin{cases} \mathcal{L}_k\,\hat{u}_{0,k} = \hat{C}_k & \text{in } \Omega_0\,, \\ \hat{u}_{0,k} = 0 & \text{on } \Gamma_0 \cup \Gamma_1\,, \\ \hat{u}_{0,k}\,x_1 - \text{periodic on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{5.3.2}$$

The solution of this problem is explicit wrt to the data $\hat{C}_k$ and the frequency k, it reads for every $x \in \Omega_0$:

$$\hat{u}_{0,k} = \frac{\hat{C}_k}{ik}\left(1 + A\,e^{rx_2} + B\,e^{-rx_2}\right), \tag{5.3.3}$$

where

$$r := \frac{\sqrt{2k}}{2}(1+i), \quad A := \frac{e^{-r}-1}{e^r - e^{-r}}, \quad B := \frac{1-e^r}{e^r - e^{-r}}\,.$$

In order to estimate the error made when we consider the solution $\hat{u}_{0,k}$ as an approximation of $\hat{u}_{\epsilon,k}$, we have to extend $\hat{u}_{0,k}$ to the whole rough domain Ω_ε. It suffices that it is continuous, since we need H^1 functions for *a priori* error estimates. In the literature, either the solution is extended by a constant in the rough layer [117] or one constructs a linear extension using the Taylor expansion around the point $(x_1, 0)$ [4]. In order to correct these errors at the next order, in the first case one corrects then the jump of the derivative, and

in the second case one should lift the Dirichlet error [40]. Here, we chose to extend $\hat{u}_{0,k}$ by a linear function in $\Omega_\varepsilon \setminus \Omega_0$:

$$\hat{u}_{0,k} := \left\{ \begin{array}{ll} \hat{u}_{0,k} & \text{in } \Omega_0, \\ \mathcal{M}_k\, x_2 & \text{in } \Omega_\varepsilon \setminus \Omega_0, \end{array} \right. , \text{ where } \mathcal{M}_k := \frac{\partial}{\partial x_2}\hat{u}_{0,k}(x_1,0) = \frac{\hat{C}_k\, r\,(2-e^r-e^{-r})}{ik\,(e^{-r}-e^r)}.$$

5.3.2 Zero order error estimates

We detail here the error estimates. Identical proofs should also be used for higher order approximations below: we detail here every step. Denote χ_Ω the characteristic function of the domain Ω, δ_{Γ_0} the Dirac measure concentrated on Γ_0.

Proposition 1. There exist two positive constants c_1 and c_2, depending only of the mode $\hat{C}_k$ and the Sobolev's inequalities, such that:

$$\|\hat{u}_{\epsilon,k} - \hat{u}_{0,k}\|_{H^1(\Omega_\varepsilon)} \leq c_1\,\sqrt{\varepsilon}\ , \quad \|\hat{u}_{\epsilon,k} - \hat{u}_{0,k}\|_{L^2(\Omega_0)} \leq c_2\,\varepsilon\,. \tag{5.3.4}$$

Proof. The first part of the proof is based on standard *a priori* estimates. The existence and uniqueness of $\hat{u}_{\epsilon,k}$ are well known and derive from the Lax-Milgram theorem. We focus on the error, namely we set $R_0^\varepsilon := \hat{u}_{\epsilon,k} - \hat{u}_{0,k}$. Since the extension $\hat{u}_{0,k}$ of $\hat{u}_{0,k}$ in the rough domain satisfies:

$$\left\{ \begin{array}{ll} \mathcal{L}_k\ \hat{u}_{0,k} = \hat{C}_k\,\chi_{\Omega_0} + ik\mathcal{M}_k\,x_2\,\chi_{\Omega_\varepsilon\setminus\Omega_0} & \text{in } \Omega_\varepsilon\,, \\ \hat{u}_{0,k} = 0 & \text{on } \Gamma_1\,, \\ \hat{u}_{0,k} = \mathcal{M}_k\,x_2 & \text{on } \Gamma_\varepsilon\,. \end{array} \right. \tag{5.3.5}$$

Then the zeroth order error solves:

$$\left\{ \begin{array}{ll} \mathcal{L}_k\ R_0^\varepsilon = \hat{C}_k\,\chi_{\Omega_\varepsilon\setminus\Omega_0} - ik\mathcal{M}_k\,x_2\,\chi_{\Omega_\varepsilon\setminus\Omega_0} & \text{in } \Omega_\varepsilon\,, \\ R_0^\varepsilon = 0 & \text{on } \Gamma_1\,, \\ R_0^\varepsilon = -\mathcal{M}_k\,x_2 & \text{on } \Gamma_\varepsilon\,. \end{array} \right. \tag{5.3.6}$$

We remark that a part of the error comes from the source term localized in $\Omega_\varepsilon \setminus \Omega_0$, and another part comes from the non homogeneous boundary term on Γ_ε. We set the lift:

$$s = -\mathcal{M}_k\,x_2\,\chi_{\Omega_\varepsilon\setminus\Omega_0}\ , \quad \tilde{R}_0^\varepsilon = R_0^\varepsilon - s\,.$$

Then:

$$\left\{ \begin{array}{ll} \mathcal{L}_k\ \tilde{R}_0^\varepsilon = \hat{C}_k\,\chi_{\Omega_\varepsilon\setminus\Omega_0} + \mathcal{M}_k\,\delta_{\Gamma_0} & \text{in } \Omega_\varepsilon\,, \\ \tilde{R}_0^\varepsilon = 0 & \text{on } \Gamma_1\,, \\ \tilde{R}_0^\varepsilon = 0 & \text{on } \Gamma_\varepsilon\,, \end{array} \right. \tag{5.3.7}$$

where the derivatives are computed in the sense of distributions. Then, on the one hand, using Poincaré inequality, we have:

$$\begin{array}{rcl} \left| \displaystyle\int_{\Omega_\varepsilon} \mathcal{L}_k \tilde{R}_0^\varepsilon\,\overline{\tilde{R}_0^\varepsilon}\,dx \right|^2 & = & \left| ik\,\|\tilde{R}_0^\varepsilon\|^2_{L^2(\Omega_\varepsilon)} + \|\nabla\tilde{R}_0^\varepsilon\|^2_{L^2(\Omega_\varepsilon)} \right|^2 \\ & = & k^2\,\|\tilde{R}_0^\varepsilon\|^4_{L^2(\Omega_\varepsilon)} + \|\nabla\tilde{R}_0^\varepsilon\|^4_{L^2(\Omega_\varepsilon)} \\ & \geq & c\,\|\tilde{R}_0^\varepsilon\|^4_{H^1(\Omega_\varepsilon)}\,. \end{array} \tag{5.3.8}$$

On the other hand, for any test function $\phi \in H_0^1(\Omega_\varepsilon)$:

$$\int_{\Omega_\varepsilon} \mathcal{L}_k \tilde{R}_0^\varepsilon \, \phi \, dx = \hat{C}_k \int_{\Omega_\varepsilon \setminus \Omega_0} \phi \, dx + \mathcal{M}_k \int_{\Gamma_0} \phi \, dx \, .$$

Then, using Cauchy-Schwarz and Poincaré like inequalities, we obtain the upper bound:

$$\left| ik \, \|\tilde{R}_0^\varepsilon\|^2_{L^2(\Omega_\varepsilon)} + \|\nabla \tilde{R}_0^\varepsilon\|^2_{L^2(\Omega_\varepsilon)} \right| \leq c \, \left(|\hat{C}_k| \, \varepsilon + |\mathcal{M}_k| \, \sqrt{\varepsilon} \right) \, \|\tilde{R}_0^\varepsilon\|_{H^1(\Omega_\varepsilon)} \, , \tag{5.3.9}$$

where c is a non negative constant depending on the Poincaré inequality. And $|\mathcal{M}_k|$ is controlled as follows:

$$\begin{aligned} |\mathcal{M}_k| &= \frac{\left| \hat{C}_k r \, (2 - e^r - e^{-r}) \right|}{|ik \, (e^r - e^{-r})|} \\ &\leq \frac{|\hat{C}_k|}{\sqrt{k}} \left(\frac{2}{|e^r - e^{-r}|} + 1 \right) \\ &\leq 2 \, |\hat{C}_k| \, . \end{aligned} \tag{5.3.10}$$

Finally, combining (5.3.8)-(5.3.10), we get the H^1 -error estimate.

For the L^2 error, we use the concept of a *very weak solution*. Namely, one solves the dual problem: for a given $\phi \in L^2(\Omega_0)$, ϕ being x_1 periodic on $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$, find $\hat{v} \in H^2(\Omega_0)$ such that

$$\begin{cases} \overline{\mathcal{L}}_k \, \hat{v} = \phi & \text{in } \Omega_0 \, , \\ \hat{v} = 0 & \text{on } \Gamma_1 \cup \Gamma_0 \, , \\ \hat{v} \text{ is } x_1\text{-periodic} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}} \, . \end{cases} \tag{5.3.11}$$

Then, considering the $L^2(\Omega_0)$ scalar product $\big(. \, , .\big)$, and using the Green formula:

$$\begin{aligned} (R_0^\varepsilon \, , \, \phi) = \int_{\Omega_0} R_0^\varepsilon \, \overline{\mathcal{L}}_k \hat{v} &= -ik \int_{\Omega_0} R_0^\varepsilon \, \hat{v} + \int_{\Omega_0} \nabla R_0^\varepsilon \, \nabla \hat{v} - \int_{\partial\Omega_0} R_0^\varepsilon \, \frac{\partial \hat{v}}{\partial n} \\ &= \left\langle \hat{v} \, , \, \frac{\partial R_0^\varepsilon}{\partial n} \right\rangle_{\Gamma_{\text{in}} \cup \Gamma_{\text{out}}} - \left(R_0^\varepsilon \, , \, \frac{\partial \hat{v}}{\partial n} \right)_{\Gamma_0 \cup \Gamma_1} \, , \end{aligned} \tag{5.3.12}$$

where the brackets refer to the dual product in $\big(H^{-1} \, , \, H^1\big) \, (\partial\Omega_0)$ and the rest of the products are in L^2 either on Γ_0 or in Ω_0. Then one computes:

$$\begin{aligned} |(R_0^\varepsilon \, , \, \phi)| &\leq \|R_0^\varepsilon\|_{L^2(\Gamma_0)} \left\| \frac{\partial \hat{v}}{\partial n} \right\|_{L^2(\Gamma_0)} \\ &\leq \sqrt{\varepsilon} \, \|\nabla R_0^\varepsilon\|_{L^2(\Omega_\varepsilon \setminus \Omega_0)} \, \|\phi\|_{L^2(\Omega_0)} \\ &\leq \sqrt{\varepsilon} \, \|\nabla R_0^\varepsilon\|_{L^2(\Omega_\varepsilon)} \, \|\phi\|_{L^2(\Omega_0)} \\ &\leq \varepsilon^{\frac{3}{2}} \, \|\phi\|_{L^2(\Omega_0)} . \end{aligned}$$

This ends the proof of the proposition, by taking the sup over all $\phi \in L^2(\Omega_0)$. But between the first and the second estimate above, we assumed that the solutions of the regular problem (5.3.11) satisfy a kind of Rellich estimates (see [155], chap. 5) :

$$\left\| \frac{\partial \hat{v}}{\partial n} \right\|_{L^2(\Gamma_0)} \leq c \, \|\phi\|_{L^2(\Omega_0)} \, . \tag{5.3.13}$$

In order to prove this, we decompose ϕ on the Hilbert basis $(e^{2\pi i n x_1} \, e^{2\pi i m x_2})_{n,m}$ of $L^2(\Omega_0)$. Separating the variables, define $\phi_n(x_2)$ the coordinates of ϕ in the $(e^{2\pi i n x_1})_n$ Hilbert basis of $L^2(0,1)$, and $a_{n,m}$ its coordinates in the basis $(e^{2\pi i n x_1} \, e^{2\pi i m x_2})_{n,m}$. Then ϕ can be written under the form:

$$\phi = \sum_{n \in \mathbb{Z}} \phi_n(x_2) \, e^{2\pi i n x_1} = \sum_{n,m \in \mathbb{Z}} a_{n,m} \, e^{2\pi i n x_1} \, e^{2\pi i n x_2},$$

therefore:

$$\|\phi\|^2_{L^2(\Omega_0)} = \sum_{n,m \in \mathbb{Z}} |a_{n,m}|^2 \, .$$

In the same way, one can decompose $\hat{v}$ on the basis: $\hat{v} = \sum_{n \in \mathbb{Z}} \hat{v}_n(x_2) \, e^{2\pi i n x_1}$. Then the first equation of system (5.3.11) can be rewritten under the form of an infinite system of ordinary differential equations:

$$\forall l \in \mathbb{Z} \, , \quad (ik + 4\pi^2 l^2) \, \hat{v}_l - \hat{v}_l'' = \phi_l \, .$$

And the solution, for a fixed l, is given by:

$$\hat{v}_l = A \, e^{bx_2} + B \, e^{-bx_2} + \hat{v}_{p,l},$$

where $\hat{v}_{p,l}$ stands for the particular solution and reads

$$\hat{v}_{p,l} := \sum_{m \in \mathbb{Z}} \frac{-a_{l,m}}{4\pi^2 m^2 + b^2} \, e^{2\pi i m x_2},$$

while the constants satisfy:

$$\begin{cases} A - B = \tanh(b) \, \hat{v}_{p,l}(0) - \dfrac{\hat{v}_{p,l}(1)}{\sinh(b)}, \\ b^2 = 4\pi^2 l^2 + ik. \end{cases}$$

Then, since

$$\left\| \frac{\partial \hat{v}}{\partial n} \right\|^2_{L^2(\Gamma_0)} = \sum_{l \in \mathbb{Z}} \left| \hat{v}_l'(0) \right|^2 ,$$

it remains to estimate:

$$|\hat{v}_l'(0)|^2 = \left| b \, (A - B) + \hat{v}_{p,l}'(0) \right|^2 \leq 2 \left(|b(A-B)|^2 + |\hat{v}_{p,l}'(0)|^2 \right) \, . \tag{5.3.14}$$

The intermediate variable b solves in $\mathbb{C}$ the equation $b^2 = 4\pi^2 l^2 + ik$ which implies that

$$b := b_r + ib_i, \quad b_r := \pm\sqrt{2\pi^2 l^2 + \sqrt{4\pi^4 l^4 + k^2}}, \quad b_i := \pm\sqrt{-2\pi^2 l^2 + \sqrt{4\pi^4 l^4 + k^2}},$$

so that

$$|b|^2 = b_r^2 + b_i^2 = 2\sqrt{4\pi^4 l^4 + k^2} \geq 2, \quad \forall l \in \mathbb{Z}, \quad \forall k \in \mathbb{Z}^*$$

We return to the rhs of (5.3.14), the first term of the rhs can be split into two parts:

$$\begin{aligned} |b|^2 |\hat{v}_{p,l}(0)|^2 |\tanh(b)|^2 &\leq \left| \sum_m \frac{a_{m,l}}{4\pi^2 m^2 + b^2} \right|^2 |b|^2 |\tanh(b)|^2 \\ &\leq 2 \left(\sum_m |a_{m,l}|^2 \right) \left(\sum_m \frac{|b|^2}{4\pi^4 m^4 + |b|^4} \right) |\tanh(b)|^2. \end{aligned}$$

For sake of conciseness we set:

$$I := \left(\sum_{m \in \mathbb{Z}} \frac{|b|^2}{4\pi^4 m^4 + |b|^4} \right),$$

then it is equivalent to write

$$I = \frac{1}{|b|^2} + 2 \sum_{m \geq 1} \frac{|b|^2}{4\pi^4 m^4 + |b|^4} =: I_1 + I_2$$

If x is a positive real, we set $m := E[x]$ where $E[\cdot]$ is the integer part of its argument, one then has

$$\begin{aligned} I_2 &\leq \int_1^\infty \frac{|b|^2}{4\pi^4 (x-1)^4 + |b|^4} dx = \int_0^\infty \frac{|b|^2}{4\pi^4 x^4 + |b|^4} dx \\ &\leq \int_0^1 + \int_1^\infty \frac{|b|^2}{4\pi^4 x^4 + |b|^4} dx \\ &\leq c \left(\frac{1}{|b|^2} + \frac{1}{|b|} \right), \end{aligned}$$

so that finally $I \leq c$, the constant c being independent on either l or k. Because $b_r \neq 0$ one has that

$$\begin{aligned} |\tanh(b)| &= \frac{e^{2b_r} + e^{-2b_r} + 2\cos(b_i)}{e^{2b_r} + e^{-2b_r} - 2\cos(b_i)} \\ &\leq \left(4 + \sum_{q=1}^\infty \frac{(2b_r)^{2q}}{(2q)!} \right) / \left(\sum_{q=1}^\infty \frac{(2b_r)^{2q}}{(2q)!} \right) \\ &\leq c, \end{aligned}$$

where c does not depend on k nor on l. We now treat the second part of the first term of the rhs of (5.3.14): in a similar way one gets again using Cauchy-Schwartz

$$\begin{aligned}\frac{|\hat{v}_{p,l}(1)|^2|b|^2}{|\sinh(b)|^2} &\leq \sum_m |a_{m,l}|^2 \left(\sum_m \frac{|b|^2}{4\pi^4 m^4 + |b|^4}\right) \frac{1}{|\sinh(b)|^2} \\ &\leq c\,\|\phi_l\|^2_{L^2(0,1)},\end{aligned}$$

where again c is a generic constant independent on k, l. The estimates of $|\hat{v}'_{p,l}(0)|$ (last term of the rhs of (5.3.14)) follow the same lines. □

5.3.3 First order correction

We have already seen that the zeroth order approximation contains two distinct sources of errors: a part is due to the order of the extension $\hat{u}_{0,k}$ in $\Omega_\varepsilon \setminus \Omega_0$ and another part comes from the non homogeneous rest on Γ_ε. In order to correct the non zero value of $\hat{u}_{0,k}$ on the rough boundary Γ_ε, we introduce the corrector β, defined on the microscopic cell $Z^+ \cup \Gamma \cup P$

$$\begin{cases} \Delta\beta = 0 & \text{in } Z^+ \cup P\,, \\ \beta = -y_2 & \text{on } P_0\,, \\ \beta \text{ is } y_1\text{-periodic}\,. \end{cases} \tag{5.3.15}$$

We define the microscopic average along the fictitious interface Γ:

$$\overline{\beta} = \frac{1}{2\pi}\int_0^{2\pi} \beta(y_1, 0)\, dy_1\,.$$

The existence and uniqueness of β, and its properties, as the exponential convergence towards $\overline{\beta}$ when y_2 tends to infinity, are described in [40] and references therein. Because β tends to $\overline{\beta}$ when y_2 goes to infinity, we subtract this constant in the final asymptotic ansatz. As the constant should be relevant only far from the roughness we correct the ansatz by adding to $\hat{u}_{1,k}$ a "counter-flow" approximation solving:

$$\begin{cases} \mathcal{L}_k\, \hat{u}_{1,k} = 0 & \text{in } \Omega_0\,, \\ \hat{u}_{1,k} = 0 & \text{on } \Gamma_1\,, \\ \hat{u}_{1,k} = \overline{\beta}\mathcal{M}_k & \text{on } \Gamma_0\,, \\ \hat{u}_{1,k} \text{ is } x_1\text{-periodic} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{5.3.16}$$

The solution is again explicit:

$$\begin{aligned}\hat{u}_{1,k} &= \frac{-\overline{\beta}\mathcal{M}_k}{e^r - e^{-r}} \left(e^{-r}\, e^{rx_2} - e^r\, e^{-rx_2}\right) \\ &= \overline{\beta}\,\mathcal{M}_k\, \frac{\sinh(r(1+x_2))}{\sinh(r)}\,.\end{aligned} \tag{5.3.17}$$

Now we are in the position to define the full boundary layer approximation :

$$\hat{\mathcal{U}}_{\epsilon,k} := \hat{u}_{0,k} + \varepsilon\mathcal{M}_k \left(\beta\left(\frac{x}{\varepsilon}\right) - \overline{\beta}\right) + \varepsilon\hat{u}_{1,k}\,.$$

5.3.4 First order estimates

The gain obtained when introducing the microscopic corrector is of order $\sqrt{\varepsilon}$. Indeed, the following error estimates hold.

Proposition 2. There exist two positive constants c_3 and c_4, depending only on the mode $\hat{C}_k$ and not on the frequency k, such that:

$$\|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\|_{H^1(\Omega_\varepsilon)} \leq c_3\,\varepsilon \ , \quad \|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\|_{L^2(\Omega_0)} \leq c_4\,\varepsilon^{3/2}\,. \tag{5.3.18}$$

Proof. Denote $R_\varepsilon := \hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}$ the error to estimate. It is solution of the problem:

$$\begin{cases} \mathcal{L}_k\, R_\varepsilon = \hat{C}_k \chi_{\Omega_\varepsilon\setminus\Omega_0} - ik\mathcal{M}_k x_2 \chi_{\Omega_\varepsilon\setminus\Omega_0} - ik\mathcal{M}_k\varepsilon\left(\beta(\frac{x}{\varepsilon}) - \overline{\beta} + \overline{\beta}\chi_{\Omega_\varepsilon\setminus\Omega_0}\right) - \varepsilon\mathcal{M}_k\overline{\beta}\delta_{\Gamma_0}, & \text{in } \Omega_\varepsilon \\ R_\varepsilon = -\varepsilon\,\mathcal{M}_k\,\left(\beta(\frac{x_1}{\varepsilon},\frac{1}{\varepsilon}) - \overline{\beta}\right) & \text{on } \Gamma_1\,, \\ R_\varepsilon = 0 & \text{on } \Gamma_\varepsilon\,, \\ R_\varepsilon \text{ is } x_1\text{-periodic on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{5.3.19}$$

The existence and uniqueness of R_ε are standard. We focus again on the *a priori* estimates: test the system above by $\overline{R}_\varepsilon$ and estimate the lhs from below as in (5.3.8), then estimate from above the rhs. The last step includes new terms wrt the zeroth order approximation, listed below :

$$\begin{cases} A_1 \;=\; \hat{C}_k \displaystyle\int_{\Omega_\varepsilon\setminus\Omega_0} \overline{R}_\varepsilon\, dx, \\ A_2 \;=\; -ik\,\mathcal{M}_k \displaystyle\int_{\Omega_\varepsilon\setminus\Omega_0} x_2\,\overline{R}_\varepsilon\, dx, \\ A_3 \;=\; -ik\,\mathcal{M}_k\,\varepsilon \displaystyle\int_{\Omega_\varepsilon} (\beta(\frac{x}{\varepsilon}) - \overline{\beta})\,\overline{R}_\varepsilon\, dx\,, \\ A_4 \;=\; -ik\,\mathcal{M}_k\,\varepsilon\,\overline{\beta} \displaystyle\int_{\Omega_\varepsilon\setminus\Omega_0} \overline{R}_\varepsilon\, dx, \\ A_5 \;=\; -\varepsilon\,\overline{\beta}\,\mathcal{M}_k \displaystyle\int_{\Gamma_0} \overline{R}_\varepsilon\, dx_1. \end{cases} \tag{5.3.20}$$

Then, estimating these terms, one gets

$$|\sum_{j=1}^{5} A_j| \;\leq\; \varepsilon^{\frac{3}{2}}\, c\, \|\nabla R_\varepsilon\|_{L^2(\Omega_\varepsilon)}$$

which ends the proof for the *a priori* estimates. Again very weak estimates give:

$$\begin{aligned}\|R_\varepsilon\|_{L^2(\Omega_0)} &\leq \|R_\varepsilon\|_{L^2(\Gamma_1\cup\Gamma_0)} + \varepsilon|k\mathcal{M}_k|\left\|\beta\left(\frac{\cdot}{\varepsilon}\right)-\overline{\beta}\right\|_{L^2(\Omega_0)}\\ &\leq c\left(e^{-\frac{1}{\varepsilon}}+\sqrt{\varepsilon}\|\nabla R_\varepsilon\|_{L^2(\Omega_\varepsilon\setminus\Omega_0)}+\varepsilon^{\frac{3}{2}}\right)\\ &\leq c\,\varepsilon^{\frac{3}{2}}\,.\end{aligned}$$

□

5.4 Derivation of Wall-laws

5.4.1 Averaging the ansatz

We aim to derive a system of equations defined on the smooth domain Ω_0, for which the effect of the roughness is included as a macroscopic boundary condition on Γ_0. First, averaging wrt the fast variable in the horizontal direction, we get:

$$\overline{\hat{\mathcal{U}}_{\epsilon,k}} = \hat{u}_{0,k} + \varepsilon\hat{u}_{1,k} := \overline{u}_{\varepsilon,k}.$$

Though, the averaging process cancels the oscillations, the averaged ansatz still contains a first order macroscopic correction $\hat{u}_{1,k}$ accounting for averaged first order effects. This new averaged quantity solves a problem in the smooth limiting domain Ω_0:

$$\begin{cases}\mathcal{L}_k\overline{u}_{\varepsilon,k} = \hat{C}_k & \text{in } \Omega_0\,,\\ \overline{u}_{\varepsilon,k} = 0 & \text{on } \Gamma_1\,,\\ \overline{u}_{\varepsilon,k} = \varepsilon\,\mathcal{M}_k\,\overline{\beta} & \text{on } \Gamma_0\,,\\ \overline{u}_{\varepsilon,k} \text{ is } x_1\text{-periodic} & \text{on } \Gamma_{\text{in}}\cup\Gamma_{\text{out}}\,.\end{cases} \tag{5.4.1}$$

We compute the L^2-error estimate between the exact solution $\hat{u}_{\epsilon,k}$ of problem (5.3.1) and the averaged first order approximation $\overline{u}_{\varepsilon,k}$.

Proposition 3. There exists one positive constant c_5 , depending only of the mode $\hat{C}_k$ such that:

$$\|\hat{u}_{\epsilon,k}-\overline{u}_{\varepsilon,k}\|_{L^2(\Omega_0)} \leq c_5\,\varepsilon^{3/2}\,. \tag{5.4.2}$$

Proof. We write a triangular inequality:

$$\|\hat{u}_{\epsilon,k}-\overline{u}_{\varepsilon,k}\|_{L^2(\Omega_0)} \leq \|\hat{u}_{\epsilon,k}-\hat{\mathcal{U}}_{\epsilon,k}\|_{L^2(\Omega_0)} + \|\hat{\mathcal{U}}_{\epsilon,k}-\overline{u}_{\varepsilon,k}\|_{L^2(\Omega_0)}\,.$$

The second term in the rhs is explicit :

$$\hat{\mathcal{U}}_{\epsilon,k}-\overline{u}_{\varepsilon,k} \;=\; \varepsilon\,\mathcal{M}_k\,\left(\beta\left(\frac{x}{\varepsilon}\right)-\overline{\beta}\right)\,.$$

One thus estimates this quantity directly in the $L^2(\Omega_0)$ norm. Thanks to the multiscale structure of this corrector one gets by a simple change of variable and thanks to the specific boundary layer properties of β that

$$\left\|\beta\left(\frac{\cdot}{\varepsilon}\right)-\overline{\beta}\right\|_{L^2(\Omega_0)} \leq \sqrt{\varepsilon}\left\|\beta-\overline{\beta}\right\|_{L^2(Z^+\cup\Gamma\cup P)}$$

which ends the proof. □

5.4.2 Implicit wall-law

In order to derive an implicit wall-law, we rewrite the boundary condition satisfied by $\overline{u}_{\varepsilon,k}$ on Γ_0:

$$\begin{aligned}\overline{u}_{\varepsilon,k} \;=\; \varepsilon\,\mathcal{M}_k\,\overline{\beta} \;\; &= \varepsilon\,\overline{\beta}\,\frac{\partial}{\partial x_2}\left(\hat{u}_{0,k}+\varepsilon\hat{u}_{1,k}-\varepsilon\hat{u}_{1,k}\right)\\ &= \varepsilon\,\overline{\beta}\,\frac{\partial\overline{u}_{\varepsilon,k}}{\partial x_2}-\;\varepsilon^2\,\overline{\beta}\frac{\partial\hat{u}_{1,k}}{\partial x_2} \quad \text{ on } \Gamma_0\,.\end{aligned} \tag{5.4.3}$$

Hence, since the term $\partial\hat{u}_{1,k}/\partial x_2$ can be bounded independently from the frequency k, we derive a first order implicit wall-law. Indeed,

$$\frac{\partial\hat{u}_{1,k}}{\partial x_2}(x_2=0) \;=\; -\hat{C}_k\left(\frac{e^r+e^{-r}}{e^r-e^{-r}}\right)^2. \tag{5.4.4}$$

So, when $k\neq 1$:

$$\left|\frac{\partial\hat{u}_{1,k}}{\partial x_2}(x_2=0)\right| \leq |\hat{C}_k|\left(\frac{1}{1-e^{-\sqrt{2}}}\right)^2. \tag{5.4.5}$$

We set the following approximate problem, posed in the smooth domain Ω_0:

$$\begin{cases}\mathcal{L}_k\hat{V}_{\varepsilon,k}=\hat{C}_k & \text{in } \Omega_0\,,\\ \hat{V}_{\varepsilon,k}=0 & \text{on } \Gamma_1\,,\\ \hat{V}_{\varepsilon,k}=\varepsilon\,\overline{\beta}\,\frac{\partial\hat{V}_{\varepsilon,k}}{\partial x_2} & \text{on } \Gamma_0\,,\\ \hat{V}_{\varepsilon,k} \text{ is } x_1\text{-periodic} & \text{on } \Gamma_{\text{in}}\cup\Gamma_{\text{out}}\,.\end{cases} \tag{5.4.6}$$

It remains to show that this first order implicit wall-law has a solution and is an approximation in the smooth domain Ω_0 of the rough problem (5.3.1). The existence of solution in $H^1_{\Gamma_1}(\Omega_0)$ (H^1-functions vanishing on Γ_1) for problem (5.4.6) is not discussed here (see for example [40]), but the error estimate are given in the following theorem.

Theorem 5.4.1. There exists two positive constants c_6 and c_7, depending only of the mode $\hat{C}_k$ and not on the frequency k such that:

$$\|\hat{u}_{\epsilon,k}-\hat{V}_{\varepsilon,k}\|_{L^2(\Omega_0)}\leq c_6\,\varepsilon^{3/2} \text{ and } \quad \left\|\hat{u}_{\epsilon,k}-\hat{V}_{\varepsilon,k}\right\|_{H^1(\Omega_0)}\leq c_7\sqrt{\varepsilon}. \tag{5.4.7}$$

Proof. We split the error into two parts:

$$\|\hat{u}_{\epsilon,k} - \hat{V}_{\varepsilon,k}\|_{L^2(\Omega_0)} \leq \|\hat{u}_{\epsilon,k} - \overline{u}_{\varepsilon,k}\|_{L^2(\Omega_0)} + \|\overline{u}_{\varepsilon,k} - \hat{V}_{\varepsilon,k}\|_{L^2(\Omega_0)}\,.$$

The first term is controlled thanks to Proposition 3. For the second one, let us define $\Theta := \overline{u}_{\varepsilon,k} - \hat{V}_{\varepsilon,k}$ and consider the boundary value problem it satisfies:

$$\begin{cases} \mathcal{L}_k\,\Theta = 0 & \text{in } \Omega_0\,, \\ \Theta = 0 & \text{on } \Gamma_1\,, \\ \Theta = \varepsilon\,\overline{\beta}\,\left(\frac{\partial \hat{u}_{0,k}}{\partial x_2} - \frac{\partial \hat{V}_{\varepsilon,k}}{\partial x_2}\right) & \text{on } \Gamma_0\,, \\ \Theta \text{ is } x_1\text{-periodic} & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{out}}\,. \end{cases} \tag{5.4.8}$$

We re-express the boundary condition on Γ_0 introducing a Robin like condition, namely:

$$\Theta - \varepsilon\,\overline{\beta}\,\frac{\partial \Theta}{\partial x_2} = \varepsilon\,\overline{\beta}\,\left(\frac{\partial \hat{u}_{0,k}}{\partial x_2} - \frac{\partial \overline{u}_{\varepsilon,k}}{\partial x_2}\right) = -\varepsilon^2\overline{\beta}\frac{\partial \hat{u}_{1,k}}{\partial x_2} \quad \text{on } \Gamma_0\,, \tag{5.4.9}$$

where the rhs is now explicitly known. One sets

$$a_k(\theta, v) = (\nabla\theta, \nabla v)_{\Omega_0} + ik(\theta, v)_{\Omega_0} + \left(\frac{\theta}{\varepsilon\overline{\beta}}, v\right)\,,$$

and it is easy to show that this bi-linear form is bi-continuous and coercive. The variational problem becomes now

$$a_k(\theta, v) = -\varepsilon^2\left(\frac{\partial \hat{u}_{0,k}}{\partial x_2}, v\right)_{\Gamma_0}, \quad \forall v \in H^1_{\Gamma_1}(\Omega_0)\,,$$

which gives directly by *a priori* estimates that

$$\|\nabla\theta\|_{L^2(\Omega_0)} \leq c\,\varepsilon^2, \quad \|\theta\|_{L^2(\Gamma_0)} \leq c\,\varepsilon^3.$$

One then uses the very weak estimates in order to estimate θ in the $L^2(\Omega_0)$ norm and concludes thanks to the last trace estimate. For the *a priori* part we simply decompose the error using every result established above to get:

$$\begin{aligned} \left\|\hat{u}_{\epsilon,k} - \hat{V}_{\varepsilon,k}\right\|_{H^1(\Omega_0)} &\leq \left\|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\right\|_{H^1(\Omega_0)} + \left\|\hat{\mathcal{U}}_{\epsilon,k} - \overline{u}_{\varepsilon,k}\right\|_{H^1(\Omega_0)} + \left\|\overline{u}_{\varepsilon,k} - \hat{V}_{\varepsilon,k}\right\|_{H^1(\Omega_0)} \\ &\leq \left\|\hat{u}_{\epsilon,k} - \hat{\mathcal{U}}_{\epsilon,k}\right\|_{H^1(\Omega_\varepsilon)} + c\sqrt{\varepsilon}\|\nabla_y\beta\|_{L^2(Z^+\cup\Gamma\cup P)} + \varepsilon^2 \\ &\leq c\,(\varepsilon + \sqrt{\varepsilon} + \varepsilon^2) \leq c'\,\sqrt{\varepsilon}\,. \end{aligned}$$

□

5.5 Numerical results

5.5.1 Discretization

In this part we aim to prove numerically that wall-laws perform better approximation than the zeroth order guess. For this sake we define an explicit shape of the roughness setting f in (5.2.1) to be :

$$f(y_1) := -\frac{(1+\cos(y_1))}{2} - \delta,$$

with δ being a positive constant equal to $5 \cdot 10^{-2}$. The periodicity of the bottom shape and of the boundary conditions on $\Gamma_{\text{in}} \cup \Gamma_{\text{out}}$ allows to discretize only a single rough period, i.e. we set

$$\Omega_{\#,\varepsilon,-} := \{x_1 \in]0, 2\pi\varepsilon[\text{ and } x_2 \in]\varepsilon f(x_1/\varepsilon), 0[\}, \quad \Omega_{\#,\varepsilon,+} :=]0, 2\pi\varepsilon[\times]0, 1[, \quad \Omega_{\#,\varepsilon} := \Omega_{\#,\varepsilon,+} \cup \Omega_{\#,\varepsilon,-},$$

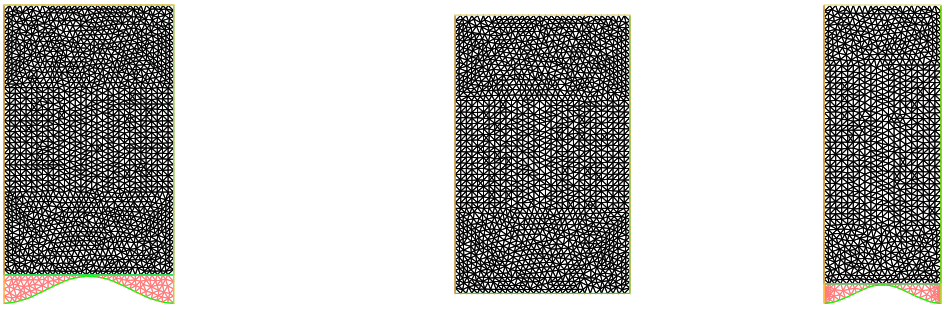


Figure 5.3: Meshes $\Omega_{\#,\varepsilon}$, $\Omega_{\#,\varepsilon,+}$, when $\varepsilon = 0.1$ and $Z^+ \cup \Gamma \cup P \cap \{y_2 \leq L = 10\}$ (coarse grids, see below for actual mesh sizes)

The mesh is periodic, *i.e.* the vertices on Γ_{in} are associated to elements containing edges on Γ_{out} (see p. 142 of the freefem++ documentation for further information on this facility). For a given ε, the meshes of $\Omega_{\#,\varepsilon}$ and $\Omega_{\#,0}$ are conforming on the upper part $\{x_2 \geq 0\}$. We take several values of ε, namely we set $i \in \{1, \ldots, 10\}$ and $\varepsilon = q^i$, where $q := 0.85$. In order to avoid discretization errors we set $n^\varepsilon := 90/\varepsilon^\alpha, \alpha = 0.2$ nodes on the horizontal fictitious boundary, and linearly proportional numbers of nodes on the other boundaries. This gives a mesh size h (maximal diameter of a triangulation, see p.88 [111]) depicted in fig. 5.4 (right) as a function of ε. Thus there exists a constant c independent of ε such that $h \leq c\varepsilon$. We fix a frequency $k = 10$ for which $\hat{C}_k \equiv 10$, we compute numerical approximations of

- $\hat{u}_{\varepsilon,k,h}$ solving the discretized problem (5.3.1)
- $\hat{u}_{0,k,h}$, the zeroth order Poiseuille-like approximation solving system (5.3.2)

- $\hat{V}_{\varepsilon,k,h}$, the implicit discrete wall-law.

In order to compute the cell problem and $\overline{\beta}$, the constant at infinity related to the specific roughness f, we discretize a cell problem defined on a truncated domain : find β_L solving

$$\begin{cases} -\Delta\beta_L = 0 & \text{in } Z^+ \cup \Gamma \cup P \cap \{y_2 < L\} \\ \partial_{\mathbf{n}}\beta = 0 & \text{on } \{y_2 = L\} \\ \beta = -y_2 & \text{on } P_0 \end{cases}$$

It is shown in [118] that the solution β_L converges exponentially fast, when $L \to \infty$, towards the solution of (5.3.15). So solving the problem above provides a good approximation of $\overline{\beta}_L$. And we use this numerical value in the boundary condition on Γ_0 in (5.4.6) in order to compute $\hat{V}_{\varepsilon,k,h}$. The code is written in `freefem++` language [111]: it is very well suited for solving complex valued variational problems with finite elements. Our code is available through Internet [1].

5.5.2 Error estimates

We compute numerical equivalent norms for *a priori* and very weak estimates. We plot this results in the log-log scale for various sizes ε (in abscissa) in fig. 5.4. We recover better

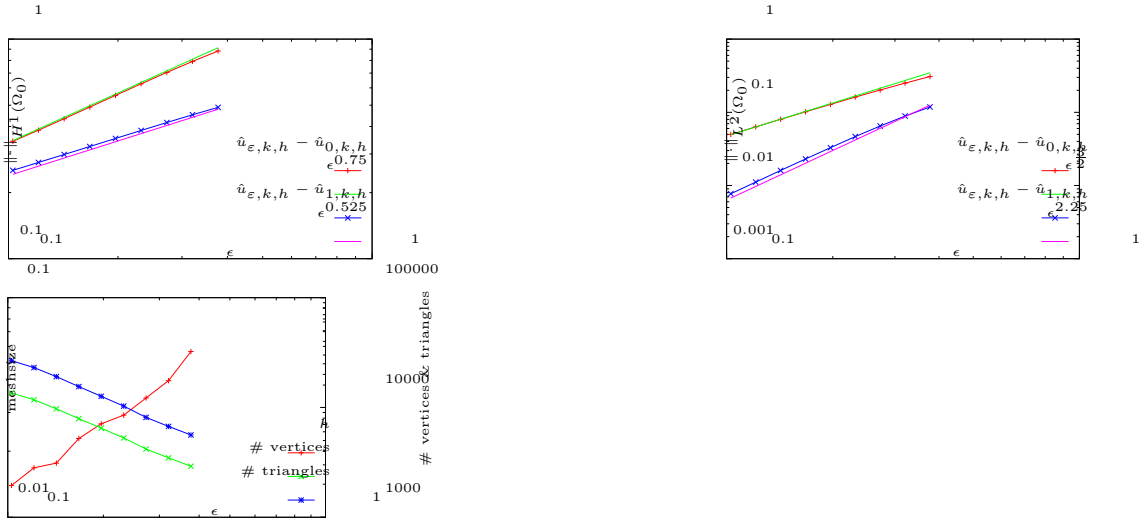


Figure 5.4: Numerical error estimates in $H^1(\Omega_0)$ (left) and $L^2(\Omega_0)$ (middle) norms, and mesh parameters (right)

orders of convergence than expected: the very weak estimates provide $\varepsilon^{\frac{3}{2}}$ convergence for the Poiseuille profile while they give $\varepsilon^{\frac{9}{4}}$ for the wall-law. The $H^1(\Omega_0)$ norm ($\sim \varepsilon^{\frac{3}{4}}$) is better than expected for the Poiseuille profile while surprisingly the error is worse for the

[1] `http://ljk.imag.fr/membres/Vuk.Milisic/Software/complexWallLaw.edp`

wall-law wrt this norm. This is due to the $\sqrt{\varepsilon}$ Dirichlet norm of gradient of boundary layers as already shown in Theorem 5.4.1. This numerical test shows that for this geometry case this latter estimate is almost optimal.

Annexe C

Direct simulations

In this appendix we focus on extending the results of [148] in the numerical direction. Namely, we present some preliminary numerical simulations which aim to give orders of magnitudes in terms of numerical costs of direct 3D simulations. More precisely, we determine actual limits, when running three-dimensional blood flow simulations of the non-homogenized stented arteries. We solve the stationary Stokes equations for an artery containing a saccular aneurysm. We examine the relation between discretization parameter, problem scale, and computation time required to solve the stationary Stokes equations. Finite element model of a two-layer 32 wire stent was constructed. We demonstrate that its coarse mesh could not be accurately incorporated in the finest discretization of the blood medium. Finally, a simplified ten-wire stent model was build. The results of the stented versus the unstented vessel show substantial difference in flow pattern inside the aneurysmal pouch. Concluding remarks and possible perspectives are given at the end.

C.1 Numerical investigation: Saccular side aneurysm

The objective of this numerical experiment is to determine actual computational limits, when running three-dimensional blood flow simulations of the non-homogenized stented vessels. Extremely small wire cross-section, $\varnothing$=0.1 mm, complex, almost random, spacing between braided wires could not be properly modeled in actual computational reality. Even when such models have been developed, industrial computer aided design (CAD) programs, mesh generators and finite element analysis tools are not well optimized for processing complex free-form geometries. However, it is important to analyze modeling and discretization limits, spatial resolution, memory needs and computation time required to guarantee an accurate and reliable hemodynamic simulation of stented vessels. The authors hope that the sequence of direct simulations brings an additional design insight, and could be used as a reference solution for the further three-dimensional homogenization research.

Parametric, three-dimensional model of a blood vessel with a side saccular aneurysm was built using commercial software, CATIA V5. The aneurysm has an ellipsoidal extended shape of 22 mm over 17 mm across its largest and smallest diameters, respectively. It is

attached to the parent vessel of 60 mm long within a constant diameter of 10 mm, fig. C.1, (left). These values are representative of the femoral artery. A braided tabular assembly of thin metallic wires is lodged against the lumen of the vessel to serve as a porous barrier disrupting blood flow into the aneurysm, fig. C.1, (right). A finite element model of a two-layer stent, braided of 32 wires has been built, fig. C.1, (bottom). Similar to commercial stents, but not being an exact replica, it has a 10 mm diameter, a length of 30 mm, and a wire diameter of $\varnothing$=0.1 mm. Its coarse mesh, 5-7 tetrahedral faces per wire cross-section, contains more than half a million 4-node tetrahedral elements.

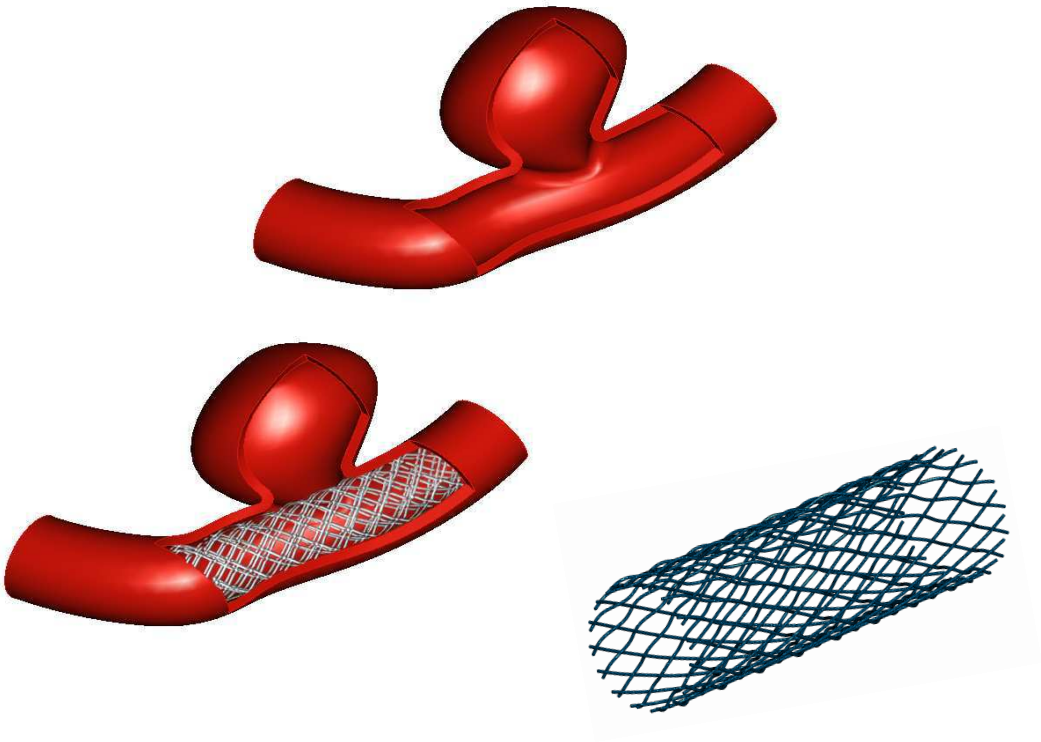
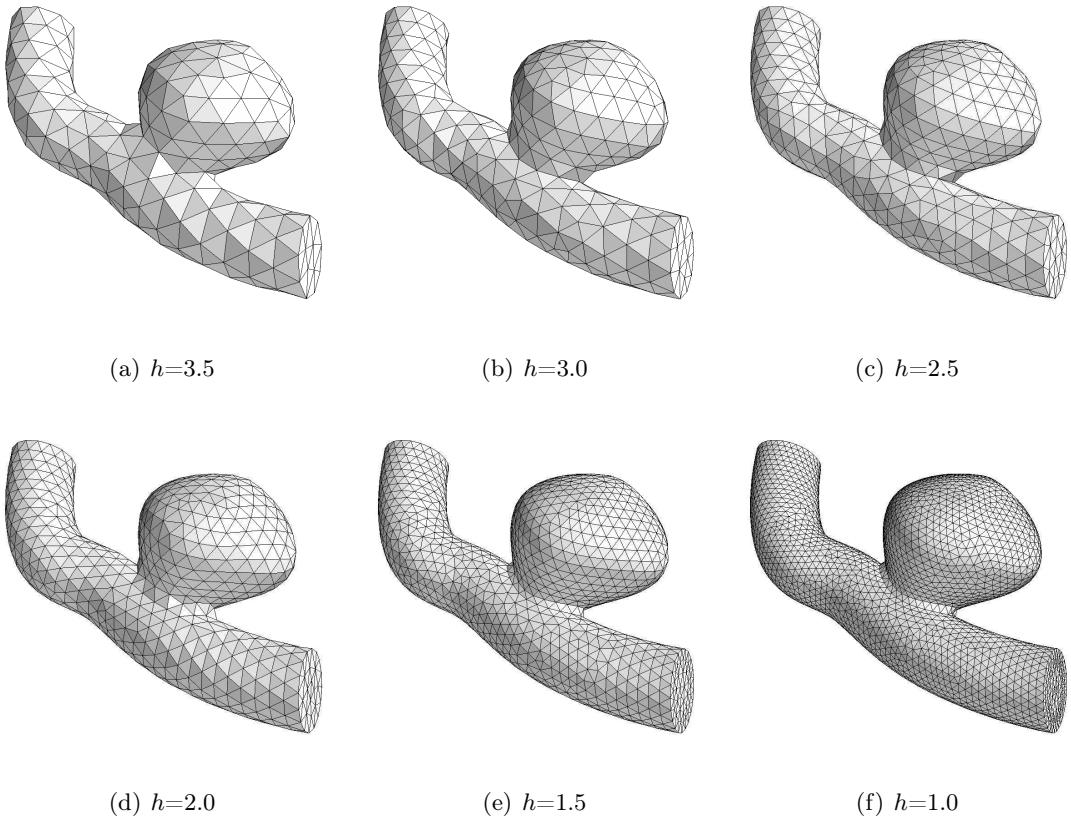


Figure C.1: CAD model of a parent vessel with a side wall aneurysm (top left). Schematic illustration of a wire multi-layer stent, which reduces blood flow into the aneurysm (top right). The cutout is for visualization purpose only. A coarse finite element model of a wire two-layer stent counts 546K tetrahedral elements (bottom).

First, we discretize the unstented artery vessel, imposing a uniform node spacing for the whole medium. Tetrahedral meshes were generated by an advancing front, followed by a tetrahedral filler technique, in order to produce high-quality, quasi-uniform meshes with a low element size variance. Consecutive levels of mesh refinement are presented in fig. C.2. Computation time, required to simulate one or several cardiac cycles could then be related to a spatial mesh resolution, by solving a given hemodynamic problem for each of the presented discretization. `Freefem++` open source finite element code was

used to compute a steady-state solution of the Stokes equations. The velocity-pressure fields were discretized by the Taylor-Hood element ((P2/P1) finite element basis). The blood was assumed to behave like an incompressible Newtonian fluid, with a constant dynamic viscosity of $3.5 \cdot 10^{\breve{}3}$ Pa·s, and an homogeneous density of 1060 kg·m^{-3}. We do not consider the compliance of arterial walls due to the complexity of the numerical modeling and requirement for a fluid structure interaction environment to solve a coupled problem. The assumption of rigid wall is based on [130], where the authors conclude that a presence of wall motion does not have significant influence of the global fluid dynamic characteristics of the femoral artery bifurcation. The inflow boundary condition is based on the constant pressure profile of 80 mmHg. We imposed the usual non-slip boundary conditions on the vessel wall, while a pressure drop of 0.07 mmHg was prescribed on the outflow boundary. In addition, the tangential velocity component was set to be zero on the non-slip boundaries, $\mathbf{u} \cdot \tau_{|\Gamma_{\text{in}},\Gamma_{\text{out}}} = 0$. These boundary conditions together with a pressure gradient establish a steady laminar flow with a Reynolds number Re=443, and a flow rate of 11.51 ml/sec.



(a) h=3.5 (b) h=3.0 (c) h=2.5

(d) h=2.0 (e) h=1.5 (f) h=1.0

Figure C.2: Different levels of mesh refinement. Quasi-uniform meshes with a prescribed discretization element size h in millimeters. An overall mesh, data are reported in table C.1.

We made use of a non-parallel version of a conjugate gradient solver with a convergence criteria of $10^{\breve{}8}$. To preserve the positive-definiteness of the global matrix, a penalization term of order $10^{\breve{}12}$ is introduced (see for instance p. 210-214 [83]). Identical results have been obtained by the GMRES iterative solver without penalization term, though it requires more memory to operate. An overall mesh, finite element, and computation data are organized in table C.1. For a given discretization parameter h, it reports a number of produced mesh nodes, tetrahedral elements, degrees of freedom, non-zero coefficients of the symmetric finite element matrix, numerically computed flow rates in units of ml/sec. Solution time, in seconds, required to solve the stationary Stokes equations is reported in the extreme right column. The computational results reported in table C.1 reveal that even the finest discretization of a fluid medium would not be sufficient to properly incorporate a coarse finite element model of the wire two-layer stent, presented in fig. C.1. The finest mesh discretization parameter is 5 times larger than a wire diameter.

Remark: The tasks were executed on the laboratory cluster, powered by 16 Intel Xeon E5462 @2.80GHz processors; 12Gb of available RAM memory are designated for each two

h	**nodes**	**tetrahedra**	**dof**	**!=0 coef.**	**flow rate**	**cpu time**
3.5	443	1664	8918	385982	9.87	17
3.0	685	2778	14272	632774	10.17	39
2.5	1184	5225	25628	1166637	10.61	126
2.0	2012	9278	44483	2052252	11.10	324
1.5	4725	23456	108144	5103210	11.29	1363
1.0	14276	75051	335807	16124234	11.49	10499
0.8	28257	153582	675171	32749351	11.51	39314
0.5	109211	616710	2657609	130403575	11.51	378162

Table C.1: Mesh data summary, non-zero coefficients of the symmetric finite element matrix, computed flow rate [ml/sec], computation time [sec] of the steady-state Stokes equations.

processors. For the finest mesh, h=0.5, 4Gb of memory were allocated. Reported cpu time represents the total cpu time taken by one single processor to obtain the converged solution. We note that since each discretization was built in the stand alone way, each processor worked independently and there was no communication or synchronization overhead in the calculations. Each processor was assigned only one mesh and one variational problem to be resolved. We have repeated each computation several times, observing negligible variance in computation time.

In the second part of our numerical experiment, we simplify the original stent model, by replacing it with a pattern of unattached ring-like struts across an aneurysm throat. A similar two-dimensional version has been recently proposed in [148]. Two modeling techniques were tested to place stent wires. The first technique was to construct stent wires, completely enclosed by the blood medium. Wire centers were displaced into the parent vessel from the outer boundary by 3/2 of the wire radius. An automated mesh generator had difficulties to properly define all enclosed surfaces, and to complete a meshing procedure. Moreover, this technique produced extremely small elements, located between stent wires and vessel boundary. Therefore, a unique six-wire model was constructed, fig. C.3 (left). Almost worthless, it takes 214 hours to solve the Stokes equations, using a non-parallel iterative solver. A model related data was summarized in table C.2.

wire $\varnothing$	h	**nodes**	**tetrahedra**	**dof**	**!=0 coef.**	**flow rate**	**cpu time**
0.8	0.1 - 0.5	$218.9{\cdot}10^3$	$1.2{\cdot}10^6$	$5.3{\cdot}10^6$	$261.4{\cdot}10^6$	8.34	769384

Table C.2: Stented aneurysm with completely enclosed stent wires: mesh data summary, non-zero coefficients of the symmetric finite element matrix, computed flow rate [ml/sec], computation time [sec] of the steady-state Stokes equations.

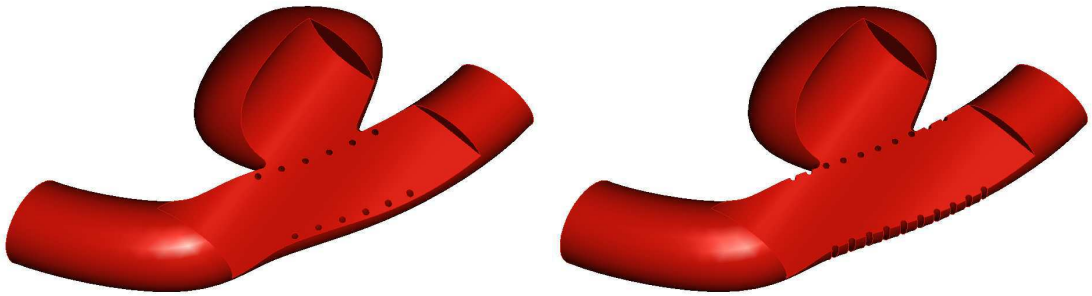


Figure C.3: Three-dimensional CAD model of a saccular aneurysm within struts pattern across the aneurysm throat. Wire diameter is $\varnothing$=0.8 mm. The cutout exposes struts location, and is for visualization purpose only. Stents struts are completely enclosed by the blood medium (left), struts are partially displaced outside of the blood medium (right).

The second strategy was to partially enclose stent wires by the blood medium; thus, wire centers were displaced into the vessel by 1/6 of the wire radius. Wires were cloned along the parent vessel direction with respect to its curvature. The distance between two wire centers is 2 mm, fig. C.3 (right). Four separate ten-wire models were constructed. Keeping the same distance between wire centers, we have consequently decreased a wire diameter, from 0.9 to 0.6 mm. We note that it is, however, about 10 times larger than the actual wire diameter used for commercial stents. Locally refined, adaptative meshes were built using the octree algorithm, imposing a nodal spacing of 0.12 mm around stent struts. The transitional element distribution between the respective regions of refined and global mesh density is presented in fig. C.4.

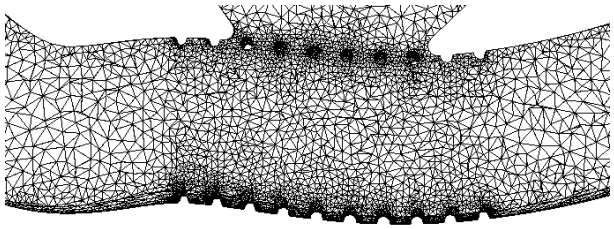


Figure C.4: Finite element model of a saccular aneurysm within ten partially enclosed struts of $\varnothing$=0.8 mm. Zoom view over the stented region. A cross-section of the stent strut is represented by approximately 16 elements. Discretization parameter h=0.7 for a global domain, h=0.12 near the stent struts.

The results of computations show that the presence of a stent induces a truly remarkable

wire ⌀	h	**nodes**	**tetrahedra**	**dof**	**!=0 coef.**	**flow rate**	**cpu time**
0.9	0.12 - 0.7	95196	481845	$2.2{\cdot}10^6$	$104.5{\cdot}10^6$	9.90	235311
0.8	0.12 - 0.7	97405	495366	$2.2{\cdot}10^6$	$107.3{\cdot}10^6$	10.10	180206
0.7	0.12 - 0.7	99913	511385	$2.3{\cdot}10^6$	$110.6{\cdot}10^6$	10.23	214308
0.6	0.12 - 0.7	100184	515291	$2.3{\cdot}10^6$	$111.3{\cdot}10^6$	10.49	233170

Table C.3: Stented aneurysm with partially enclosed stent wires: wire diameter, mesh data summary, non-zero coefficients of the symmetric finite element matrix, computed flow rate [ml/sec], computation time [sec] of the steady-state Stokes equations.

change of a blood flow near the throat region, fig. C.5. In the case of a stented vessel the streamlines are not bent towards the aneurysm pouch, but remain similar to the bulk flow behavior. A parabolic flow profile is observed at the extreme ends of the vessel. The flow rates were computed at the upstream and downstream boundaries. The presence of the stent struts decreased the flow rate in the parent vessel. It averages the pressure inside the aneurysmal sac (this fact was already proved rigorously in [148] in 2D), and eliminates neck singularities, see fig. C.6 (bottom left). Velocity vectors, depicted in fig. C.7 illustrate that after stent placement, the aneurysmal vortex was no longer present. This confirms in 3D results theoretically proved in 2D in [148]. Adaptive refinement and extremely fine mesh found to be insufficient to properly model commercial multi-layer stents. It is evident that direct finite element simulations could give an additional insight, a better understanding of blood flow nature within a specific stent design, but we actually need much more computational power to simulate a pulsatile flow, where hundreds of time steps should be computed within one cardiac cycle. For this reason a work in preparation [149] aims at incorporating homogenized interface conditions and at providing some quantitative averaged results useful for clinical purposes.

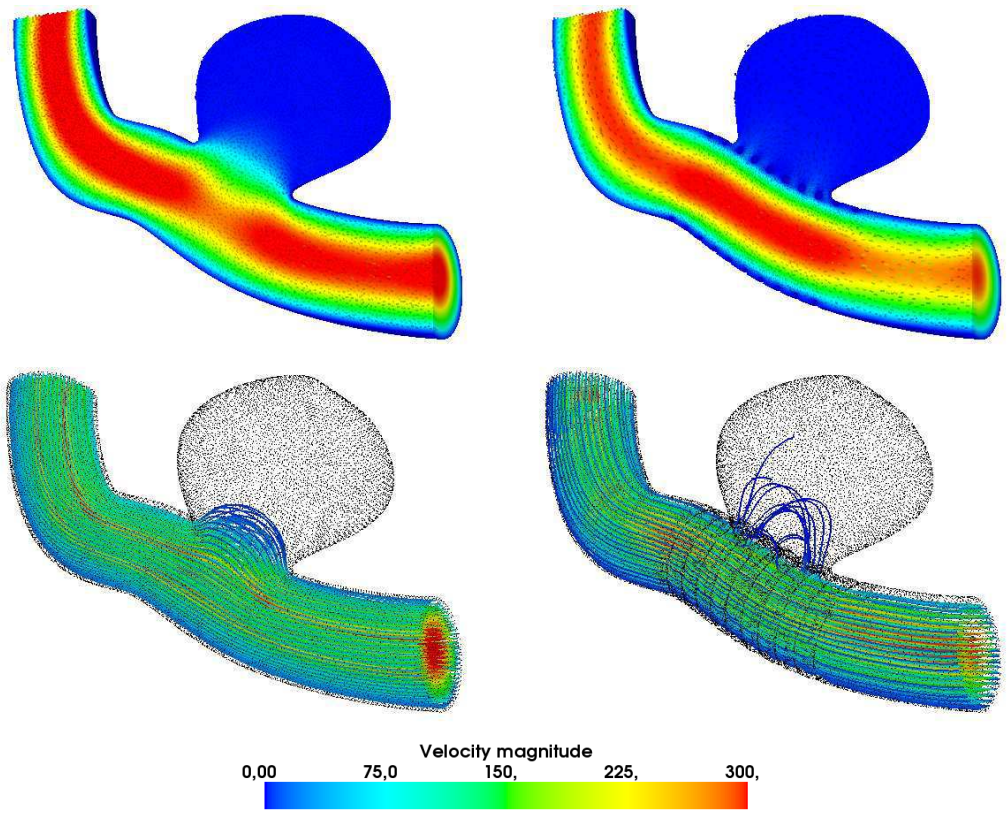


Figure C.5: Stationary velocity field (top) and streamlines (bottom) computed before (left) and after (right) stent treatment.

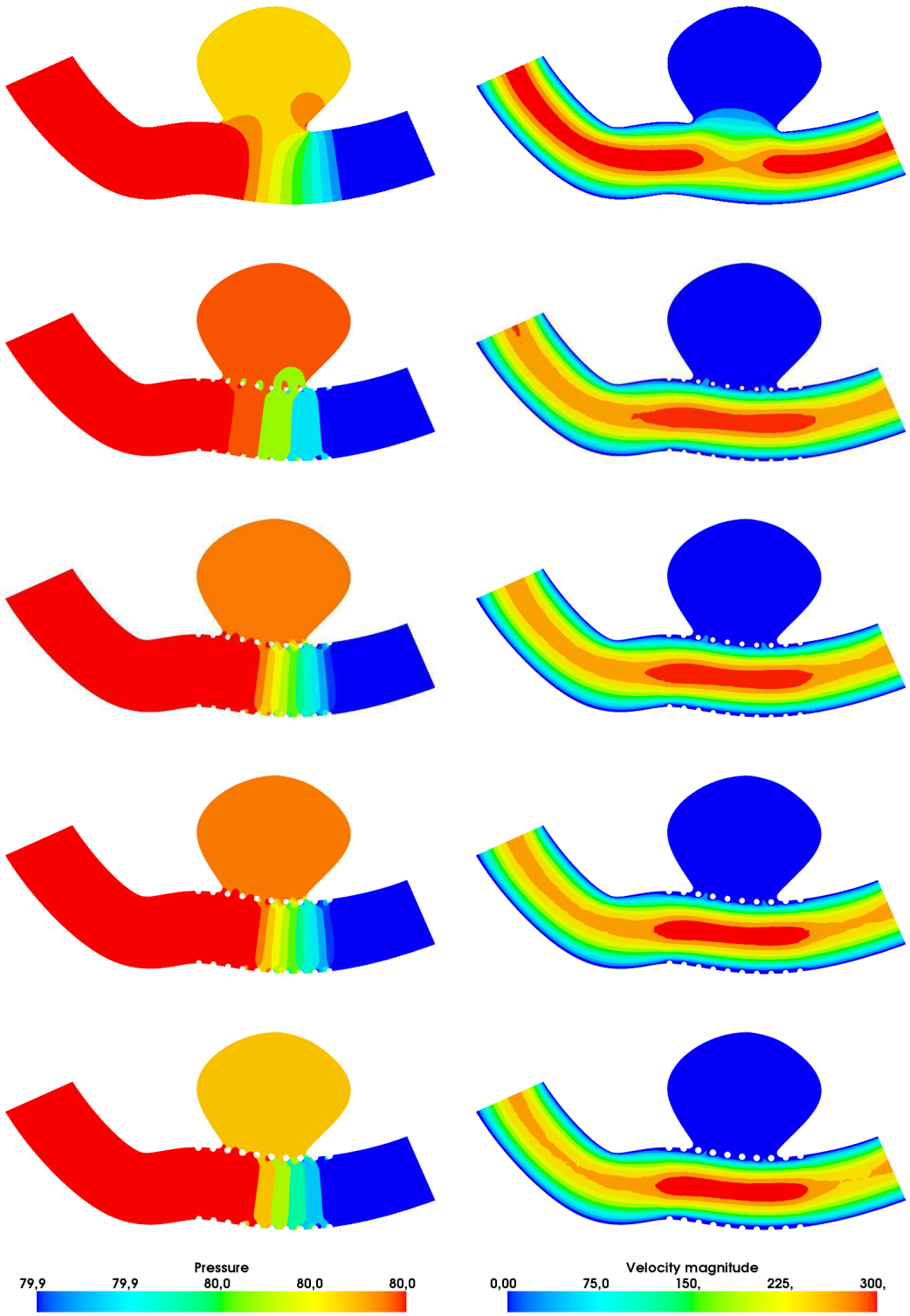
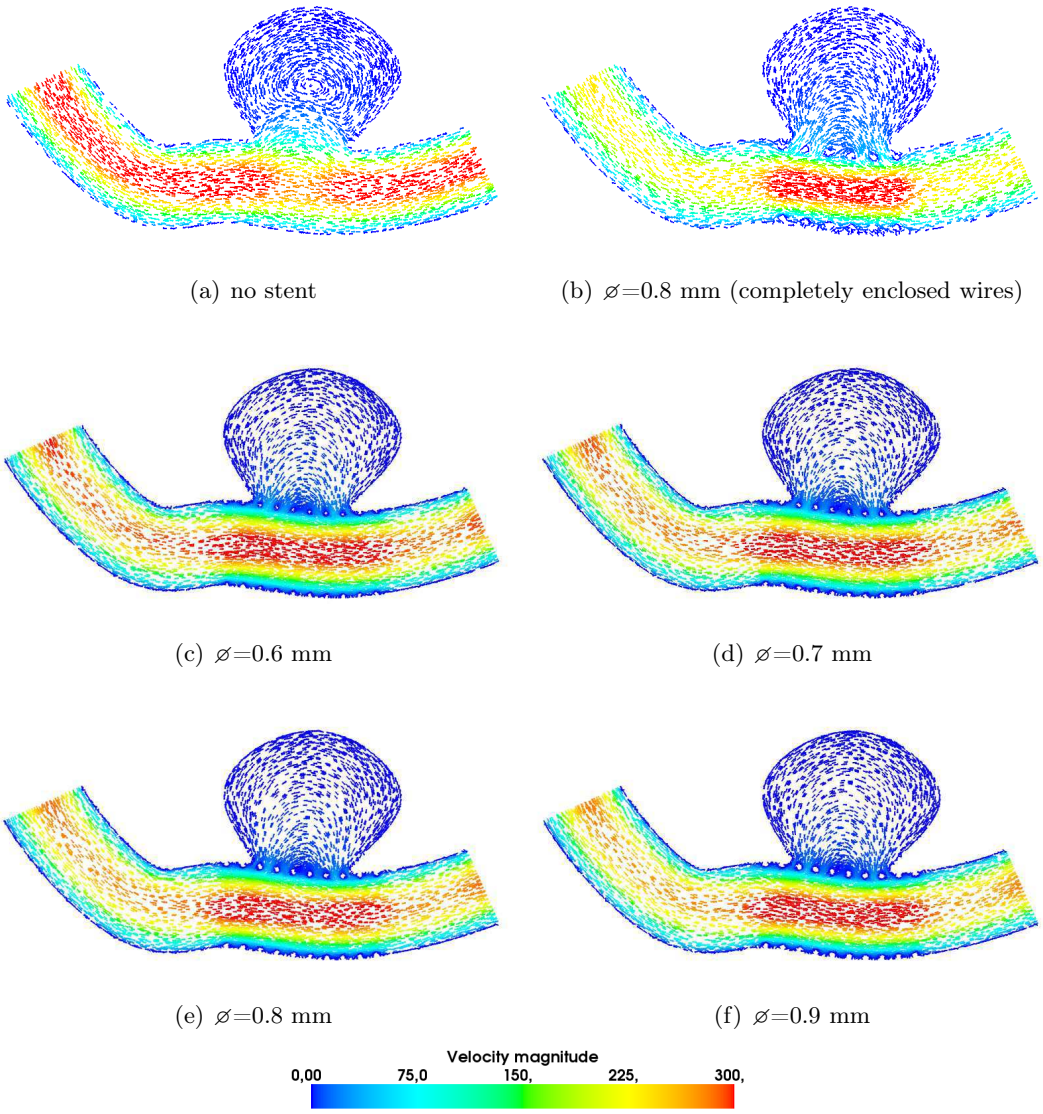


Figure C.6: Sequence of pressure (left), velocity (right) solution contours. From top to bottom: unstented vessel, wire ⌀=0.6 mm, 0.7 mm, 0.8 mm, and 0.9 mm, respectively.



(a) no stent

(b) ⌀=0.8 mm (completely enclosed wires)

(c) ⌀=0.6 mm

(d) ⌀=0.7 mm

(e) ⌀=0.8 mm

(f) ⌀=0.9 mm

Figure C.7: Velocity vectors colored by magnitude (not scaled arrow symbols).

Bibliographie

[1] Achdou,Y. and Pironneau, O., Domain Decomposition and wall laws, C. R. Acad. Sci. Paris (1995)

[2] Achdou,Y. and Pironneau, O. and Valentin, F., Analysis of the First and Second Order Wall Laws for Rough Domains by Domain Decomposition, Technical report, INRIA (1998)

[3] Achdou, Y. and Le Tallec, P. and Valentin, F. and Pironneau, O., Constructing wall laws with domain decomposition or asymptotic expansion techniques, Comput. Methods Appl. Mech. Eng. (1998)

[4] Achdou, Y. and Pironneau, O. and Valentin, F., Effective boundary conditions for laminar flows over periodic rough boundaries, J. Comput. Phys. (1998)

[5] Allaire, G., Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes, Arch. Rational Mech. Anal. (1991)

[6] Amirat, Y. and Simon, J., Influence of rugosity in laminar hydrodynamics, C.R. Acad. Sci. Paris (1996)

[7] Aoki, T., Inamuro, T. and Onishi, Y., Slightly rarefied gas flow over a body with small accomodation coefficient, *J. Phys. Soc. Japan.* (1979).

[8] Aregba-Driollet D. and Natalini, R., Convergence of relaxation schemes for conservation laws, *Appl. Anal.* **1-2** (1996), pp. 163–193.

[9] Audusse, E., Modélisation hyperbolique et analyse numérique pour les écoulements en eaux peu profondes, PhD Thesis, Université Paris VI (2004)

[10] Audusse, E., Bouchut F., Bristeau, M.-O., Klein R. and Perthame, B., A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM J. Sc. Comp. (2004)

[11] Audusse, E. and Bristeau, M.-O., A 2d well-balanced positivity preserving second order scheme for Shallow Water Flows on unstructured meshes, J. of Comp. Phys. (2005)

[12] Audusse, E., A multilayer Saint-Venant model : derivation and numerical validation, Discrete Contin. Dyn. Syst. Ser. B (2005)

[13] Audusse, E. and Bristeau, M.-O., Finite-volume solvers for a multilayer Saint-Venant system, Int. J. Appl. Math. Comput. Sci. (2007)

[14] Audusse, E. and Bristeau, M. O. and Decoene, A., Numerical simulations of 3D free surface flows by a multilayer Saint-Venant model, Internat. J. Numer. Methods Fluids (2008)

[15] Audusse, E., Bristeau, M. O., Perthame, B. and Sainte-Marie, J., A multilayer Saint-Venant system with mass exchanges for Shallow Water flows :derivation and numerical validation, M2AN Math. Model. Numer. Anal. (2010)

[16] Augsburger, L., Flow changes investigation due to the insertion of a braided stent in an inertia driven flow aneurysm model using experimental methods, Technical report, Laboratoire d'Hémodynamique et de Technologie Cardiovasculaire, EPFL, Suisse (2008)

[17] Aw, A. Klar, A., Materne, T. and Rascle, M., Derivation of continuum traffic flow models from microscopic follow the leader models, *SIAM J. Appl. Math.* **63** (2002), pp. 259–278.

[18] Aw A. and Rascle, M., Resurrection of second order models of traffic flow ?, *SIAM J. Appl. Math.* **60** (2000), pp. 916–938.

[19] Azèrad, P. and Guillén F., Equations de Navier-Stokes en bassin peu profond : l'approximation hydrostatique, C.R. Acad. Sci. Paris (1999)

[20] Azèrad, P. and Guillén F., Mathematical justification of the hydrostatic approximation in geophysical fluid dynamics, SIAM J. Math. Anal. (2001)

[21] Bardos, C., Golse, F. and Levermore, D., Fluid dynamic limits of kinetic equations. I. Formal derivations, *J. Statist. Phys.* **63** (1991), pp. 323–344.

[22] Bardos, C., Golse, F. and Levermore, D., Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation, *Comm. Pure Appl. Math.* **46** (1993), pp. 667–753.

[23] Bardos, C., Golse, F. and Levermore, D., Macroscopic limits of kinetic equations, *IMA Vol. Math. Appl.* **29** (1991), pp. 1–12.

[24] Basson, A. and Gérard-Varet, D., Wall laws for fluid flows at a boundary with random roughness, Comm. Pure Appl. Math. (2008)

[25] Bennoune, M. Lemou, M. and L. Mieussens, L., Uniformly stable numerical schemes for the Boltzmann equation preserving the compressible Navier–Stokes asymptotics, *J. Comput. Phys.* **227** (2008), pp. 3781–3803.

[26] Bermudez, A., Rodriguez, C. and Vilar, MA., Solving shallow water equations by a mixed implicit finite element method, IMA J. Numer. Anal. (1991)

[27] Berthelin, F., Tzavaras, A.E. and Vasseur A., From discrete velocity Boltzmann equations to gas dynamics before shocks, J. Stat. Phys. (2009)

[28] Bessems, D., Rutten, M. and Van de Vosse, F., A wave propagation model of blood flow in large vessels based on boundary layer theory, *submitted in* J. Fluid Mech. (2005)

[29] Bianchini, S., Hanouzet, B. and Natalini, R., Asymptotic behavior of smooth solutions for partially dissipative hyperbolic systems with a convex entropy. *Comm. Pure Appl. Math.* **60** (2007), pp. 1559-1622.

[30] Bianchini, S., Hyperbolic limit of the Jin-Xin relaxation model, *Comm. Pure Appl. Math.* **47** (1994), pp. 787-830.

[31] Bird, R.B., Armstrong, R.C. and Hassager, O., Dynamics of polymeric liquids, volume 1 : fluid mechanics - second edition, J. Wiley and Sons (1987)

[32] Boltzmann, L., Lectures on gas theory - reprinted version, Dover Publications (1995)

[33] Bonnetier, E. and Bresch, D. and Milišić, V., High order multi-scale wall laws : part II, the non periodic case, Advances in Mathematical Fluid Dynamics (2009)

[34] Bonnetier, E. and Bresch, D. and Milišić, V., *A priori* convergence estimates for a rough Poisson-Dirichlet problem with natural vertical boundary conditions, Advances in Mathematical Fluid mechanics (2010)

[35] Bouchut F. and Westdickenberg M., Gravity driven shallow water models for arbitrary topography, Comm. in Math. Sci. (2004)

[36] Bouchut, F. Golse, F. and Pulvirenti, M., Kinetic equations and asymptotic theory, *Gauthiers-Villars* (2000).

[37] Boutounet M., Chupin L., Noble P. and Vila, J.–P., Shallow water flows for arbitrary topography, Comm. Math. Sciences, 6 (2008)

[38] Brenier, Y. and Roesch, M., Reconstruction d'écoulements incompressibles à partir de données lagrangiennes, Actes du 29ème Congrès d'Analyse Numérique :CANum'97 (Larnas, 1997), ESAIM Proc. (1998)

[39] Bresch, D. and Milišić, V., Towards implicit multi-scale wall laws, C. R. Acad. Sciences, Série Mathématiques (2009)

[40] Bresch, D. and Milišić, V., High order multi-scale wall laws : part I, the periodic case, Quart. Appl. Math. (2010)

[41] Bresch, D., Guillén-González, F., Masmoudi, N. and Rodríguez-Bellido, M.A., Asymptotic derivation of a Navier condition for the primitive equations, Asympt. Anal. (2003)

[42] Bresch, D., Desjardins, B. and Lin, C.-K., On some compressible fluid models : Korteweg, Lubrication, and Shallow Water systems, Comm. Partial Diff. Eq. (2003)

[43] Bresch, D. and Desjardins, B., Existence of global weak solutions for a 2D viscous shallow water equations and convergence to the quasi-geostrophic model, Comm. Math. Phys. (2003)

[44] Bresch, D. and Noble, P., Mathematical justification of a shallow water model, Meth. and Appl. of Anal. (2007)

[45] Brezis, H., Analyse fonctionnelle, Masson (1983)

[46] Bristeau, M. O. and Sainte-Marie, J., Derivation of a non-hydrostatic shallow water model ;comparison with Saint-Venant and Boussinesq systems, Discrete Contin. Dyn. Syst. Ser. B (2008)

[47] Bristeau, M. O. and Coussin, B., Boundary conditions for the shallow water equations solved by kinetic schemes, INRIA report (2001)

[48] James E. Broadwell, Study of rarefied shear flow by the discrete velocity method. *J. Fluid Mech.* (1964).

[49] Buffard, T., Gallouët, T. and Hérard J-M., Un schéma simple pour les équations de Saint-Venant, C. R. Acad. Sci. Paris (1998)

[50] Buffard, T., Gallouët, T. and Hérard J-M., A sequel to a rough Godunov scheme : application to real gases, Computers and Fluids (1999)

[51] Bui, A. T., Existence and uniqueness of a classical solution of an initial boundary value problem of the theory of shallow waters, SIAM J. Math. Anal. (1981)

[52] Cabannes, H., Gatignol, R. and Luo, L.-S., The discrete Boltzmann equation, Lecture notes, University of California, Berkeley (1980)

[53] Caflisch, Russel E. The fluid dynamic limit of the nonlinear Boltzmann equation (1980).

[54] Caflish, R., Jin, S. and Russo, G., Uniformly accurate schemes for hyperbolic systems with relaxation, *SIAM J. Numer. Anal.* **34** (1997) 246–281.

[55] Caiazzo, A., Fernández, M.A., Gerbeau, J.-F. and Martin, V., Projection schemes for fluid flow through a porous interface, SIAM J. Sci. Comput. (2011)

[56] Carrillo, J.-A., Goudon, T. and Lafitte, P., Simulation of fluid and particle flows : Asymptotic preserving schemes for bubbling and flowing regimes, J. Comput. Phys. (2008)

[57] Carrillo, J.-A., Goudon, T., Lafitte, P. and Vecil, F., Numerical schemes of diffusion asymptotics and moment closures for kinetic equations, J. Sci. Comput. (2008)

[58] Castro, M., Macías, J. and Parés, C., A Q-scheme for a class of systems of coupled conservation laws with source term. Application to a two-layer 1-D shallow water system, M2AN Math. Model. Numer. Anal.(2001)

[59] Castro, M. J. and García-Rodríguez, J. A. and González-Vida, J. M. and Parés, C., A parallel 2D finite volume scheme for solving the bilayer shallow-water system : modellization of water exchange at the Strait of Gibraltar, Parallel computational fluid dynamics, Elsevier B. V., Amsterdam (2005)

[60] Cercignani, C., Theory and Application of the Boltzmann equation, Springer-Verlag, New York (1988)

[61] Chalabi, A., On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms *Mathematics of Computation*, **66**, (1997) pp. 527–545

[62] Chalabi, A., Convergence of relaxation schemes for hyperbolic conservation laws with stiff source terms *Mathematics of Computation*, **68**, (1999), pp. 955–970

[63] Chalabi, A. and Qiu, Y., Relaxation schemes for hyperbolic conservation laws with stiff source terms : application to reacting Euler equations *Journal of Scientific Computing*, **15**, (2000), pp. 395–416

[64] CHatelon, F.J., Muñoz-Ruiz, M.L., and Orenga, P., On a bi-layer shallow water problem, Nonlinear Anal.(2004)

[65] Chen, G.Q., Liu, T.P. and Levermore, C.D., Hyperbolic conservation laws with stiff relaxation terms and entropy, *Comm. Pure Appl. Math.* **47** (1994), no. 6, 787–830.

[66] Chen,Q., Miao C. and Zhang, Z., Well-posedness for the viscous shallow water equations in critical spaces, SIAM J. Math. Anal. (2008)

[67] Choquet, C. and Mikelić, A., Laplace transform approach to the rigorous upscaling of the infinite adsorption rate reactive flow under dominant Peclet number through a pore, Appl. Anal. (2008)

[68] Cioranescu, D. and Donato, P., An Introduction to Homogenization, Oxford university press (1999)

[69] Clopeau, Th. and Mikelić, A. and Robert, R., On the vanishing viscosity limit for the 2D incompressible Navier-Stokes equations with the friction type boundary conditions, Nonlinearity. (1998)

[70] Colin, S. and Lalond, P. and Caen, C., Validation of a second order slip flow model in rectangular microchannels, Heat transfert engineering (2004)

[71] Colin, T., The Cauchy problem and the continuous limit for the multilayer model in geophysical fluid dynamics, SIAM J. Math. Anal. (1997)

[72] Conca, C., Étude d'un fluide traversant une paroi perforée. I. Comportement limite près de la paroi, J. Math. Pures Appl. (9) (1987)

[73] Conca, C., Étude d'un fluide traversant une paroi perforée. II. Comportement limite loin de la paroi, J. Math. Pures Appl. (9) (1987)

[74] Csiszár, I., Information-type measures of difference of probability distributions and indirect observations, Stud. Sci. Math. Hung.-2- (1967)

[75] Decoene, A., Bonaventura, L., Miglio, E. and Saleri, F., Asymptotic derivation of the section-averaged shallow water equations for river hydraulics, MOX-Report (2007)

[76] De Dieu Zabsonré, J. and Narbona-Reina, G., Existence of a global weak solution for a 2D viscous bi-layer Shallow Water model, Nonlinear Anal. (2009)

[77] Degond, P. and Jin, S., A smooth transition model between kinetic and diffusion equations, *SIAM J. Numer. Anal.* **41** (6) (2005) 2671-2687

[78] Degond, P., Jin, S. and Liu, J.-G., Mach-number uniform asymptotic-preserving gauge schemes for compressible flows. *Bull. Inst. Math. Acad. Sin. (N.S.)* **2** (2007), pp. 851–892.

[79] Dimarco, G. and Pareschi, L., Hybrid multiscale methods I. Hyperbolic relaxation problems. Comm. Math. Sciences, **4**, (2006) pp. 155–177.

[80] Dimarco, G. and Pareschi, L., Hybrid multiscale methods II. Kinetic equations. SIAM Multiscale Model. Simul. (2008)

[81] Dimarco, G. and Pareschi, L., Exponential Runge-Kutta methods for stiff kinetic equations, (to appear) SIAM J. Numer. Anal.

[82] Dubois, F. and Le Floch, P., Boundary conditions for nonlinear systems of conservation laws, J. of Diff. Eq. (1988)

[83] Ern, A. and Guermond, J.-L., Theory and Practice of Finite Elements, Springer-Verlag (2004)

[84] Eymard R., Gallouët, T. and Herbin R., Finite volume methods, Handbook of numerical analysis (2000)

[85] Evans, Lawrence C., Partial differential equations, American Mathematical Society (2010)

[86] Fernández, M.A., Gerbeau, J.-F. and Martin, V., Numerical simulation of blood flow through a porous interface, M2AN Math. Model. Numer. Anal. (2008)

[87] Ferrari, S. and Saleri, F., A new two-dimensional Shallow Water model including pressure effects and slow varying bottom topography, M2AN Math. Model. Numer. Anal. (2004)

[88] Filbet, F. and Russo, G., High order numerical methods for the space non-homogeneous Boltzmann equation. *J. Comput. Phys.* **186**, (2003) pp. 457–480.

[89] Filbet, F., Pareschi, L. and Toscani, G., Accurate numerical methods for the collisional motion of (heated) granular flows. *J. Comput. Phys.* **202**, (2005) pp. 216–235.

[90] Filbet, F., Mouhot, C. and Pareschi, L., Solving the Boltzmann equation in $N \log_2 N$. *SIAM J. Sci. Comput.* **28**, (2006) pp. 1029–1053

[91] Filbet, F., An asymptotically stable scheme for diffusive coagulation-fragmentation models, Comm. Math. Sciences, **6**, (2008) pp. 257–280.

[92] Filbet, F. and Jin, S., A class of asymptotic preserving schemes for kinetic equations and related problems with stiff sources, J. Comp. Physics, **229**, (2010)

[93] Filbet, F. and Jin, S., An asymptotic preserving scheme for the ES-BGK model for the Boltzmann equation, J. Sci. Comp. **46**, (2011)

[94] Filbet, F. and Rambaud, A., Analysis of an Asymptotic Preserving scheme for relaxation systems, Submitted (2011)

[95] Flori, F., Orenga, P. and Peybernes, M., Sur un problème de Shallow Water bicouche avec conditions aux limites de Dirichlet, C.R. Acad. Sci. Paris (2005)

[96] Gabetta, E., Pareschi, L. and Toscani, G., Relaxation schemes for nonlinear kinetic equations, *SIAM J. Numer. Anal.* **34** (1997), 2168–2194

[97] Galdi, P.G., An introduction to the mathematical theory of the NS equations, vol I & II, Springer (1994)

[98] Gallouët, T., Hérard J-M. and Seguin, N., Some approximate Godunov schemes to compute shallow-water equations with topography, Computers and Fluids (2003)

[99] Gérard-Varet, D., Formal derivation of boundary layers in fluid mechanics, J. Math. Fluid Mech. (2005)

[100] Gérard-Varet, D., and Paul, T. Remarks on boundary layers expansions, Comm. in Part. Diff. Eq. (2008)

[101] Gerbeau, J.-F. and Perthame, B., Derivation of viscous Saint-Venant system for laminar shallow water ; numerical validation, Discrete Contin. Dyn. Syst. Ser. B (2001)

[102] Godlewski, E. and Raviart P.-A., Numerical approximations of hyperbolic systems of conservation laws, Appl. Math. Sc. (1996)

[103] Golse, F., Jin, S. and Levermore, C.D., The Convergence of Numerical Transfer Schemes in Diffusive Regimes I : The Discrete-Ordinate Method, SIAM J. Num. Anal. 36 (1999) pp. 1333-1369

[104] González-Vida, J. M. and Castro, Manuel J. and García-Rodríguez, J. A. and Macías, J. and Parés, C., Simulation of tidal currents in the Strait of Gibraltar using two-dimensional two-layer shallow-water models, Bol. Soc. Esp. Mat. Apl. S$\vec{e}$MA (2008)

[105] Goudon, T. and Lafitte, P., Splitting Schemes for the Simulation of Non Equilibrium Radiative Flows, preprint (2007)

[106] Gosse, L. and Toscani, G., An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations, C.R. Math. Acad. Sci. Paris 334 (2002)

[107] Gosse, L. and Toscani, G., Space localization and well-balanced schemes for discrete kinetic models in diffusive regimes, SIAM J. Numer. Anal. (2004)

[108] Hanouzet, B. and Natalini, R., Weakly coupled systems of quasilinear hyperbolic equations, *Differential Integral Equations* **6** (1996), 1279–1292.

[109] Hao, C., Hsiao, L. and Li, H., Cauchy problem for viscous rotating shallow water equations, J. Differ. Equ.(2009)

[110] Haspot, B., Cauchy problem for viscous shallow water equations with a term of capillarity, M3AS (2010)

[111] Hecht, F. and Pironneau, O. and Le Hyaric, A. and Ohtsuka K., Freefem++ Laboratoire Jacques-Louis Lions, Universite Pierre et Marie Curie (2005)

[112] Jäger, W. and Mikelić, A., On the boundary conditions at the contact interface between a porous medium and a free fluid, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) (1996)

[113] Jäger, W. and Mikelić, A., On the effective equations for a viscous incompressible fluid flow through a filter of finite thickness, Comm. Pure Appl. Math. (1998)

[114] Jäger, W. and Mikelić, A., Homogenization of the Laplace equation in a partially perforated domain, Berdichevsky, V. et al., World Scientific. Ser. Adv. Math. Appl. Sci. (1999)

[115] Jäger, W. and Mikelić, A., On the interface boundary condition of Beavers, Joseph, and Saffman, SIAM J. Appl. Math. (2000)

[116] Jäger, W. and Mikelić, A., On the roughness-induced effective boundary condition for an incompressible viscous flow, J. Diff. Equa. (2001)

[117] Jäger, W. and Mikelić, A., Couette flows over a rough boundary and drag reduction, Commun. Math. Phys. (2003)

[118] Jäger, W. and Mikelić, A. and Neuss, N., Asymptotic analysis of the laminar viscous flow over a porous bed, SIAM J. Sci. Comput. (2001)

[119] Jin, S., Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations, *SIAM J. Sci. Comput.* **21** (1999) pp. 441—454,

[120] Jin, S., Runge-Kutta Methods for Hyperbolic Conservation Laws with Stiff Relaxation Terms, J. Computational Physics, **122** (1995), 51-67.

[121] Jin, S. and Katsoulakis, M.A., Hyperbolic Systems with Supercharacteristic Relaxations and Roll Waves, SIAM J. Appl. Math. **61** (2000)

[122] Jin, S. and Levermore, C.D., The Discrete-Ordinate Method in Diffusive Regime, Transp. Theory Stat. Phys. 20 (1991), 413-439.

[123] Jin, S. and Levermore, C.D., Numerical schemes for hyperbolic conservation laws with stiff relaxation terms, J. Comput. Phys. 126 (1996), no. 2, 449–467.

[124] Jin, S., Pareschi, L. and Toscani, G., Diffusive Relaxation Schemes for Discrete-Velocity Kinetic Equations, SIAM J. Num. Anal. 35 (1998) 2405-2439

[125] Jin, S., Pareschi, L. and Toscani, G., Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM J. Numer. Anal.* **38** (2000), 913–936

[126] Jin, S. and Xin, Z., The relaxation schemes for systems of conservation laws in arbitrary space dimensions, Comm. Pure Appl. Math. (1995)

[127] Johnson, M., Noble, P. and Zumbrun, K., Nonlinear Stability of Viscous Roll Waves, SIAM J. Math. Anal. 43 (2011)

[128] Katsaounis T. and Makridakis C., Relaxation models and finite element schemes for the shallow water equations, Hyperbolic problems : Theory, Num., Appl. (2003)

[129] Kawashima, S. and Shizuta, Y., Systems of equations of hyperbolic-parabolic type with application to the discrete Boltzmann equation. *Hokkaido Math. J.* **14**, 435-457 (1984).

[130] Kim, Y.-H. and Kim, J.-E. and Yasushi, I. and Shih, A. M. and Brott, B. and Anayiotos, A., Hemodynamic Analysis of a Compliant Femoral Artery Bifurcation Model using a Fluid Structure Interaction Framework, Annals of Biomedical Engineering (2008)

[131] Kloeden, P.E., Global existence of classical solutions in the dissipative shallow water equations, SIAM J.Math. Anal. (1985)

[132] Kružkov, S. N., First order quasilinear equations with several independent variables, Mat. Sb. (N.S.) (1970)

[133] LeVeque, Randall J., Finite volume methods for hyperbolic problems, Cambridge University Press (2002)

[134] LeVeque, Randall J., Numerical methods for conservation laws, Birkhäuser Verlag (1992)

[135] Levermore, C.D. and Wagner, B.A. Robust fluid dynamical closures of the Broadwell model. *Phys. Lett. A.* **174**, 220–228 (1993).

[136] Lewandowski, R., Analyse Mathématique et Océanographie, Masson (1997)

[137] Lions, P.L., Mathematical Topics in Fluid Mechanics (1), Oxford University Press (1996)

[138] Lions, J.L., Temam, R. and Wang S., New formulation of the primitive equations of the atmosphere and applications, Nonlinearity (1992)

[139] Lions, J.L., Temam, R. and Wang S., On the equations of the large-scale ocean, Nonlinearity (1992)

[140] Lions, J.L., Temam, R. and Wang S., Mathematical study of the coupled models of atmosphere and ocean, J. Math. Pures Appl. (1995)

[141] Liou,T.-M., Liou, S.-N. and Chu, K.-L., Intra-aneurysmal flow with helix and mesh stent placement across side-wall aneurysm pore of a straight parent vessel, Journal of biomechanical engineering (2004)

[142] Liu, T.P., Hyperbolic conservation laws with relaxation, *Comm. Math. Phys.*, **1**, pp. 153–175, (1987),

[143] Marche, F., Derivation of a new two-dimensional viscous shallow water model with varying topography, bottom friction and capillary effects, European Journal of Mechanic (2007)

[144] Marche, F., Theoretical and numerical study of shallow water models ; applications to nearshore hydrodynamics, Phd thesis, Université Bordeaux 1 (2005)

[145] Mascia, C. and Natalini, R., On relaxation hyperbolic systems violating the Shizuta-Kawashima condition. *Ar. Rational Mech. Anal.* **41** (2010).

[146] Matsumura, A. and Nishida, T., The initial value problem for the equations of motion of viscous and heat-conductive gases, Journal of Mathematics of Kyoto University (1980)

[147] Milišić, V., Very weak estimates for a rough Poisson-Dirichlet problem with natural vertical boundary conditions, Methods and Applications of Analysis (2009)

[148] Milišić, V., Blood-flow modelling along and trough a braided multi-layer metallic stent, submitted (2010)

[149] Milišić, V., Homogenized stents for blood flow simulations in cardio-vascular junctions and aneurysms, In preparation

[150] Milišić, V., Pichon Gostaf, K. and Rambaud A., Asymptotic analysis of blood flow in stented arteries : time dependency and direct simulations, ESAIM proceedings (2011)

[151] Muñoz-Ruiz, M.L., On a nonhomogeneous bi-layer shallow water problem : smoothness and uniqueness results, Nonlinear Anal.(2003)

[152] Natalini, R., Convergence to equilibrium for the relaxation approximations of conservation laws, Comm. Pure Appl. Math., **8**, (1996)

[153] Natalini, R., A discrete kinetic approximation of entropy solutions to multidimensional scalar conservation laws J. of Diff. Eq., v. 148 (1998)

[154] Natalini, R., Recent results on hyperbolic relaxation problems. CRC Monogr. Surv. Pure Appl. Math. (1999)

[155] Nečas, J., Les méthodes directes en théorie des équations elliptiques, Masson et Cie, Éditeurs, Paris (1967)

[156] Neuss, N. and Neuss-Radu, M. and Mikelić, A., Effective laws for the Poisson equation on domains with curved oscillating boundaries, Applicable Analysis (2006)

[157] Orenga, P., Un théorème d'existence de solutions d'un problème de shallow water, Arch. Rational Mech. Anal. (1995)

[158] Pareschi, L. and Russo, G., Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation, J. Sci. Comp. **25**, (2005)

[159] Pedlosky, J., Geophysical fluid dynamics, Springer (1987)

[160] Perthame, B. and Simeoni, C., A kinetic scheme for the Saint-Venant system with a source term, Calcolo, 38 (2001)

[161] Peybernes, M., Analyse de problèmes mathématiques de la mécanique des fluides de type bicouche et à frontière libre, Phd Thesis (2006)

[162] Poiseuille, J.L.M., Recherche sur la force du coeur aortique, Phd thesis (1828)

[163] Proft, J., Multi-Algorithmic Numerical Strategies for the Solution of Shallow Water Models, PhD Thesis, TICAM Report (2002)

[164] Quarteroni, A. and Valli, A., Domain decomposition methods for partial differential equations, Oxford Science Publication, Numerical Mathematics and Scientific Computation (1999)

[165] Rambaud A., A dynamic multilayer shallow water model, Submitted (2011)

[166] Richardson, L. F., Weather prediction by numerical process, Cambridge University press, second edition (2007)

[167] Barré de Saint-Venant, A.-J.-C., Théorie du mouvement non permanent des eaux avec applications aux crues des rivières et à l'introduction des marées dans leur lit, C. R. Acad. Sci. (1871)

[168] Sanchez-Palencia, E. and Zaoui, A. , Homogenization techniques for composite media, Springer-Verlag (1987)

[169] Serre, D., Systèmes hyperboliques de lois de conservation, Parties I et II, Diderot, Paris (1996)

[170] Smith, Hal L., Systems of ordinary differential equations which generate an order preserving flow. A survey of results, SIAM Rev. (1988)

[171] Stoker J.J., Water waves, the mathematical theory with applications, Wiley (1958)

[172] Sundbye, L., Global existence for Dirichlet problem for the viscous shallow water equations, J. Math. Anal. Appl. (1996)

[173] Sundbye, L., Global existence for the Cauchy problem for the viscous shallow water equations, Rocky Mt. J. Math. (1998)

[174] Taylor, Michael E., Partial differential equations. III, Springer-Verlag (1997)

[175] Temam R., Navier-Stokes equations : theory and numerical analysis, AMS Chelsea Pub. (2001)

[176] Temam R. and Ziane M., Some mathematical problems in geophysical fluid dynamics, Handbook of math. fluid dyn., North-Holland (2004)

[177] Villani, C., Handbook of Mathematical Fluid Dynamics, Chapter 2, A review of mathematical topics in collisional kinetic theory, North-Holland (2002)

[178] Wang, W. and Xu, C., The Cauchy problem for viscous shallow water equations, Rev. Mat. Iberoamericana (2005)

[179] Whitham G. B., Linear and nonlinear waves, John Wiley and Sons Inc., New York (1999)

[180] Wiener, N., The homogeneous chaos, Am. J. Math. (1938)

[181] Wu, G. and Zhang, B., Local well-posedness of the viscous rotating shallow water equations with a term of capillarity, ZAMM Z. Angew. Math. Mech. (2010)